SCHOLARLY PUBLICATIONS

A CURRENT AWARENESS BULLETIN
OF RESEARCH OUTPUT

@DTU

(122nd Edition)

FEBUARY 2023

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 Twenty Second** 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 **February, 2023**. 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 Grounded Flux Controlled Incremental/Decremental Memristor Emulator, ***6.Navnit Kumar***, ***3.Manjeet Kumar*** and ***3.Neeta Pandey***, Electronics, DTU

2. A low power single bit-line configuration dependent 7T SRAM bit cell with process-variation-tolerant enhanced read performance, ***6.Bhawna Rawat*** and ***3.Poornima Mittal***, Electronics, DTU

3. A Modified Bypass Circuit for Improved Reliability of PV Module Validated with Real-Time Data, ***6.Kashika Baranwal***, ***3.Prem Prakash*** and ***3.Vinod Kumar Yadav***, Electrical, DTU

4. A Novel Cost-efficient Parallel Binary Adder without Crossover in QCA with Energy Dissipation Analysis, ***6.Hemanshi Chugh*** and ***3.Sonal Singh***, ECE, DTU

5. A review of dissimilar aluminum alloys welding via Friction stir welding, ***6.Sumit Jain*** and ***3.R.S. Mishra***, Mechanical, DTU

6. A Survey on Quantum Computing for Internet of Things Security, ***6.Diksha Chawla*** and ***3.Pawan Singh Mehra***, CSE, DTU

7. A two stream face anti-spoofing framework using multi-level deep features and ELBP features, ***6.Aashania Antil*** and ***3.Chhavi Dhiman***, Electronics, DTU

8. Analyzing the barriers for aquaponics adoption using integrated BWM and fuzzy DEMATEL approach in Indian context, ***3.Girish Kumar***, Ram C. Bhujel, ***8.Aniket Aggarwal***, ***8.Divyansh Gupta***, ***8.Ashish Yadav*** and Mohammad Asjad, Mechanical, DTU

9. Blockchain-Based Secure Localization against Malicious Nodes in IoT-Based Wireless Sensor Networks Using Federated Learning, ***6.Gebrekiros Gebreyesus Gebremariam***, ***3.J. Panda*** and ***3.S. Indu***, Electronics, DTU

10. Can Gratitude Counter Workplace Toxicity? Exploring the Mediating Role of Psychological Capital (PsyCap), Ayatakshee Sarkar, ***3.Naval Garg***, DK Srivastava and BK Punia, DSM, DTU

11. Carboxy-methyl tamarind kernel gum based bio-hydrogel for sustainable agronomy, ***6.Ritu Malik***, ***3.Sudhir G. Warkar*** and Reena Saxena, Applied Chemistry, DTU

12. Constraints on the Parameterized Deceleration Parameter in FRW Universe, Amine Bouali, ***7.Himanshu Chaudhary***, Ujjal Debnath,, Tanusree Roy and G.Mustafa, Applied Mathematics, DTU

13. Conv-CapsNet: capsule based network for COVID-19 detection through X-Ray scans, ***8.Pulkit Sharma***, ***8.Rhythm Arya***, ***8.Richa Verma*** and ***3.Bindu Verma***, IT, DTU

14. Deep Convolutional Architectures for Uncertainty Quantification and Forecast in Inundation Problems, Azzeddine Soulaïmani, Azzedine Abdedou, ***8.Yash Kumar***, and ***8.Pratyush Bhatt***, DTU

15. Dependence of Matrix Elements on Hydrostatic Pressure and Temperature, ***6.Mona***, ***6.Aakash***, ***6.Priyanka*** and ***3.Rinku Sharma***, Applied Physics, DTU

16. Design and Evaluation of aWater-Based, Semitransparent Photovoltaic Thermal TrombeWall, ***8.Sheel Bhadra***, Niloy Sen, Akshay K K, Harmeet Singh and Paul G. O'Brien, Mechanical, DTU

17. Design and Modal Analysis of Optical Fibers with Multiple Cores and Multiple Cladding Fiber, ***6.Ankit Kumar***, Electroncis, DTU

18. Design and Simulation of the Spiral Micromixer with Chaotic Advection, ***6.R. Prakash***, ***3.M. Zunaid*** and ***3.1.Samsher***, Mechanical, DTU

19. Does workplace toxicity influence turnover intentions among Indian healthcare employees? Investigating the moderating role of gratitude, ***3.Naval Garg***, Manju Mahipalan and Nidhi Sharma, USME, DTU

20. ECG Signals- Early detection of Arrhythmia using Machine Learning approaches, ***8.Jahnavi Nandanwar***, ***8.Jasmeet Singh*** & ***3.Sanjay Patidar***, Software Engineering, DTU

21. Effect of fiber synergy on strength & durability of High-Performance Concrete, ***6.Mohit Gupta***, ***3.Dr. Ritu Raj***, ***3.Dr. Anil Kumar Sahu*** & Dr. Sapna Yadav, Civil, DTU

22. Effect of orange peel oil on the performance and emission characteristics of diesel engine fueled with quaternary blends, ***7.Dushyant Mishra***, ***3.Naveen Kumar*** and ***3.Rajiv Chaudhary***, Mechanical, DTU

23. Emotion Recognition in Speech by Multimodal Analysis of Audio and Text, ***8.Siddhant Bikram Shah***, Shubham Garg and Aikaterini Bourazeri, CSE, DTU

24. Empirical Review of Various Thermography-Based Computer-Aided Diagnostic Systems for Multiple Diseases, ***3.TRASHA GUPTA***, ***3.RAJNI JINDAL*** and ***3.INDU SREEDEVI***, Electronics, DTU

25. Energy Efficient Multiple Visible Light Communication Access Points (VAP), ***8.Vishal Kumar*** and ***3.Rohit Kumar,*** Applied Mathematics, DTU

26. Estimating Medical Insurance Cost using Linear Regression with HyperParameterization, Decision Tree and Random Forest Models, ***3.Sanjay Patidar***, ***8.Saksham Dudi*** and ***8.Rohit***, Software Engineering, DTU

27. Experimental approach and optimization of hot air drying tomato waste using response surface methodology, Djamel Mennouche, Tarik Hadibi, Abdelghani Boubekri, Amina Boudouaya, Abdellouahab Benseddik, Müslüm Arıcı, ***3.Anil Kumar*** and Wang Yunfeng, Mechanical, DTU

28. FedEmo: A Privacy-Preserving Framework for Emotion Recognition using EEG Physiological Data, ***8.Mohd Ayaan Anwar***, ***8.Manan Agrawal***, ***6.Neha Gahlan***, ***3.Divyashikha Sethia***, ***8.Gaurav Kumar Singh*** and ***8.Rishabh Chaurasia***, CSE and Software Engineering, DTU

29. Flexible Organic Solar Cell to Power Modern Cardiac Pacemakers: Versatile for All Age Groups, Skin Types and Genders, ***6.Paritosh Chamola*** and ***3.Poornima Mittal***, Electronics, DTU

30. Fourier Analysis Based Respiration Rate Estimation Using Corrupted Photoplethysmogram Signal, Pankaj, Ashish Kumar, Rama Komaragiri and ***3.Manjeet Kumar***, Electronics, DTU

31. Impact of Competition in Credit Rating Industry: Evidence From India, ***3.Chandan Sharma***, ***3.Archana Singh***, and ***3.Rajan Yadav***, DSM, DTU

32. Impact of Word Embedding Methods on Software Vulnerability Severity Prediction Models, ***3.Ruchika Malhotra*** and ***6.Vidushi***, Software Engineering, DTU

33. Influence of SiC microparticles and multi-pass FSW on weld quality of the AA6082 and AA5083 dissimilar joints, ***6.SUMIT JAIN***, ***3.R. S. MISHRA*** and HUSAIN MEHDI, Mechanical, DTU

34. Integration of Pre-Treatment with UF/RO Membrane Process for WasteWater Recovery and Reuse in Agro-Based Pulp and Paper Industry, ***6.Sumit Dagar***, ***3.Santosh Kumar Singh*** and Manoj Kumar Gupta, Environmental, DTU

35. Intelligent Model for Smartphone Addiction Assessment in University Students using Android Application and Smartphone Addiction Scale, ***7.1.Anshika Arora***, Pinaki Chakraborty, M.P.S. Bhatia and Aditya Puri, Software Engineering, DTU

36. Intrusion detection based on machine learning and statistical feature ranking techniques, ***8.Akhil Kumar*** and ***3.Shailender Kumar***, CSE, DTU

37. LDML: A Proposal to Reduce Leakage Power in DML Circuits, ***6.Neetika Yadav***, ***3.Neeta Pandey*** and ***3.Deva Nand***, Electronics, DTU

38. Localization and turbulence of beam-driven whistler wave with magnetosonic wave in magnetopause, ***6.Jyoti***, ***3.Suresh C. Sharma***, and R. P. Sharma, A Physics, DTU

39. Low Power, Highly Stable and Enhanced Read Speed 7T SRAM, ***7.Deberjeet Usham*** and ***3.Malti Bansal***, Electroncis, DTU

40. Machine Learning Based Obfuscated Malware Detection in the Cloud Environment with Nature-Inspired Feature Selection, ***6.Mohd. Rehan Ghazi*** and ***3.N. S. Raghava***, Electronics, DTU

41. Measuring circularity of a manufacturing organization by using sustainable balanced scorecard, ***6.Abhishek Sahu***, ***3.Saurabh Agrawal*** and Chandra Prakash Garg, Mechanical and DSM, DTU

42. Mechanical and Erosion Characterization of Untreated and Solution-Treated Nitrogen-Alloyed (23-8N) Austenitic Stainless Steel, ***6.Mohit Vishnoi***, ***3.Qasim Murtaza***, and ***3.Paras Kumar***, Mechanical, DTU

43. mlpack 4: a fast, header-only C++ machine learning library, Ryan R. Curtin, Marcus Edel, Omar Shrit, Shubham Agrawal, Suryoday Basak, James J. Balamuta, Ryan Birmingham, ***8.Kartik Dutt***, Dirk Eddelbuettel, Rishabh Garg , Shikhar Jaiswal, Aakash Kaushik, Sangyeon Kim, Anjishnu Mukherjee, Nanubala Gnana Sai, Nippun Sharma, Yashwant Singh Parihar, Roshan Swain and Conrad Sanderson, DTU

44. Multiple Feature Selection Frameworks based on Evolutionary Computing and Ensemble Learning for Software Defect Prediction, ***3.Ruchika Malhotra***, ***8.Anmol Budhiraja***, ***8.Abhinav Kumar Singh***, ***8.Ishani Ghoshal*** and ***3.Shweta Meena***, Software Engineering, DTU

45. Nano ZnO embedded poly (ethylene glycol) diacrylate cross-linked carboxymethyl tamarind kernel gum (CMTKG)/poly (sodium acrylate) composite hydrogels for oral delivery of ciprofloxacin drug and their antibacterial properties, ***6.Indu Rani***, ***3.Sudhir G. Warkar***, ***3.Anil Kumar***, Applied Chemistry, DTU

46. ON A CLASS OF NON-UNIVALENT FUNCTIONS ASSOCIATED WITH A PARABOLIC REGION, ***7.MRIDULA MUNDALIA*** AND ***3.S. SIVAPRASAD KUMAR***, MATHEMATICS, DTU

47. Optimization and manufacturing of friction drilling tools by using solid works and Ansys, ***7.Ebrahim Mohibi*** and ***3.Dr. A.K Madan***, Mechanical, DTU

48. Performance and emission characteristics analysis of LPG-Karanja biodiesel on CI engine with optimization, ***8.Sheel Bhadra***, ***7.Hansham Dewal*** and ***3.Naveen Kumar***, Mechanical, DTU

49. Performance Assessment and Modeling Techniques for Domestic Solar Dryers, Shimpy, ***8.Mahesh Kumar*** and ***3.Anil Kumar***, Mechanical and Environment, DTU

50. Recent Prospects of Carbonaceous Nanomaterials-Based Laccase Biosensor for Electrochemical Detection of Phenolic Compounds, ***7.Sakshi Verma***, ***6.Deeksha Thakur***, Chandra Mouli Pandey and ***3.Devendra Kumar***, Applied Chemistry, DTU

51. Role of nuclear energy in carbon mitigation to achieve United Nations net zero carbon emission: evidence from Fourier bootstrap Toda-Yamamoto, ***6.Akanksha Singh***, Shiv Lal, ***3.Nand Kumar***, ***3.Rajan Yadav*** and ***6.Shweta Kumari***, Humanities and DSM, DTU

52. Running vacuum in Brans & Dicke theory: a possible cure for the 8 and H0 tensions, Javier de Cruz Perez2;1 Joan Sola Peracaula and ***3.C. P. Singh***, Mathematics, DTU

53. Selectively probing ferric ions in aqueous environments using protonated and neutral forms of 7-azaindole as a multiparametric chemosensor, ***3.Mohan Singh Mehata*** and ***6.Aneesha***, Applied Physics, DTU

54. Snail: An Eco-Friendly Autonomous Steam Mopping Robot for Cleaning and Disinfection of Floors, Sathian Pookkuttath, Mohan Rajesh Elara, ***6.Madan Mohan Rayguru***, Zaki Saptari Saldi, Vinu Sivanantham and Balakrishnan Ramalingam, Electrical, DTU

55. Strategic Energy Saving Technique for Multiple Visible Light Communication Access Points (VAP), ***6.Vishal Kumar***, ***3.Rohit Kumar*** and ***6.Himani Mehta***, DTU

56. Study of the Energy Dissipation over the Type-A Piano Key Weir, ***6.Deepak Singh*** and ***3.Munendra Kumar***, Civil, DTU

57. Terahertz Wave Propagation Characteristics in Graded Teflon Based Solid-Core Photonic Crystal Fibre, ***8.Shikhar Dash***, ***8.Varun Mathur***, ***6.Nilesh Pandey*** and ***3.Ravindra K. Sinha***, Applied Physics, DTU

58. Topological Nature of Large Bulk Band Gap materials Sr3Bi2 and Ca3Bi2, Sanjeev, ***3.Mukhtiyar Singh***, Ramesh Kumar, Sunita Srivastava and Tankeshwar Kumar, Applied Physics, DTU

59. Trust factor-based analysis of user behavior using sequential pattern mining for detecting intrusive transactions in databases, ***3.Indu Singh*** and ***3.Rajni Jindal***, CSE, DTU

60. Viscous cosmology in holographic dark energy with Granda–Oliveros cut-off, ***6.Simran Kaur*** and ***3.C P Singh***, Applied Mathematics, 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*

2022 5th International Conference on Multimedia, Signal Processing and Communication Technologies (IMPACT) | 978-1-6654-7647-8/22/$31.00 ©2022 IEEE | DOI: 10.1109/IMPACT55510.2022.10029225

A Grounded Flux Controlled Incremental/Decremental Memristor Emulator

Navnit Kumar
Department of Electronics and Communication Engineering, Delhi Technological University, Delhi, India.
navnitnitmanipur@gmail.com

Manjeet Kumar
Department of Electronics and Communication Engineering, Delhi Technological University, Delhi, India.
manjeetchhillar@gmail.com

Neeta Pandey
Department of Electronics and Communication Engineering, Delhi Technological University, Delhi, India.
n66pandey@rediffmail.com

Abstract - Single Inverting Current Conveyor Transconductance Amplifier (ICCTA) based flux controlled grounded incremental/decremental memristor emulator is presented in this article. The proposed circuit uses one ICCTA, one grounded resistor, one grounded capacitor. This realization is appropriate for monolithic integration due to the usage of grounded passive devices. Analog multiplier circuit is not required to design the circuit. Additionally, a high pass filter circuit demonstrates how well the suggested memristor emulator works. Using a 180 nm CMOS process, the suggested circuit and its utility are assessed.

Keywords- Memristor Emulator, Mem-immittance Circuit, Pinch Hysteresis loop, ICCTA.

I. INTRODUCTION

The fourth [1-2] fundamental nonlinear element known as memristor is gaining considerable attention among the researchers due to its high-density storage property. Leon O. Chua developed memristor by giving the relation between electric charge and magnetic flux in 1971. Since memristor has nonlinear characteristic, non-volatile nature, and high density storge properties, it is being used in a variety of applications such as programmable gate arrays [3], programmable analogue circuits [4], neuromorphic circuits [5], adaptive learning [6], adaptive filters [7], chaotic oscillators [8], and relaxation oscillators [9-10] etc. After 37 years of proposal of memristor by L O. Chua, Hewlett-Packard (H. P.) lab created the first solid-state memristor in 2008. Due to its high price and difficult manufacturing process, this solid state memristor is not currently accessible on the market. This constraint spurs researchers to create memristor circuits based on active building blocks (ABB). Below is a literature review of an ABB-based memristor emulator.

There are few publications on the realisation of charge and flux controlled memristor circuit, using advance active blocks such as Second Generation Current Conveyor (CCII) [11-14], Operational Transconductance Amplifier (OTA) [15], Current Conveyor Transconductance Amplifier (CCTA) [16], Differential Voltage Current Conveyor Transconductance Amplifier (DVCCTA) [17], Differential Difference Current Conveyor (DDCC) [18], Voltage Difference Transconductance Amplifier (VDTA) [19,20], Voltage Differencing Buffered Amplifier (VDBA) [21], Voltage Differencing Current Conveyor (VDCC) [22], Current Backward Transconductance Amplifier (CBTA) [24]. The reported flux and charge controlled memristor emulator suffer from the following disadvantages:

(1) Memristor emulator reported in [8,9,11,12,13,14,15,26] need more than one active block.
(2) Memristor emulator introduced in [8,9,11-14,16-18,20,25,26] required more than two passive components.
(3) Memristor emulator designed in [8,9,11,12,13,14,16-18,26] need floating passive component.
(4) Memristor emulator presented in [4,8,9,13,14,18,20,24,26] use additional circuit like analog multiplier, Digital to analog converter and analog to digital converter.
(5) Memristor emulator reported in [8,9,11-15,26] operates in low frequency range.
(6) Memristor emulator illustrated in [25] uses passive inductor.

This article presents flux controlled grounded incremental/decremental memristor emulator. The proposed circuit uses one ICCTA as active components, one grounded capacitor and one grounded passive resistor as passive components. The suggested circuit can be designed without the use of additional circuitry, such as an analog multiplier circuit. In addition, Memristor based first order high pass filter circuit shows the usefulness of proposed circuit. Remaining paper is organized as follow: Circuit description of proposed circuit is discussed in section2. Simulation results of proposed memristor is discussed in section 3. The suggested circuit's application is shown in Section 4. The research article is concluded in section 5.

II. CIRCUIT DESCRIPTION

An ICCTA is a versatile active building block. The block symbol and its MOS based block realization is illustrated in Fig.1. The port relation of ICCTA is given in equation (1).

$$I_Y = 0;\ V_{X-} = -V_Y;\ I_{Z-} = I_{X-};\\ I_{0+} = g_m V_{z-};\ I_{0-} = -g_m V_{z-} \quad (1)$$

978-1-6654-7647-8/22/$31.00 ©2022 IEEE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on February 06,2023 at 05:20:01 UTC from IEEE Xplore. Restrictions apply.

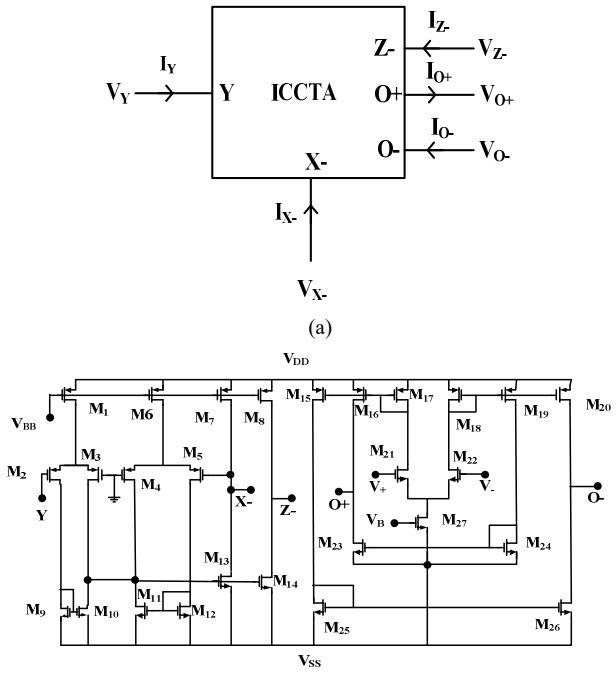


Fig. 1. ICCTA (a) Block symbol (b) MOS based implementation of ICCTA is derived from ref. [27].

A flux controlled grounded incremental/decremental memristor emulator using ICCTA illustrated in Fig.2. This circuit includes one active building block (ICCTA), one grounded resistor and one grounded capacitor.

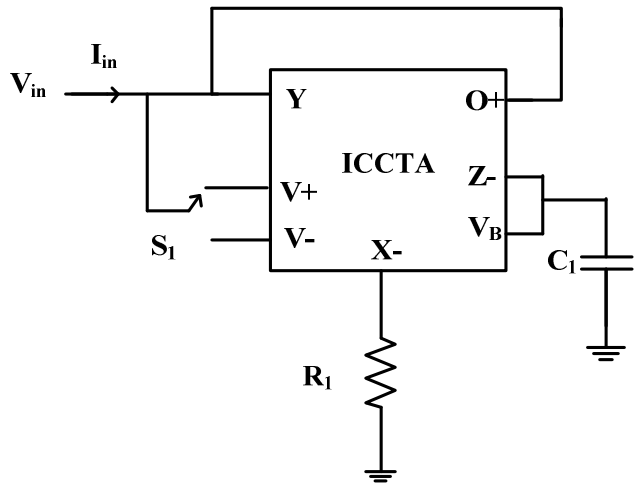


Fig. 2. Proposed memristor emulator.

Current flows through the $X-$ terminal of designed circuit can be written as:

$$I_{X-} = \frac{V_{in}}{R_1} \tag{2}$$

Voltage across capacitor connected at Z- terminal of proposed circuit may be obtained as:

$$V_C = -\frac{1}{C_1}\int \frac{V_{in}}{R_1} dt = -\frac{\varphi_{in}}{R_1 C_1} \tag{3}$$

Now, input current of memristor circuit may be obtained using equation (3) as

$$I_{in} = \mu_n C_{ox} \frac{W}{\sqrt{2}L}\left(-\frac{\varphi_{in}}{R_1 C_1} - V_{SS} - V_{th}\right) V_{in} \tag{4}$$

Therefore, memconductance of proposed memristor emulator in decremental mode (switch S_1 is connected to V_+ terminal) may written as:

$$W(\varphi(t)) = \frac{I_{in}(t)}{V_{in}(t)} = \mu_n C_{ox} \frac{W}{\sqrt{2}L}\left(-\frac{\varphi_{in}}{R_1 C_1} - V_{SS} - V_{th}\right) \tag{5}$$

Now, memconductance of proposed memristor emulator in incremental mode (switch S_1 is connected to V_- terminal) may written as:

$$W(\varphi(t)) = \frac{I_{in}(t)}{V_{in}(t)} = \mu_n C_{ox} \frac{W}{\sqrt{2}L}\left(+\frac{\varphi_{in}}{R_1 C_1} + V_{SS} + V_{th}\right) \tag{6}$$

A voltage source of Vin=$A_m sin(2\pi f_m t)$ is connected at input terminal to get frequency response of the designed memristor emulator. Where, A_m and f_m are known as amplitude and frequency. Thus, the memconductance of proposed circuit may be obtained as:

$$W(\varphi(t)) = \frac{I_{in}(t)}{V_{in}(t)} = \pm\frac{k}{\sqrt{2}}(V_{ss} + V_T) \pm k\frac{A_m cos(\omega t - \pi)}{2\sqrt{2} f C_1 R_1} \tag{7}$$

Where $k = \mu_n C_{ox}\frac{W}{L}$. Here, μ_n, C_{ox} *and* $\frac{W}{L}$ have their usual meaning. Equation (7) shows that the memconductance of the proposed emulator may be changed by varying the amplitude and frequency of the input signal.

III. Simulation Results

Theoretical analysis of proposed flux controlled memristor emulator as stated in section 2, are validated by PSPICE simulation using 180 nm TSMC technology. Supply voltage and bias voltage (V_{BB}) of ICCTA are ±1.25 V and 0.42 V respectively. Table 1 provides size of MOS transistors of ICCTA.

TABLE I. ASPECT RATIOS OF MOS TRANSISTORS

Transistor	W(µm)/ L(µm)
M_1, M_6 - M_8	2.4/0.18
M_2	0.35/0.18
M_3 - M_5	0.72/0.18
M_{11} – M_{13}	1.2/0.18
M_{15} – M_{20}	2.15/0.7
M_{21}, M_{22}	5.75/0.7
M_{23} – M_{26}	1.45/0.7

The value of resistance and capacitance is R1=3 kΩ and C1=100 pF respectively, are selected to operate at 1 MHz frequency to demonstrate the characteristics of incremental or decremental memristor emulator. Pinch hysteresis curve and transient response of decremental memristor emulator is

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on February 06,2023 at 05:20:01 UTC from IEEE Xplore. Restrictions apply.

illustrated in Fig. 3. Transient response of decremental memristor shows 0° phase difference between input current and voltage. Pinch hysteresis curve and transient response of proposed incremental memristor emulator is depicted in Fig. 4. Transient response of incremental memristor emulator shows 180° phase difference between input voltage and input current. Moreover, Fig. 5 shows pinch hysteresis curve of grounded incremental and decremental memristor emulator at different capacitance value of 100 pF, 200 pF, and 300 pF while keeping frequency f=1 MHz and amplitude of input signal Vin=1 V constant. Fig.4 shows that lobe area of pinch hysteresis curve is decreased as value of capacitance is increased. Comparison between active block based published memristor circuit and proposed memristor emulator is mentioned in Table II.

TABLE II. COMPARISON OF ACTIVE BLOCK BASED PUBLISHED MEMRISTOR CIRCUIT WITH PROPOSED MEMRISTOR EMULATOR

Reference	No. of Active Block	No. of Resistors(R) / Capacitors (C)	Parameter	Power supply	Floating/ Grounded	Tunability	Additional Device
[11]	CCII-4, OTA-3	6/1	NA	±15 V	Floating	YES	-
[12]	CFOA-4	4/4	NA	NA	Floating	NO	Diode-2
[13]	CFOA-2	4/3	NA	±10 V	Grounded	NO	Multiplier-1
[14]	CCII-4	5/1	NA	±10 V	Floating	NO	Multiplier-1
[16]	CCTA-1	3/1	180 nm	±1.5 V	Floating	YES	-
[17]	DVCCTA-1	3/1	250 nm	±1.25 V	Grounded	YES	-
[18]	DDCC-1	2/1	250 nm	±1.5 V	Floating	NO	Multiplier-1
[20]	VDTA-1	2/1	180 nm	±0.9 V	Floating	YES	Multiplier-1
[22]	VDCC-1	2/1	180 nm	±0.9 V	Grounded	YES	-
[24]	CBTA-1	2/1	180 nm	±0.9 V	Grounded	YES	Multiplier-1
[25]	VIM-1	2/1	180 nm	±1.65 V	Grounded	NO	Diode-1, Inductor-1
[26]	Op-amp-2, CFOA-2	7/1	NA	NA	Floating	NO	Multiplier-1
Proposed	ICCTA-1	1/1	180 nm	±1.2 V	Grounded	YES	-

NA: NOT AVAILABLE

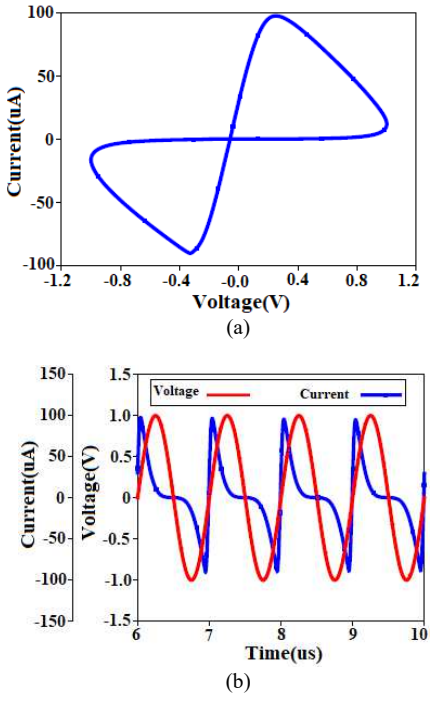


Fig. 3. Characteristic of proposed decremental memristor emulator. (a) Pinch hysteresis curve and (b)Transient response.

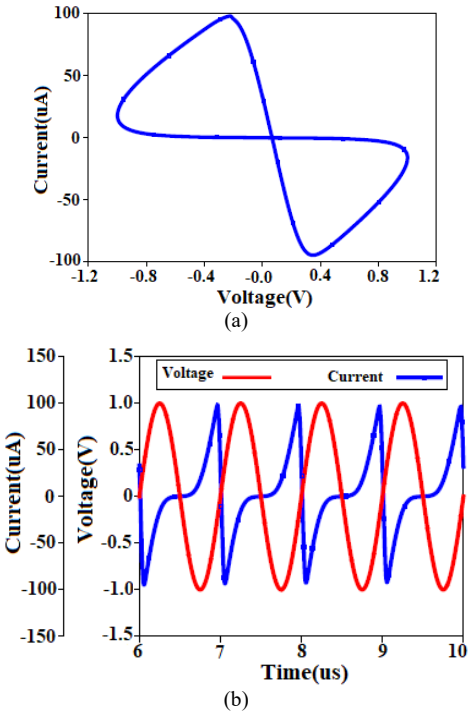
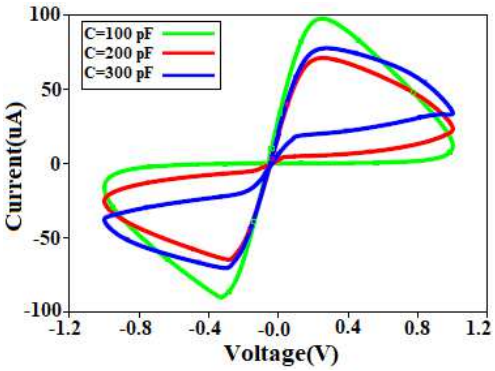
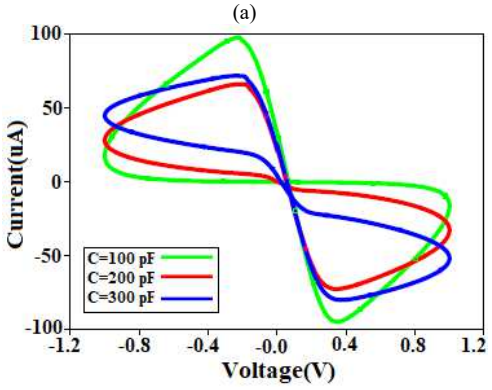


Fig. 4. Characteristic of proposed incremental memristor emulator. (a) Pinch hysteresis curve and (b)Transient response

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on February 06,2023 at 05:20:01 UTC from IEEE Xplore. Restrictions apply.



(a)

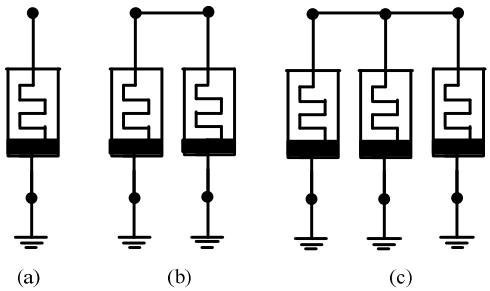


(b)

Fig. 5. Pinch hysteresis curve of proposed memristor emulator for C=100 pF to 300 pF. (a) Decremental and (b) Incremental.

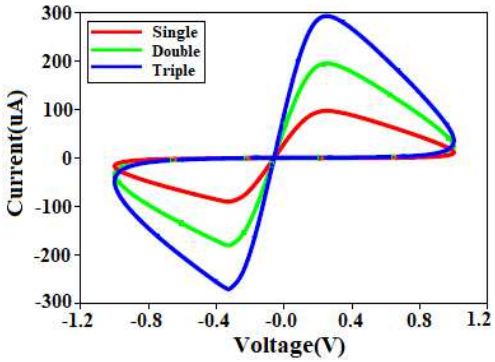
IV. APPLICATION

Workability of proposed grounded memristor emulator can be justified through different configurations of connection of proposed emulator. Parallel connection of proposed decremental memristor circuit are depicted in Figs. 6. Fig. 7 shows that pinch hysteresis curve of parallel connection of decremental memristor emulator. Since memristanc of parallel combination of two memristor and memristance of parallel combination of three memristor are half and one third memristance of the single memristor respectively. Hence Fig. 7 shows that memristor current of single configuration, parallel combination of two memristor and parallel combination of three memristor are 100μA, 200μA, and 300μA respectively.

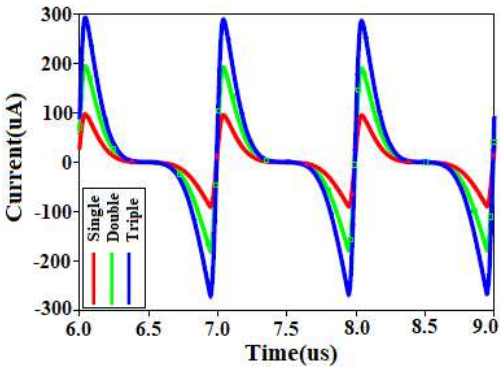


(a) (b) (c)

Fig. 6. Memristor connected in different configurations (a) Single, (b) Double and (c) Triple.



(a)



(b)

Fig. 7. Pinch hysteresis curve of proposed decremental memristor emulator for (a) Single(b) double and (c) Triple.

To demonstrate the viability of incremental and decremental memristor circuits, a first order high pass filter may be employed. The first order RC high pass filter is shown in Fig8(a). Incremental and decremental memristor based first order high pass filter is shown in Fig. 8(b) and 8(c). This figure is obtained by replacing resistance with proposed memristor in RC high pass filter circuit. Equations (8) and (9) can be used to express the cut off frequency of a memristor-based high pass filter and an RC-based high pass filter circuit, respectively.

$$f = \frac{1}{2\pi C\{R_{AVG} \pm \Delta R_m sin(2\pi f_m t + \varphi)\}} \quad (8)$$

$$f = \frac{1}{2\pi CR} \quad (9)$$

Where, R_{AVG} and ΔR_m is known as average value of memristance and variation of memristance due to applied sinusoidal signal respectively. Resistance and capacitance value of RC high pass filter are 2 kΩ, and 1 pF respectively. Components selection for high pass filter using memristor are R_1=3 kΩ, C_1=100 pF and C=1 pF. Magnitude response of voltage gain of RC high pass filter and high pass filter using memristor are illustrated in Fig.9. Cut off frequency of incremental memristor based high pass filter, RC based high pass filter and decremental memristor based high pass filter are 33.79 MHz, 79.58 MHz, and 165.92 MHz respectively.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on February 06,2023 at 05:20:01 UTC from IEEE Xplore. Restrictions apply.

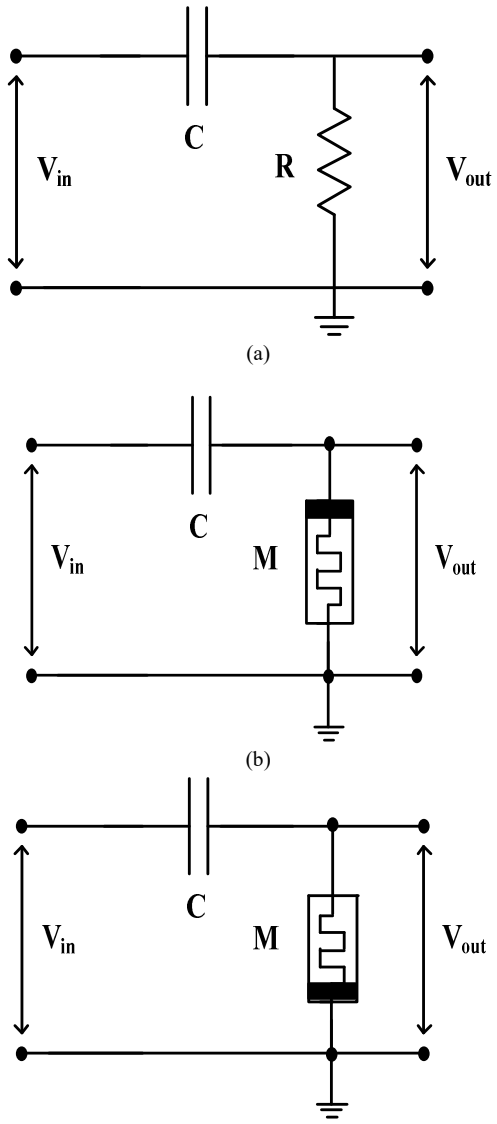


Fig. 8. (a) RC based high pass filter (b)Incremental memristor based high pass filter and (c) Decremental memristor based high pass filter.

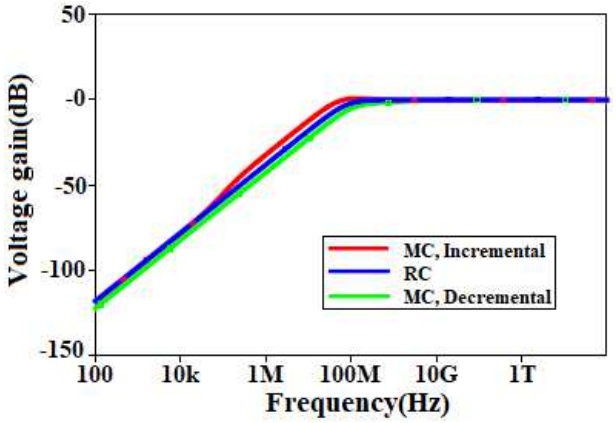


Fig. 9. Magnitude response of RC and MC based high pass filter circuit.

V. Conclusion

This article describes a flux-controlled grounded incremental/decremental memristor circuit. This circuit uses one ICCTA, one grounded passive resistor and one grounded capacitor. The designed circuit is free from additional circuit such as analog multiplier circuit, ADC circuit and DAC circuit. In addition, proposed circuit is verified in a high pass filter circuit. Cut off frequency of first order high pass filter based on memristor circuit can be tuned without disturbing memristor emulator.

References

[1] J. M. Tour, and T. He, "Electronics: The fourth element," Nature, vol. 453, pp. 42-43, 2008.

[2] L. O. Chua, "Memristor: the missing circuit element," IEEE Transactions on Circuit Theory, vol. 18, no. 5, pp. 507-512, 1971.

[3] H. A. F. Almurib, T.N. Kumar, and F. Lombardi, "Design and evaluation of a memristor-based look-up table for non-volatile field programmable gate arrays," IET Circuits, Device System, vol. 10, no.4, pp.292-300, 2016.

[4] Y.V. Pershin, and M. Di Ventra, "Practical approach to programable analog circuits with memristors," IEEE Trans. Circuits System I, Reg. paper, vol. 57, no. 8, pp. 1857-1864, 2010.

[5] G. C. Adam, B. D. Hoskins, M. Prezioso, F. Merrikh-Bayat, B. Chakrabarti, and D. B. Strukov, "3-D memristor crossbars for analog and neuromorphic computing applications," IEEE Transaction Electron Devices, vol.64, no.1, pp.312-318, 2017.

[6] F. Z. Wang, L. O. Chua, X. Yang, N. Healian, R. Tetzlaff et. al, "Adaptive Neuromorphic Architecture," Neural Networks, vol. 45, pp. 111-116, 2013.

[7] T. Driscoll et al., "Memristive adaptive filters," Applied phys. Letter, vol.97, no. 9, p. 093502, 2010.

[8] M.E.Sahin, A. S. Demirkol, H. Guler and S. E. Hamamci, "Design of a hyperchaotic memristive circuit based on wien bridge oscillator," Computers and Electrical Engineering, 88 (2020) 106826.

[9] D. Yu, H. H. C. Iu, "A Floating Memristor Emulator Based Relaxation Oscillator," IEEE Transaction on Circuits and System I: Regular Papers, vol. 61, no. 10, pp.2888-2896, Oct. 2014.

[10] M. E. Fouda, and A. G. Radwan, "Memristor-based voltage-controlled relaxation oscillators," International journal of circuit theory and applications, vol. 42, no. 10, pp. 1092-1102, 2014.

[11] H. Sozen and U. Cam, "Electronically tunable memristor emulator circuit," Analog Integrated Circuits Signal Processor, vol. 89, no. 3, pp. 655-663, 2016.

[12] M. T. Abuelama'atti and Z. J. Khalifa, "A new floating memristor emulator and its application in frequency-to-voltage conversion," Analog Integrated Circuit and Signal Processing. Vol. 86, pp. 141-147, 2016.

[13] C. S. Lopez, M. A. C. Aguilar and C. M. Montero, "A 16 Hz-160kHz memristor emulator circuit," International Journal of Electronics and Communications, vol. 69, no. 9, pp. 1208-1219, 2015.

[14] C. S. Lopez, M. A. C. Aguilar and C. M. Montero, "A Floating Analog Memristor Emulator Circuit," IEEE Transactions on Circuits and System II, vol. 61, no. 5, pp. 309-313, May 2014.

[15] G. Kanyal, P. Kumar, S. K. Paul, and A. Kumar, "OTA based high frequency tunable resistorless grounded and floating memristor emulators," International Journal of Electronics and Communications, vol. 92, pp. 124-145, Aug. 2018.

[16] R. K. Ranjan, N. Rani, R. Pal, S. K. Paul, and G. Kanyal, "Single CCTA based high frequency floating and grounded type of incremental/decremental memristor emulator and its application," Microelectronics Journal, vol. 60, pp. 119-128, Feb. 2017.

[17] R.K. Ranjan, N. Bhuwal, N. Raj and F. Khateb, "Single DVCCTA based High frequency incremental/decremental Memristor Emulator and its

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on February 06,2023 at 05:20:01 UTC from IEEE Xplore. Restrictions apply.

Application," International Journal of Electronics and Communications, vol. 82, pp. 177-190, Dec.2017.

[18] A. Yesil, Y. Babacan and F. Kacar, "A new DDCC based memristor emulator circuit and its applications," Microelectronics Journal, vol. 45, no. 3, pp. 282-287, 2014.

[19] J. Vista and A. Ranjan, "Flux Controlled Floating Memristor Employing VDTA: Incremental or Decremental Operation," IEEE Transactions on Computer Aided Design of Integrated Circuits and System, vol. 40, no. 2, pp. 364-372, Feb. 2021.

[20] P. B. Petrovic, "Floating incremental/decremental flux-controlled memristor emulator circuit based on single VDTA," Analog integrated circuits and signal processing, no. 96, pp. 417-433, March 2018.

[21] N. Yadav, S. K. Rai, and R. Pandey, "Novel Memristor using Fully Balanced VDBA and Grounded Capacitor," Iranian Journal of Science and Technology Transactions of Electrical Engineering, vol. 45, pp. 229-245, June 2021.

[22] A. Yesil, Y. Babacan, and F. Kacar, "Electronically tunable memristor based on VDCC," AEU-International Journal of Electronics and Communications, vol. 107, pp. 282-290, 2019.

[23] N. Raj, R. K. Ranjan, and F. Khateb, "Flux-controlled Memristor Emulator and its Experimental result," IEEE Transactions on Very Large Scale Integration (VLSI) System, vol. 28, no. 4, pp. 1050-1061, April 2020

[24] U. E. Ayten, S. Minaei and M. Sagbas, "Memristor emulator circuits using single CBTA," International Journal of Electronics and Communications, vol. 82, pp. 109-118, 2017.

[25] R Arundeepakvel, P. Khatter, N. Pandey, and S. Minaei, "A novel design for voltage inverting metamutator and its applications," Microelectronics Journal, vol. 113, 105096, 2021.

[26] Y. Dongsheng, L. Yan, H.H.C. Iu, and H. Yi-Hua, "Mutator for transferring a memristor emulator into meminductive and memcapacitive circuits," Chin. Phys. B, vol. 23 no.7, pp. 070702–070711, 2014.

[27] A. Kumar, and B. Chaturvedi, "Novel CMOS dual-X current conveyor transconductance amplifier realization with current mode multifunction filter and quadrature oscillator", Circuit System Signal Processing, vol. 37, pp. 1-11, 2019.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on February 06,2023 at 05:20:01 UTC from IEEE Xplore. Restrictions apply.

Analog Integrated Circuits and Signal Processing
https://doi.org/10.1007/s10470-023-02147-x

A low power single bit-line configuration dependent 7T SRAM bit cell with process-variation-tolerant enhanced read performance

Bhawna Rawat[1] · **Poornima Mittal**[1]

Received: 13 November 2021 / Revised: 28 May 2022 / Accepted: 18 January 2023
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

Abstract
Cache memory is a key component for most microprocessors in embedded system. The increasing processing load has resulted in an upsurge in the demand for low power, high performance SRAM bit cells. Consequently, in this paper a 7T bit cell is designed for feature size 32 nm and 300 mV supply voltage. The improvement in the performance of the proposed cell is validated against the results obtained for pre-existing 6T, 7T, 8T, 9T, and 10T cells. The read and hold noise margin for the cell is obtained to be 96 and 68 mV respectively, whereas the static margin for the write operation is 170 mV. To perform a successful write operation, a pulse-width of 30 ns is utilized. The power analysis reveals that the proposed cell has minimal read/write power consumption. The leakage power for the cell is 8.4 pW and 1.2 pW for Q = '0' and '1' respectively. Tolerance analysis justifies that the cell maintains its functionality and yields credible outputs under process-voltage-temperature variations for static performance metrics. The layout for the proposed 7T cell occupies 0.584 μm^2 area. This is 5.55% smaller than a single ended 6T. The area for other 7T counterparts, 8T, 9T, and 10T cells is larger than the proposed cell.

Keywords Low voltage · Single port · Leakage power · Write margin · Layout analysis

1 Introduction

The increase in demand for round the clock digital connectivity, and the explosion of everyday personal data has increased the processing demand for embedded systems. A major component of efficient embedded system is Static Random Access Memory (SRAM); formed by bit cells organized in an array. The bit cell stores the data, whereas reading and writing into the cell is facilitated by peripheral circuitry; it is composed of decoders, bit-lines, and sense amplifiers [1]. The memory core along with the peripheral circuit consumes about 30% of the total power for the system and 50% chip area [2, 3]. Two factors—static and dynamic, combine to form the total power consumed by an SRAM cell. Static power has a linear dependence on supply voltage (V_{DD}), whereas, the latter has a quadratic dependence on V_{DD} [4]. The increasing demand for portable devices with longer battery life has made power consumption reduction a key design objective. The most trivial method to lower power consumption is to lower the operational V_{DD} for the cell [5]. But this is limited by the variations caused by process-voltage-temperature (PVT) in the nanometer vicinity [6]. Also, the declining trends for V_{DD} have resulted in standby power dominating dynamic power [7]. Reduction in V_{DD} also limits the speed of operation [8]. As V_{DD} decreases, the V_{DD}–V_{TH} factor diminishes. Thereby diminishing the operational frequency for the bit cell [9]. Therefore, the power consumption per instruction increase, even though the total power consumed by the cell decreases [10]. Consequently, hampering the overall performance for the SRAM based memory. Other reasons that result in the degradation of performance include—drain induced barrier lower (DIBL), geometry defects, electro migration, oxide breakdown, hot carrier injection, and electrostatic discharge [11–13]. Thus, designing a digital circuit operational at low levels of V_{DD} is a mammoth task for designers [14].

Another aspect that limits the performance of SRAM is the slow-fast corner; here if an NMOS and PMOS transistors are equally sized the performance for the CMOS circuit is skewed. This is caused by vast gap between the mobility for the NMOS and PMOS transistors, which

✉ Poornima Mittal
poornimamittal@dtu.ac.in

Bhawna Rawat
bhawnarawat12@gmail.com

[1] Department of Electronics and Communication Engineering, Delhi Technological University, Delhi, India

Published online: 03 February 2023

gets extenuated because of the skewed corner. The collective influence of the aforementioned factors increases the vulnerability [15] of the cell towards read, write, and access time failures. Besides, embedded cache memories with larger area are more prone to such failures [16]. In this paper a seven-transistor (7T) cell based on a single bit-line architecture is proposed. The presence of a single bit-line abets in limiting the dynamic power consumed by the cell. Another reason for choice of single ended configuration over differential ended cell is its ability to withstand reduction in V_{DD} [17, 18]. It also provides the added advantage of reduced silicon overhead, thereby increasing its economic feasibility.

In this paper, different SRAM bit cells are designed at 32 nm technology node to analyze the performance of the bit cell to identify their shortcomings. The learning of that are then utilized to design a modified 7T SRAM bit cell with improved performance. Along with this introduction section, the paper is organized into the following ten sections. Section II gives a schematic review of existing SRAM bit cells. Further, Section III defines the architecture and modes of operation for the proposed 7T cell. In Section IV, the simulation results for static analysis are presented and in section V, the cell ratio analysis is detailed. Further, in section VI the dynamic write ability is elaborated, while the power analysis is given in section VII. Section VIII justifies the tolerance analysis for the cell. The layout and area comparison of the bit cells are detailed in section IX. Finally, Section X recapitulates the findings of the paper.

2 Review of pre-existing bit cells in literature

Diverse architectures for SRAM cell are reported in literature to accomplish different design objectives—low area footprint, operational at low V_{DD}, reduced power consumption, faster operation, and high bit density. Figure 1 illustrates the circuit diagram for SRAM cells that were reported in the previous decade for different applications. The conventional SRAM cell is a six transistor (6TC) bit cell, as depicted in Fig. 1(a). Mutually connected inverter pairs (PUL-PDL & PUR-PDR) form the data core of the cell. The contents of the memory are retrieved using NMOS access transistors (ACL & ACR). Its speedy operation along with smaller area footprint make it the "de facto" model for implementation. However, the declining values for V_{DD} and shrinking technology node have made implementing 6TC cell fairly difficult. At lower technology nodes, the high leakage current increases the static power consumed by cell [19, 20]. The performance of conventional 6T cell suffers majorly because of the inherent read–write conflict in the cell. Also, due to common control signal for the read and write operation, it is also prone to half select disturbance, majorly causing a miswriting operation in the half selected cell. Also, due to V_{DD} scaling, the impact of process, voltage and temperature variations on the read and write operation becomes dominant.

For SRAM memory, maximum power is consumed during charging and discharging of bit-lines [17]. To reduce the discharging event, a single ended, dual port 6T (6TS) cell was described in 2019 [21]. The transistor based diagram

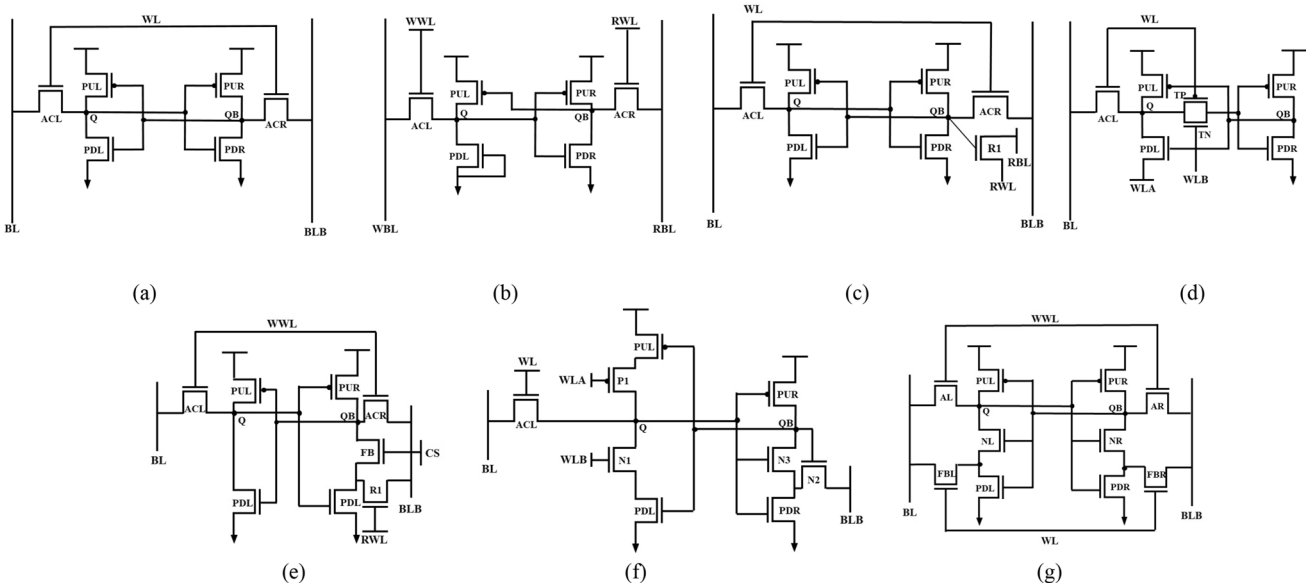


Fig. 1 Transistor based Circuit diagram for **a** 6TC, **b** 6TS, **c** 7TL, **d** 7TY, **e** 8TS, **f** 9TC, and **g** 10TK SRAM bit cells

for the cell is presented in Fig. 1(b). The memory element in 6TS consists of an inverter pair (PUR-PDR). The second inverter (PUL-PDL), is modified such that the gate and source for PDL are shorted to ground, thus it always operates in cutoff region. But, this degrade the Q = '0' margin, so a low V_{TH} transistor (LVT) is used to restore its performance. NMOS transistors (ACR & ACL) are used to access the contents of the data during the read/write operations. But in doing so, the number of control signals have increased leading to an increment in the overall capacitance of the memory. As bit cells are joined to form an array that acts as the memory core for the cache memory, a small increment in the capacitance for a bit cell has a multi-fold impact on the capacitance of the entire column. This results in increase in dynamic power for the memory and also increases designing complexity for the sense amplifier (and other peripheral circuitry). Increment in capacitance of the bit cell also deteriorates its dynamic write ability. Consequently, this has a radical effect on the dynamic performance and power consumed by the bit cell. This cell though outperforms the 6TC bit cell in terms of leakage, the overall energy consumption and area for the embedded memory is quite large.

Another method to uplift the performance of an SRAM cell is to combine 6TC cell with an independent read port to isolate the read and write operation. A similar 7T bit cell architecture (7TL) was reported by researchers in 2017 [22]. The circuit schematic for the same is illustrated in Fig. 1(c). An additional NMOS transistor (R1) is used to perform the read function; R1 is controlled by node QB. Read bit-line (RBL) via R1 to read word-line (RWL) form the read discharge path for the cell. As the gate of R1 is connected to the QB node of the memory core, it conducts only when Q = '0'. Thus reducing the read activity factor for the bit cell to 0.5. The only shortcoming for this cell is that the additional read port reduces the write drive current for the cell thereby resulting in a poor write performance.

Yet another 7T (7TY) cell configuration was suggested in 2016 [23]. The schematic for the same is depicted in Fig. 1(d). 7TY has a single-ended, single port architecture. Transmission gate (TN-TP) is used for disconnecting the feedback between the inverter pair. The cell is accessed via NMOS transistor (ACL). Four control signals are used to navigate the bit cell according to its modes of operation. These control signals are, WL, WLA, WLB and BL. But, its dual pulse based write operation reduces the speed for the bit cell. Also, the increased count of control signals; WLA and WLB results in additional peripheral circuitry for the bit cell. Consequently, having a substantial increase in the layout footprint and power consumption for the memory array.

A stable and low voltage enabled SRAM bit cell was reported in 2017 [24]. The bit cell is formed by eight transistors (8TS) and relies on a single bit-line enabled read and a dual bit-line enabled write operation. An NMOS transistor (FB) is used to disconnect the V_{DD} from ground rails; this is done to boost the single ended read stability for the cell. The transistor based diagram of the 8TS is presented in Fig. 1(e). The stacked transistor configuration also aids in reducing the leakage current for the cell. This cell has reported an improved performance but for its efficient working relies on an XOR gate common to a row. A dedicated XOR gates for each row has multifold impact on the area footprint and the overall power consumption for the memory array.

A 9T (9TC) single ended SRAM bit cell using Schmitt Trigger (ST) based inverter was reported in 2020 [25]. ST based inverter has the ability to manipulate its switching voltage subject to the direction of input transition [26]. The circuit diagram for the 9TC is illustrated in Fig. 1(f). The data storage in the cell is facilitated by a stacked inverter (PUL-P1- N1-PDL), plus a ST inverter (PUR-N2-PDL). The word-line (WL) control signal is used to access the data storage node for the bit cell. While the signal WLA and WLB are used to control P1 and N1 respectively. The WLA/WLB signals for the read and write operation are 0/1 and 1/1 respectively.

Another ST based cell was described by Kulkarni and Roy in 2012 [27]. This cell was an improved version of the cell described in 2007 [28]. This cell has a ten transistor (10TK) structure and the circuit schematic for the same is presented in Fig. 1(g). The write-word-line (WWL) control signal is used to access the data core formed by mutually connected ST inverter pair (PUL-NL-PDL & PUR-NR-PDR). While, the WL signal controls the feedback for the bit cell. This cell has an improved hold stability but due to its differential read operation the read stability for the cell suffers. While, the increased transistor count increases the power consumed by the cell [29]. Moreover, the increased area footprint makes the cell less economically efficient. A major limitation of various bit cells is that they often concentrate on enhancing only one of the performance parameters. And, usually the improvement in one comes at the cost of another. But it is imperative to design and propose a bit cell that balances the performance for all modes of operations simultaneously. The aforementioned cell are used to validating the performance of the proposed bit.

It can be concluded from the aforementioned review that the performance for SRAM bit cell has suffered drastically due to technology node scaling and supply voltage reduction. The 6TC cell is unable to perform for lower technology node due to the inherent read–write conflict, so the 6TS cell was reported. It eliminated the inherent read–write conflict, but resulted in high power requirement, thereby making it unfit for low power application. Consequently, the 7TL cell was reported to isolate the read and write port, while maintaining a low operational voltage requirement. To lower the power consumption for the bit cell, a single ended, single port cell—7TY was reported. But, it is dependent on two

additional control signals thereby increasing the interconnect load for the circuit. Increase in control signal count also increases the effective capacitance for the memory, thereby, increasing delay in the circuit. Another bit cell that resolves the read–write conflict is the 8TS cell. But, it leads to floating condition in the circuit, thereby increasing its vulnerability to error. While, to limit the power consumption for the bit cell, 9TC and 10TK cell were reported with Schmitt Trigger topology to improve the performance of the bit cell. But the area penalty for the 9 T and 10 T SRAM bit cell is extremely high, thus making them economically unfeasible. This mandates the requirement for an SRAM bit cell that efficiently resolves the read–write conflict; it operates at low values of supply voltage, and has lower layout area.

3 Proposed structure

The proposed SRAM is a seven transistor (7TP) bit cell that utilizes a single bit-line configuration. The data core of the cell comprises of mutually connected inverter pairs; inv1 (PUL-PDL) and inv2 (PUR-PDR). Unlike a conventional 6TC which relies on the mutual feedback connection between inverter pair, the 7TP SRAM bit cell disconnects the feedback via an additional LVT NMOS transistor (FB). NMOS transistors (WA and RA) facilitate the read/write operations for the cell. The transistor based circuit for 7TP cell is shown in Fig. 2. The read and write functions are controlled by signal read-line (RL) and word-line (WL) respectively. The operating conditions of the three control signals; W, WL, and RL are given in Table 1.

3.1 Hold operation

Throughout the hold operation, the effective functioning of the cell is limited to the coupled inverter pair. To maintain the cross coupled connection the feedback NMOS transistor (FB) is in conducting state for the entire duration of the hold operation.

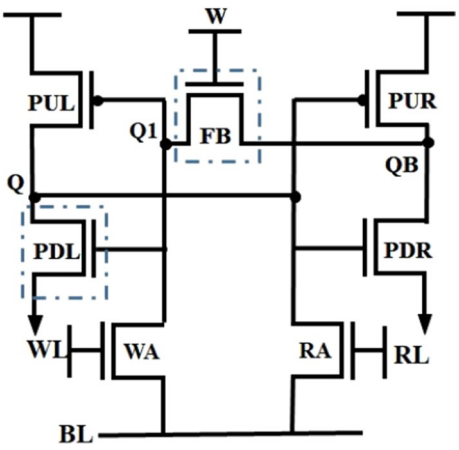


Fig. 2 The transistor based circuit diagram of the 7TP bit cell

3.2 Read operation

The 7TP cell relies on RA, controlled by RL, for the read operation. To additionally reduce the read vulnerability, the cell relies only on inv2 (PUR-PDR) for read operation. Thus, when the bit cell stores Q = '1', there is no discharge path for the bit line. This is an important aspect of the bit cell, as it reduces the activity factor for 7TP cell from 1 to 0.5. Thus assisting in limiting the dynamic power consumed by the cell when operating in read mode.

3.3 Write operation

The first step to steer the cell into write mode is to disconnect the mutually connected inverter pair. The same is done by exerting W = '0'. This effectively reduces the coupled inverter pair to a cascaded configuration dependent on BL. Firstly, the complement of the data to be stored at Q is set on BL. Then, WL signal is exerted to activate WA transistor. Thus, the data on BL is transferred to node Q1, which drives inv1 to produce the desired bit value at Q. Then Q drives inv2 to write data at QB. The additional signal W is utilized by the circuit to control the internal configuration of the bit cell. It enables the cell to convert the mutually coupled inverter pair configuration to a cascaded inverter configuration. But, this results in an increased control signal, resulting in increased interconnects in the circuits. But, it helps improve the write ability for the cell at optimal transistor sizing; therefore the area for the proposed 7T SRAM bit cell is lesser than its other counterparts (explained in Sect. 9). Therefore, the slightly increased overhead because of the extra control signal gets overshadowed by its benefits.

The RA transistor is cutoff when the cell is operating is write mode. To additionally assist the write operation, the voltage transfer characteristics for the inverters are adjusted. Also, the write operation for the cell relies on the swiftness of inv1; as the delay of a gate is inversely proportional to V_{DD}–V_{TH}. Thus, reducing the value of V_{TH} decreases the delay for the gate [30]. Therefore, PDL being an LVT NMOS ensure a swifter write operation. Optimally sized transistors along with the use of LVT NMOS for inv1 are step incorporated to build-up the write performance for the cell.

Table 1 Truth table for proposed 7 T SRAM bit cell

	W	WL	RL
Hold	1	0	0
Read	1	0	1
Write	0	1	0

4 Static analysis

The tool used for designing all the bit cells is Symica DE. The library files used for simulating the circuits is 32 nm technology node by Predictive Technology Models. The results obtained for the 7TP cell are compared against 6TC, 6TS, 7TL 7TY, 8TS, 9TC, and 10TK for validation. Each of these bit cell has already been discussed in Section II. The environment temperature is set to 27 °C for the complete study. The dimensions for all the cells are mentioned in Table 2. An important parameter that determines the stability of bit cell is the retention of stored data in standby as well as read access mode. The conventional method to measure the former is static noise margin (SNM); it is the maximum DC noise level a cell may withstand before an erroneous flip in the data of the bit cell is registered [28]. It is used to quantify the cell's stability and its immunity to noise [31]. It is obtained by graphically superimposing the transfer characteristics of the inverter core resulting in a two lobe based curve [32], called the butterfly curve. The SNM is the measure of side of the biggest square that perfectly fits inside the smaller lobes of the butterfly curve [33].

4.1 Hold noise margin

An SRAM cell for most part of its functioning remains in the hold mode. Therefore, the first parameter under analysis is the hold SNM (HSNM). The SNM for each bit cell is calculated as the side of the largest square that completely fits inside the smaller lobe of the butterfly curve. For a symmetrical butterfly curve, either of the lobes may be considered for calculation of SNM. To demonstrate the calculation of SNM, a square is embedded in the HSNM and RSNM curve for the butterfly curve. The same is depicted in Fig. 3(h). The HSNM for the 7TP cell is calculated as difference between the x-axis (y-axis) coordinates for red square embedded in the hold butterfly curve (123–27 or 279–183 = 96 mV). The RSNM for the 7TP cell is calculated as difference between the x-axis (y-axis) coordinates for blue square embedded in the read butterfly curve (133–65 or 289–221 = 68 mV). The values for HSNM of all the cells detailed in the schematic survey along with the proposed cell are presented in the Fig. 3; along with the read butterfly curve (explained in the subsequent sub-section) for each cell. All the obtained HSNM values are tabulated in Table 3. The values suggest that 7TP cell has a moderate hold performance with respect to other bit cells. The proposed cell performs better with respect to 6TC, 6TS, 7TL, and 7TY, while the HSNM of the 10TK cell is equivalent to the HSNM of the 7TP cell

Table 2 Cell Dimensions (in nm) for various SRAM bit cells

	6TC	6TS	7TL	7TY	7TP	8TS	9TC	10TK
PUL	72	120	64	64	64	160	64	200
PUR	72	180	64	64	64	160	64	200
PDL	72	180	166	96	166	160	96	200
PDR	72	180	96	96	96	160	96	200
ACL	144	120	128	96	96	160	64	200
ACR	144	120	128	96	96	160	64	200
R1	–	–	128	–	–	160		–
R2	–	–	128	–	–	160		–
TN	–	–	–	64	–	–		–
TP	–	–	–	64	–	–		–
DL	–	–		–	64	160		–
FB	–	–		–	64	160		–
WA					96			
RA					96			
P1							96	
N1							96	
N2							96	
N3							96	
FBL	–	–	–		–	–		200
FBR	–	–	–		–	–		200
NL	–	–	–		–	–		200
NR	–	–	–		–	–		200

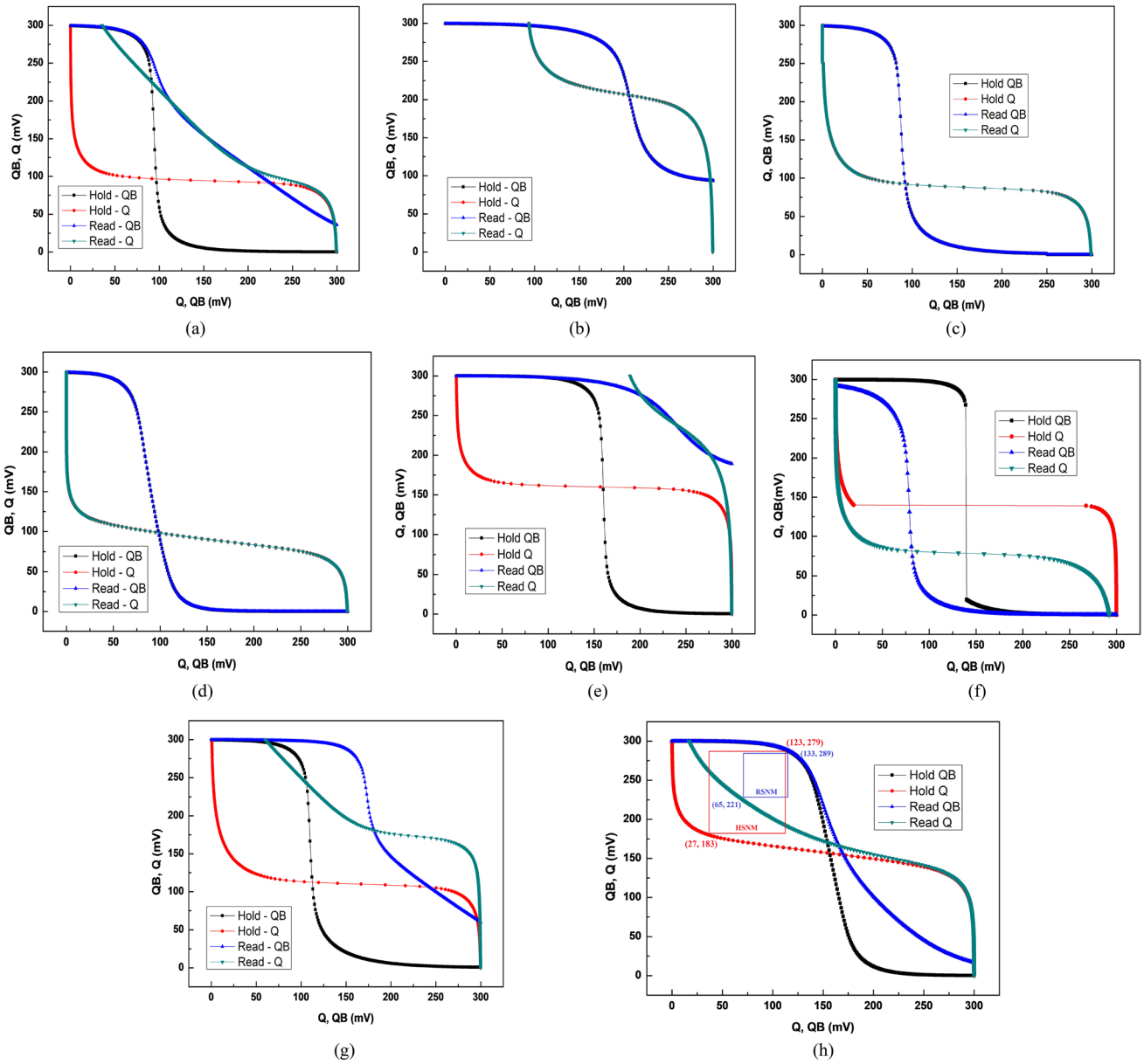


Fig. 3 Combined butterfly curves for HSNM and RSNM for **a** 6TC, **b** 6TS, **c** 7TL, **d** 7TY, **e** 8TS, **f** 9TC, **g** 10TK, and **h** 7TP SRAM cells

at 96 mV. But the 8TS and 9TC cells have a higher HSNM value compared to the 7TP cell.

4.2 Read noise margin

For the duration of the read mode, the bit value stored in the cell is retrieved using access transistors. For 6TC cell, the word-line is set high to cause conduction through the read access transistors. Thus, subjecting the bit cell to vulnerability, as there is a possibility of undesired flipping of data stored in the memory core. Therefore, it is important to measure the resilience of the cell during read mode; it is quantified as read SNM (RSNM). A differential read

Table 3 The performance of SRAM cells for static parameter (HSNM, RSNM, and WM)

Cells	HSNM (mV)	RSNM (mV)	WM (mV)
6TC	80	17	145
6TS	37	37	105
7TL	74	74	189
7TY	74	74	-
7TP	96	68	170
8TS	104	0	111
9TC	117	80	195
10TK	96	57	145

operation always has poor performance as the cell becomes more prone to noise [34] for both '0' and '1', as the cell discharges for both. This problem is more intensified by decreasing V_{DD}. Whereas, during a single ended read operation, the cell discharges for either '1' or '0', thereby reducing the possibility of error by half. Additionally, to boost the read discharge current vulnerability for the bit cell, the voltage transfer characteristics for the 7TP bit cell is modified to make it less vulnerable during the read operation. Consequently, the RSNM for the 7TP is higher in comparison to the other bit cells with a differential read configuration; the same can be inferred from Fig. 3. The HSNM and RSNM for all the bit cells are depicted in Fig. 3.

Another aspect vital for differential bit line sensing is the minimum required voltage of 50 mV [27]. The complementary bit lines of the bit cell should develop a 50 mV [27] minimum difference to sense the output. But the results show that the RSNM for 6TC is 17 mV which is way below the permissible limit. Therefore, it can be said that 6TC bit cells will fail to perform effective read operation at V_{DD} as low as 300 mV. On the other end, the cells 6TS, 7TL, 7TY, 9TC, and 10TK show RSNM of 37, 74, 74, 80, and 57 mV respectively. The proposed cell registers a RSNM of 68 mV even with a single ended port; hence it will perform a valid read operation.

4.3 Write stability

The robustness of a cell during write mode is measured as its ability to alter the stored data. Write margin (WM) is a reliable parameter for a cell's write ability [35]. By definition, WM is estimated as the difference between the voltage level of V_{DD} and WL when the bit stored in the cell flips [36]. A balanced WM is essential to ensure a successful write operation. A very low value for WM makes it challenging to perform the write operation, while a high WM escalates its vulnerability to noise. WM is measured using the methodology recommended by Islam and Hasan [37]. The WM values for all bit cells except 7TY are tabulated in Table 3. The write operation for 7TY is divided into two pulses, while, all the other cells in this study utilize a single pulse to perform the write operation therefore, its evaluation is not included, to avoid projecting it as pseudo poor.

The WM values obtained for 6TC/6TS/7TL/8TS/9TC/10TK cells register a WM of 145/105/189/111/195/145 mV respectively. Table 3 collectively summarizes the static performance data for all bit cells. The graphical comparison for the same is presented in Fig. 4. All the cells have a commendable performance during the hold operation but their RSNM value acts as the differentiator. The RSNM butterfly curve of 8TS cell collapse; thereby making it highly susceptible to read noise. Another cells with considerably low RSNM value is 6TC at 17 mV. This cell is also susceptible to read noise. The other bit cells have reliable values for RSNM metric (Table 4).

The 9TC cell has a good performance for both HSNM and RSNM but its WM is extremely high, making its write operation highly susceptible to noise. The 10TK cell has a commendable performance for the three parameters but its high layout area (explained in section IX) make it a less viable. Thus, after analyzing the static performance of the 7TP cell, it can be stated that this cell has achieved a balance between the three modes of functioning—hold, read and write.

5 Cell ratio analysis

Cell ratio for a cell is the proportion between the aspect ratio of the access transistor and the corresponding pull-down network. For a conventional cell it determines the

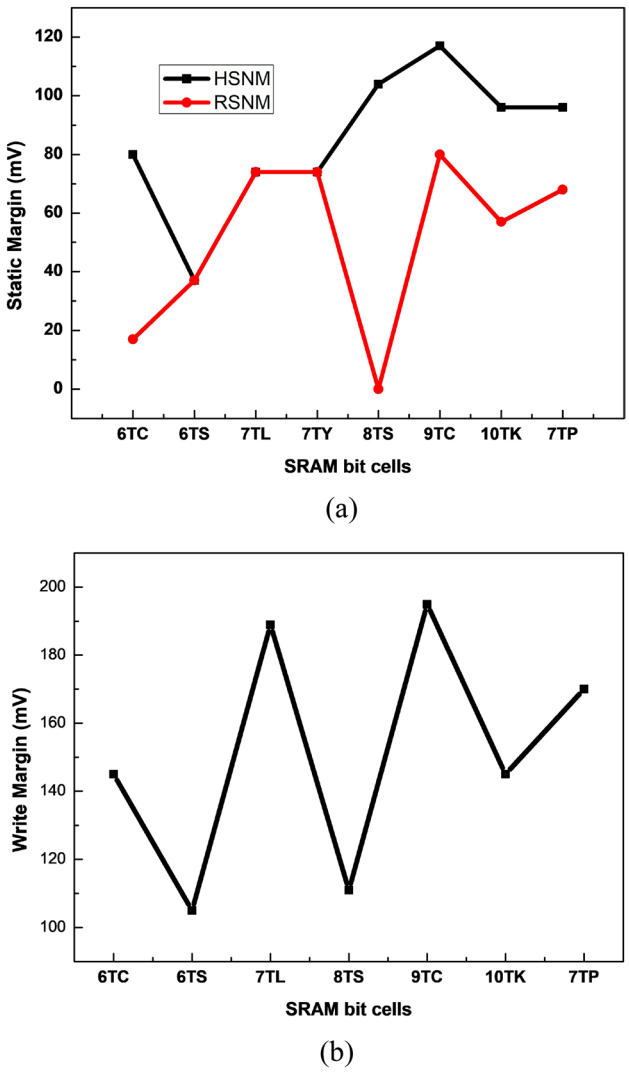


Fig. 4 Graphical representation of the static analysis for **a** HSNM and RSNM, **b** WM for various SRAM bit cells

Table 4 Performance of the 7TP cell in terms of Saturation voltage, delay, and SNM with varying cell ratio

Cell ratio	Saturation voltage (mV)	Time delay (ps)	SNM (mV)
1.5	91	59	63.6
2	76	71	87.36
2.5	70	90	90
3	62	101	84.56
3.5	56	129	78.43
4	53	156	70.56
4.5	49	180	67.2
5	46	245	61.26
5.5	43	280	56.58
6	40	330	52.8

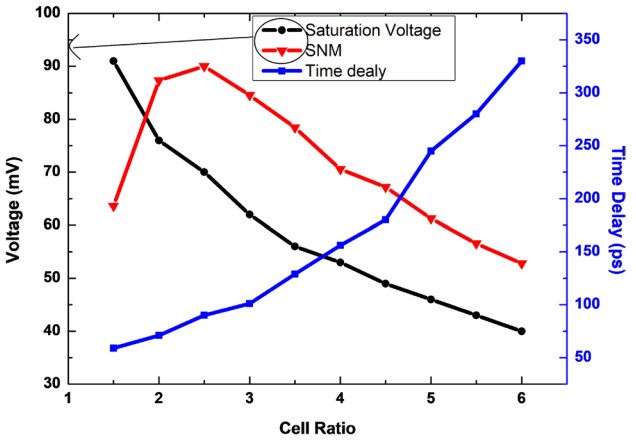


Fig. 5 Performance of the 7TP SRAM cell for varying cell ratios

optimum sizing for the best RSNM value for the cell. For this analysis the cell ratio of the 7TP bit cell is increased from 1 to 6 in steps of 0.5. The parameters that are measured with respect to the cell ratio are the HSNM, saturation voltage and the delay in attaining 40 mV (the minimum saturation voltage attained by all bit cells).

$$\text{Cell Ratio} = \frac{\textit{Aspect Ratio of Access Transistors}}{\textit{Aspect Ratio of Pull Down Network}} \quad (1)$$

The single-ended read port for the 7TP cell it carefully designed to not register a destructive read operation. For each cell ratio, there is a corresponding maximum voltage the BL rises to and it is referred to as saturation voltage. While, the delay in the cell is calculated as the time consumed by the cell to raise the bit-line voltage to 40 mV. The corresponding data is plotted on a double Y axis plot as shown in Fig. 5. An increase in the delay and saturation voltage is observed with an increment in the cell ratio for the proposed bit cell. While RSNM vs. cell ratio trend line attains peak corresponding to cell ratio of 2.5. The saturation voltage is sufficiently high, while the delay is fairly low for the same cell ratio. Therefore, 2.5 is considered as the most optimum cell ratio for the 7TP SRAM bit cell.

6 Dynamic write analysis

During static analysis, the pulse width for WL is assumed to be infinite. While, for real time operations the pulse width for write operation is finite. Thus, to characterize the real time write ability for a cell, it is primal to measure the write pulse width. It is defined as the least possible pulse width required to alter the bit stored in the cell. If it is less than that then the data in the cell will not change, causing a write failure. Figure 6 depicts the write operation for two consecutive pulse. The initial conditional for the 7TP is Q = '0' and QB = '1'. The pulse-width for the 7TP cell is found to be 30 ns.

To analyze the dynamic write ability of SRAM bit cell with increment in values of V_{DD}, the plot for V_{DD} vs pulse width is presented in Fig. 7. The plot is split for V_{DD} varying from (a) 0.3 to 0.6 V and (b) for 0.7 to 1 V. This is done for the ease of comparison as well as to demonstrate commendable dynamic write ability for the proposed cell for low as well as high V_{DD}. Figure 7(a) validates that the 7TP cell operates at the lowest pulse width of 30 ns at 300 mV. The highest dynamic pulse width is utilized by 6TS bit cell at 100 ns for V_{DD} at 300 mV. While, the performance of 6TC, 7TL, 8TS, and 10TK cell is better in comparison to 6TS but slightly higher than the 7TP cell. But these aforementioned cells have a differential ended write operation and it is an accepted notion that in comparison to the former a single ended write is more challenging. Thereby, establishing the improvement in the performance of the 7TP cell. The 9TC cell also has a single ended operation but its pulse width requirements are slightly higher than the 7TP cell.

Therefore, for lower values of V_{DD} it can be deduced that 7TP cell requires a shorter pulse width with respect to other cells. The proposed SRAM bit cell is able to maintain its credibility for V_{DD} in the range of 0.7 to 1 V. The cell displays successful write operation for least value of pulse width for all values of V_{DD}. The SRAM bit cells that have comparable performance to 7TP bit cell are the 6TC, 7TL, and 10TK SRAM bit cells but they are differential in nature so that automatically makes it easier to write into them. Therefore, it is appropriate to infer that 7TP cell has an improved performance in comparison to other cells.

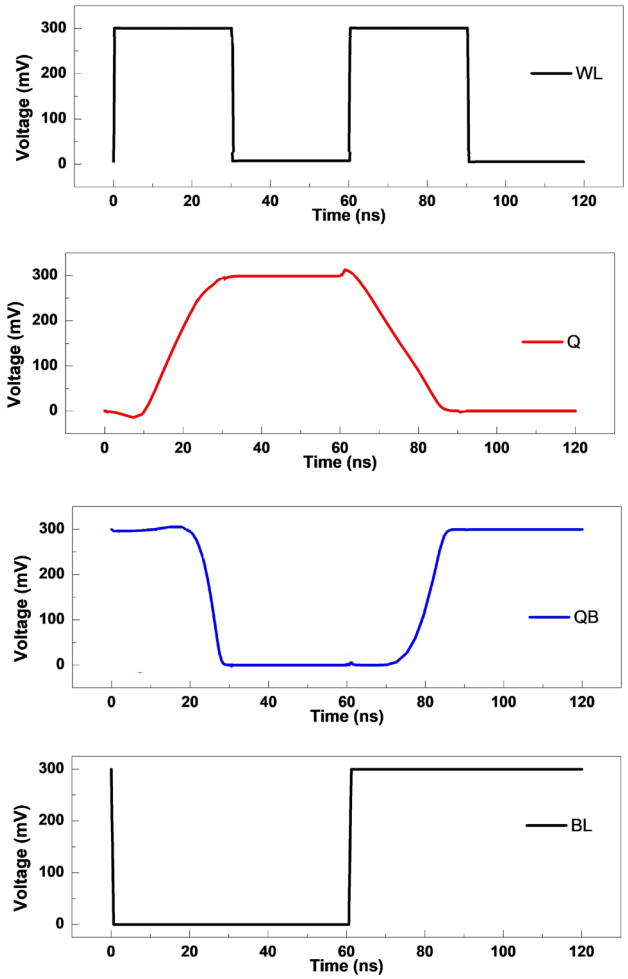


Fig. 6 Two consecutive write operation for the 7TP cell performed at pulse width of 30 ns

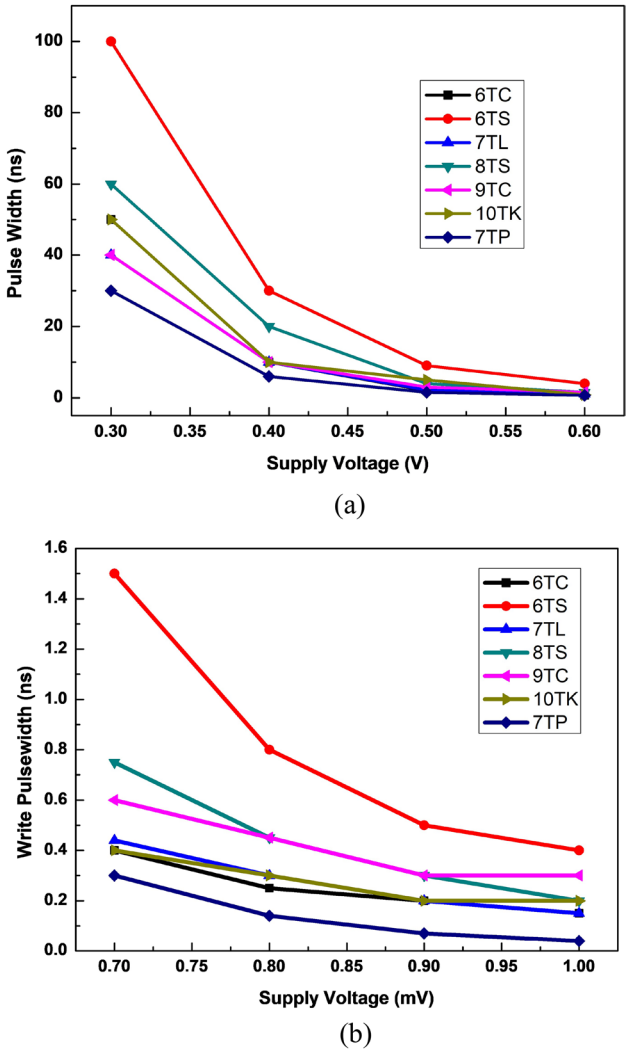


Fig. 7 Characteristic plots of pulse width as a function of V_{DD} **a** 0.3 to 0.6 V and **b** 0.7 to 1 V for SRAM cells

7 Power analysis

7.1 Read and write power analysis

A cell consumes most power during the phase it is being accessed. This can be either read or write operation. But of the two, more power is consumed during the write operation. This same conclusion can also be drawn from the Fig. 8, which gives a graphical comparison for power consumed by the cell in (a) read and (b) write mode.

For a differential read operation the power consumed is evaluated over the time required to drop the bit-line voltage to 50 mV. While, for single ended cell the pulse width is taken to be time needed for data to reach the vicinity of switching voltage. The power consumption during the read operation is highest for 8TS cell for data value '1'. The read cell though has a single ended operation the stacking configuration increases the read drive current. The device dimensions are also large for the cell, thereby increasing the power dissipation for the cell. 6TC cell also experiences a high read power for both '0' and '1', due to its differential nature. Though 10TK cell is also differential in nature, its read power is low owing to the shortened discharged path and ST inverter configuration. The remaining SRAM bit cells—6TS, 7TY, are single ended in nature and have read power for Q = '1' less than 7TP due to no discharge path for the operation. The power consumption for Q = '0' is comparable for the three cells.

During the write operation a longer pulse-width is needed to deliberately modify the data present in the cell. By default, it is assumed that the data stored inside the cell is Q = '0' and QB = '1'. For the first cycle, the input data exerted into the cell via BL is '0'. This results in '0' at node Q1, thus the node Q develops '1' and the node QB becomes '0'. Then, for

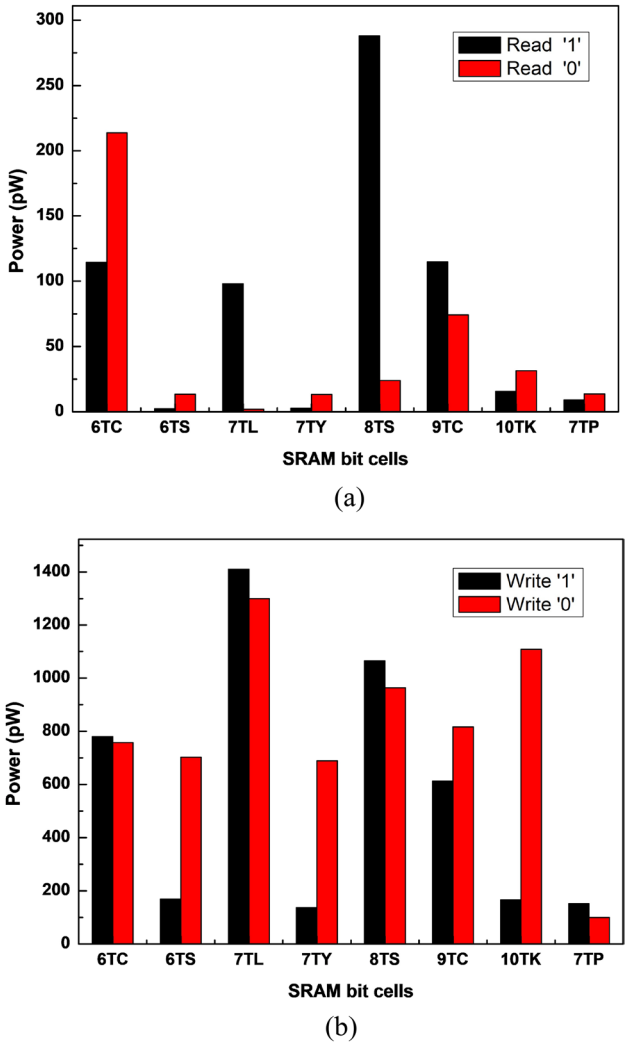


Fig. 8 Power consumption by the different SRAM cell when operating in **a** read and **b** write mode

the next cycle, the BL signal is exerted high, to develop '0' and '1' at node Q and QB respectively. So for write power analysis the operation is performed and the power consumed during the process is recorded. The power dissipated during a differential write operation is higher with respect to a single ended cell [38]. The same can be inferred from Fig. 8, as the power consumption for 6TC, 7TL, 8TS, and 10TK is considerably high. Single ended cells are generally biased towards either '0' or '1', resulting in high power consumption for the other value. This characteristic of a single-ended SRAM cell is attributed to its coupled inverter based memory core. But the write operation for the 7TP cell modifies the inverter core to form a cascaded inverter configuration. Also, the use of LVT MOS reduces the power requirement for the cell. Thereby making it more power efficient.

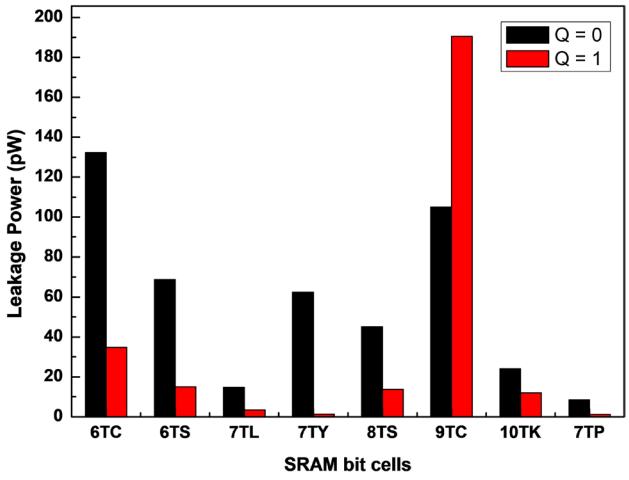


Fig. 9 Graphical comparison for power consumed during the hold state for SRAM bit cells

7.2 Leakage power analysis

Most cells in the SRAM array are maintained in hold state. Hence, leakage power consumed by a cell during the hold mode forms a key component of total power consumption for a cell [37]. Additionally, in submicron technologies, the increase in leakage is attributed to nanometer feature size for the device. Therefore, it is essential to study the leakage power for a cell. To identify the leakage power for the bit cell, the leakage current is simulated and the standby power is calculated as V_{DD} times of leakage current [39]. A graphical comparison of power consumed because of leakage in the cell is depicted in Fig. 9. It is inferred from Fig. 9 that the leakage power recorded for the 7TP cell is very low. The leakage power recorded for 7TP cell is 8.4 pW and 1.2 pW for Q = '0' and '1' respectively. The only other cell with comparable leakage power is 7TL, as it also has a single-ended nature. The highest leakage power is registered for 9TC; this is a consequence of its additional control signals being ON during the hold state as well, causing a high current to flow through the circuit to keep the inverter core intact.

The high values for leakage of 6TC is attributed to its differential nature and for 6TS is due to PDL being always in ON state. For 10TK cell, though differential in nature has low leakage due to its stacked inverter configuration along with the disable control signals help reduce the leakage power consumption for the cell [40].

8 Tolerance analysis

8.1 Process variation tolerance

SRAM bit cell have observed a very sharp decline in technology node resulting in serious variability issue at the circuit level [41]. Additionally, the increasing size of cache memory has resulted in upsurge in process variations among cells. This causes degradation in performance of adjacent bit cell due to asymmetrical characteristics [15]. Therefore, statistical methods are utilized to detect this degree of variation in the performance all the cells. Inter-die V_{TH} mismatch is taken as the main statistical parameter for evaluation. To ensure reliability, a failure probability of $P_{fail} < 10^{-10}$ is required [41]. The probability is Gaussian in nature and is calculated based on μ-6σ parameter. Here, μ is the mean value of V_{TH} and σ is its standard deviation. The factors responsible for process variation are line edge roughness, chemical–mechanical polishing, random dopant fluctuation, and lithography effects [42]. Most of these parameters are neither predictable nor can they be controlled. But the location and distribution of these factors influences the V_{TH} [42]. Therefore, in this work, V_{TH} is used as the primary source to predict the variability tolerance of the 7TP cell.

The Monte Carlo simulations for static parameters of 7TP are performed for 10,000 data point. The results for the same are depicted in Fig. 10. The process variation results yield the least value for HSNM and RSNM to be 87 and 48 mV respectively. While the WM obtained for maximum variation is 195 mV. The Monte Carlo simulation of the 7TP bit cell for 10,000 data points is compared with respect to other cells. The result for each cell deteriorate in keeping with 6σ global variations, the worst case value attained for the cells are demonstrated in Fig. 10(d). The 6TC, 6TS, and 8TS cell register a drop in HNSM, while their RSNM curve collapses. Thus, certifying a high read failure probabilty for the cells. Whereas, the 7TL and 9TC cells experience increases in WM to 200 and 210 mV respectively, which is considerably high, making them highly susceptible to noise. The

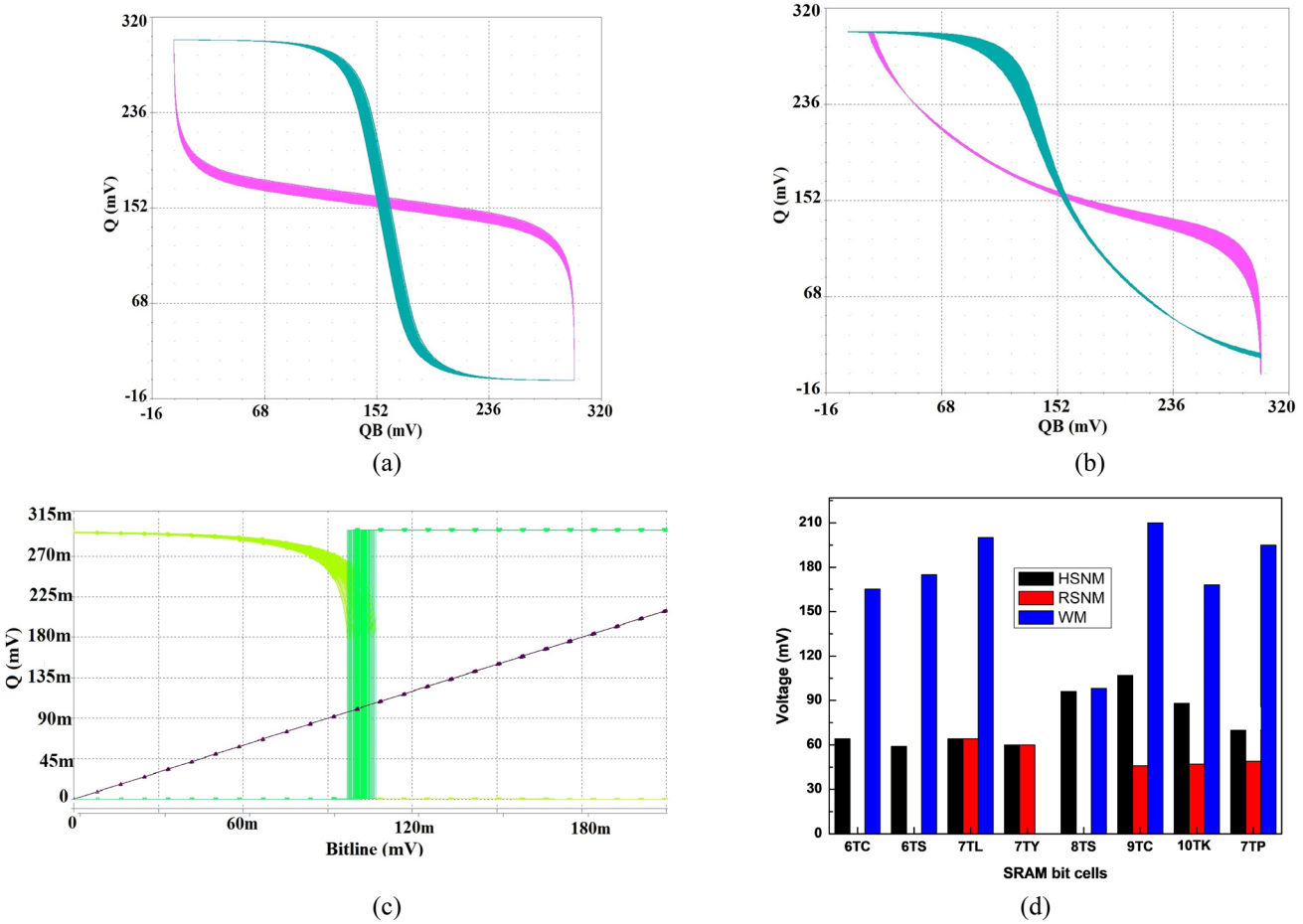


Fig. 10 Effect of process variation on **a** Hold SNM, **b** Read SNM, **c** WM for 7TP cell, and **d** graphical comparison for minimum values obtained at 10,000 Monte Carlo simulations for bit cells

6TC, 6TS and 10TK cells have a reliable WM in the vicinity of 165 mV but their poor RSNM values make them highly prone to read failure.

The MonteCarlo simulations reveal the cells vulnerability to failure subject to 6σ global variations. So it can be deduced from the analysis that all cells register a decline in their perfromance but the 7TP SRAM bit cell maintain reliable perfromance.

8.2 Voltage tolerance

The static analysis is included for the proposed cell for voltage levels lower than 0.3 V up to 0.14 V. The values as depicted in the Fig. 11 reflect that the cell is unable to perform write operation at 0.16 V or lower voltages. The least voltage at which the cell is able to maintain its functionality is 0.18 V.

However, the static performance at this voltage is so low that sensing the read for single ended cell would become highly difficult. Therefore, in keeping with stable performance for all the modes of operation, 0.3 V is taken as the least voltage for operation for the bit cell.

8.3 Temperature tolerance

SRAM circuits are required to function under a wide range of temperature. A device temperature may surge due to relentless operation or environmental circumstances. Therefore, to account for the temperature variation encountered by the cache memory, a temperature variation analysis is performed for the proposed cell. This is accomplished to ensure the performance reliability of the cell. The analysis is performed for temperature varying from 25 to 85 °C.

The results for the analysis is presented in Fig. 12. With the variation in temperature a deviation in performance of static parameters for the cell is observed. The cell maintains

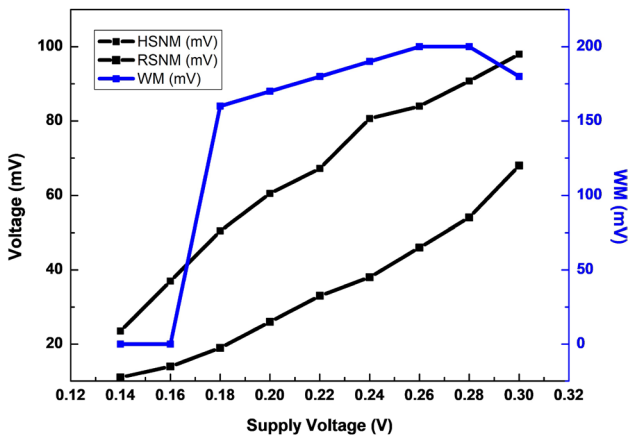


Fig. 11 Variation in RSNM and WM for 7TP cell with varying V_{DD}

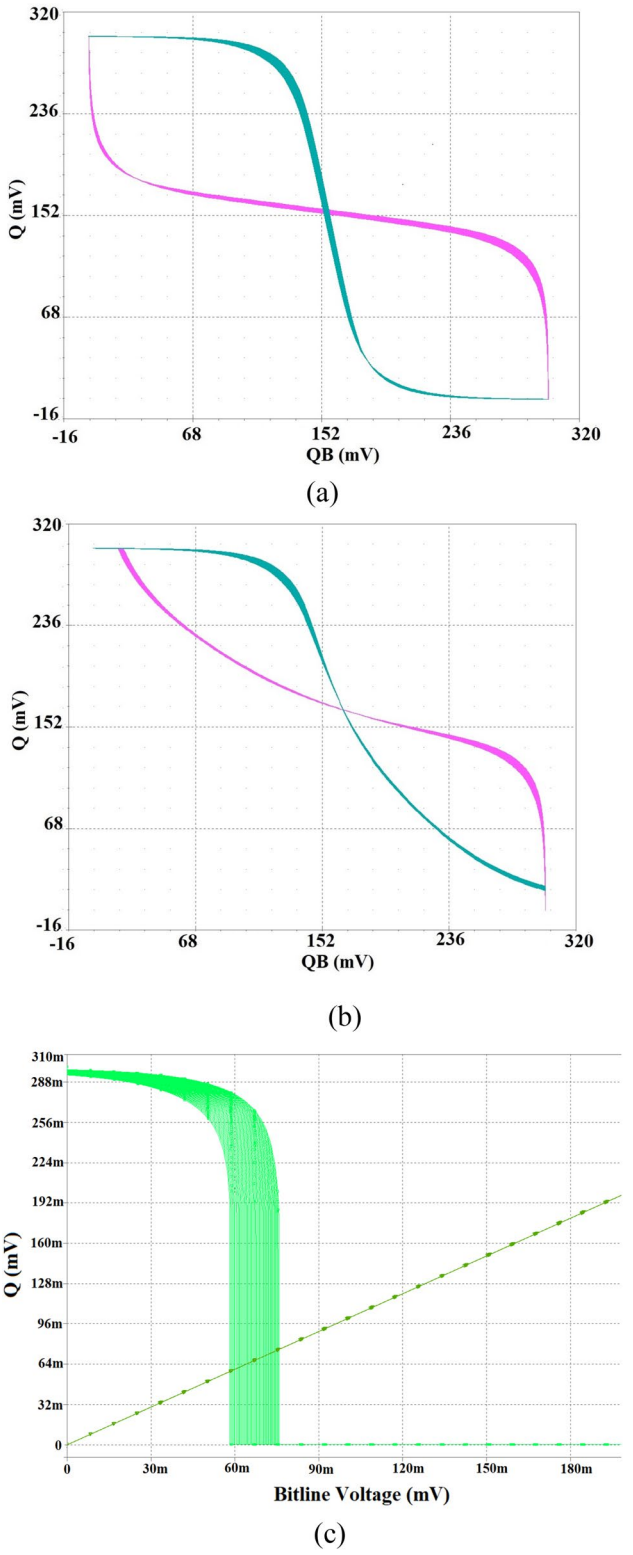


Fig. 12 Deviation in performance of **a** HSNM, **b** RSNM, and WM for 7TP cell due to temperature variation

HSNM and RSNM margins at 90 and 58 mV respectively; which is a decline of 0.1 and 0.16 mV/°C from its hold and read static performance respectively. While, the WM increases to 230 mV. This data justifies the performance of 7TP bit cell as reliable with variation in temperature.

8.4 Temperature tolerance

SRAM circuits are required to function under a wide range of temperature. A device temperature may surge due to relentless operation or environmental circumstances. Therefore, to account for the temperature variation encountered by the cache memory, a temperature variation analysis is performed for the proposed cell. This is accomplished to ensure the performance reliability of the cell. The analysis is performed for temperature varying from 25 to 85 °C.

The results for the analysis is presented in Fig. 12. With the variation in temperature a deviation in performance of static parameters for the cell is observed. The cell maintains HSNM and RSNM margins at 90 and 58 mV respectively; which is a decline of 0.1 and 0.16 mV/°C from its hold and read static performance respectively. While, the WM increases to 230 mV. This data justifies the performance of 7TP bit cell as reliable with variation in temperature.

9 Layout and area

SRAM cells are arranged into large arrays to form cache memory. Thus, in order to integrate large number of bit cells into the cache memory of the SoC, it is imperative to have minimal area for a bit cell. The layout design for the bit cells are given in Fig. 13. The dimensions of each cell are mentioned alongside the layout. Also, the length (μm), width (μm) and area (μm^2) for each cell are plotted in Fig. 15 for an easy comparison.

The layout for the 4×4 array for the 7TP cell is depicted in Fig. 14. The layout is designed to demonstrate the routing directions for supply lines and control signals for the cell. The V_{DD}, ground potential, W and RL1 signal are routed vertically, while the BL and WL signal for the cell are routed horizontally. The write operation for the bit cell is controlled via control signals WL and W, which form independent row and column based signals respectively. While, the read operation is controlled by the RL signal. The 4×4 array based layout for the bit cell helps understand the array based implementation for the SRAM memory. For ease of comparison the bit cell area along with the cell dimension (length and width) are plotted in Fig. 15.

The SRAM bit cell with the least area of 0.456 μm^2 is 6TC cell. This is a key factor that makes 6TC the most

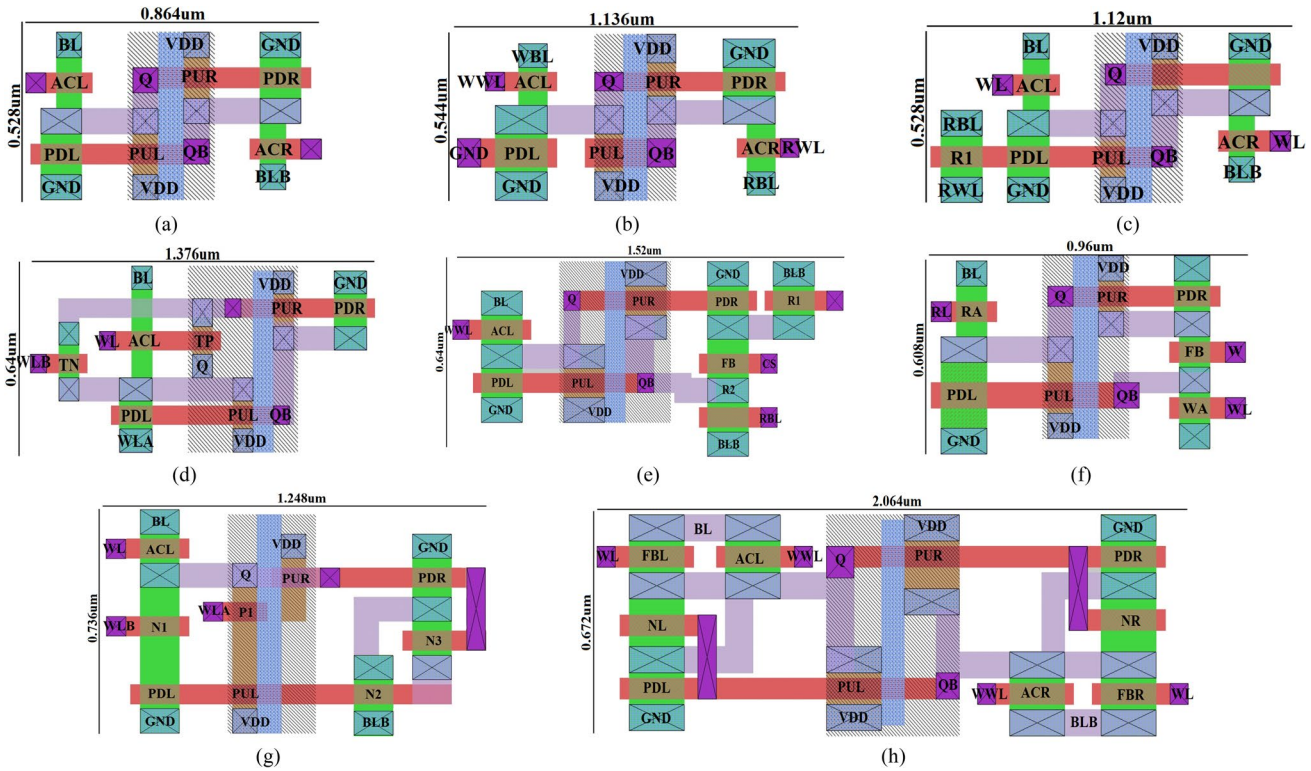


Fig. 13 Layout for the **a** 6TC, **b** 6TS, **c** 7TL, **d** 7TY, **e** 8TS, **f** 7TP, **g** 9TC, and **h** 10 T SRAM bit cells

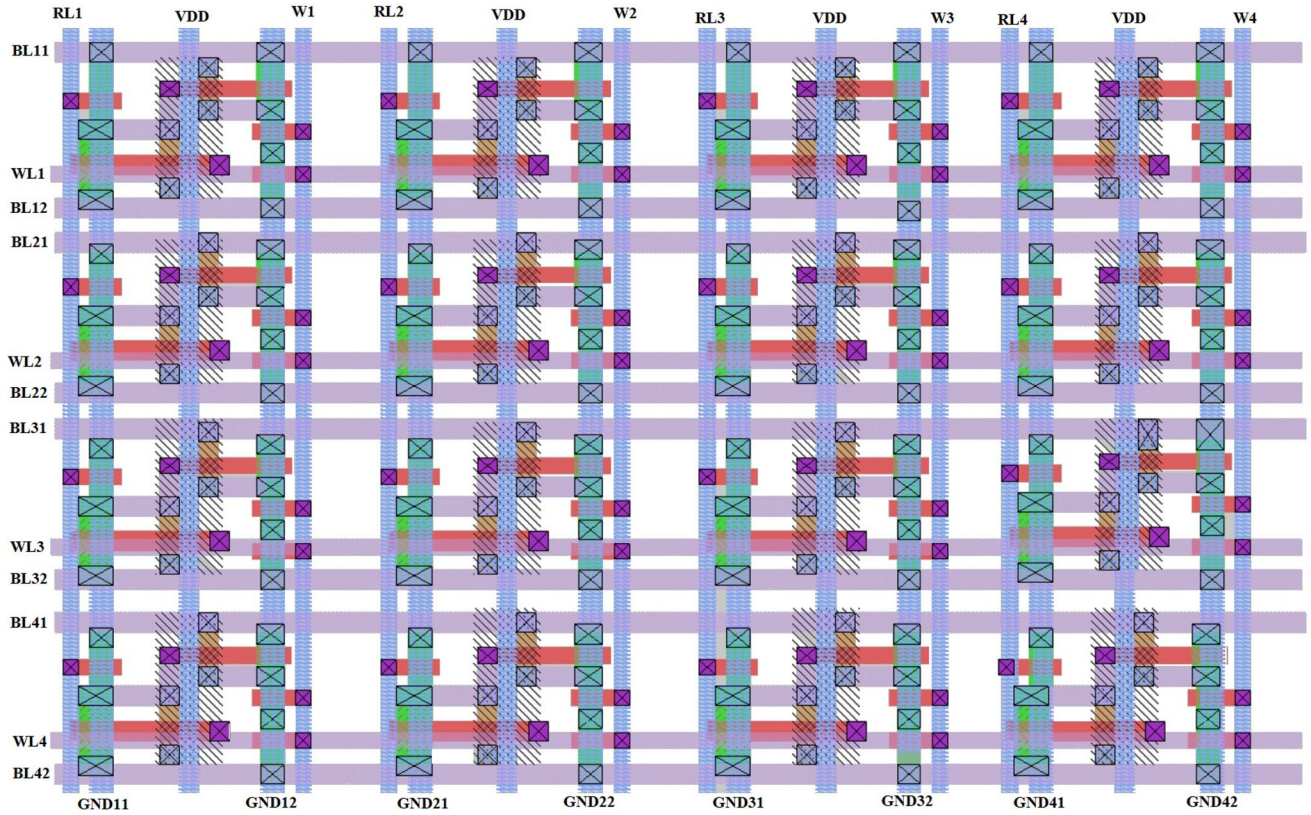


Fig. 14 A 4×4 block of the proposed 7TSE SRAM bit cell showcasing the direction of control signal and supply line routing

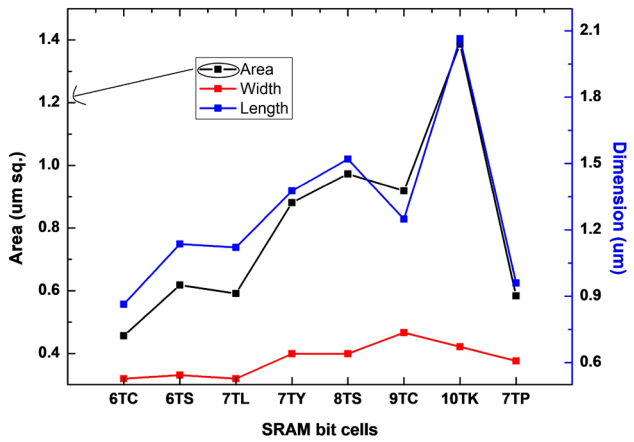


Fig. 15 Graphical comparison for area and dimensions of SRAM cells

preferred memory cell for the industry. But this is only the case for 6TC, the 6TS bit cell though has the same transistor count, but its area is significantly higher at 0.618 μm^2. The 7TP bit cell on the other hand occupies on 0.584 μm^2 which is 27.8% more than 6TC but is 5.55% less than the 6TS cell. Also, it is the smallest SRAM bit cell in terms of area after 6TC. The area for the other bit cells; 6TS, 7TL, 7TY, 8TS, 9TC and 10TK is greater as compared to 7TP cell by 5.87%, 1.31%, 50.8%, 66.67%, 57.36%, and 137.63% respectively.

10 Conclusion

In this work, a single bit-line enabled, single port 7T bit cell is proposed. The feature size for the cell is 32 nm and is simulated at 300 mV supply voltage. The single bit-line architecture of the proposed cell decreases its switching activity factor. The performance of the proposed 7T cell is validated against its 6T, 7T, 8T, 9T, and 10T pre-existing counterparts. The comparison yields that the static performance for the 7TP cell for hold operation is higher than its 6T, 7T, and 10T counterparts, while the read operation is better than all except other 7T cells. Whereas, the write margin for the cell 170 mV, which is fairly balanced and is near the ideal value of $V_{DD}/2$. The cell requires a pulse-width of 30 ns to register a successful write operation, which is fastest amongst the other pre-existing cells in the paper. All the aforementioned merits for the 7TP SRAM bit cell are obtained for minimal power consumption. The tolerance analysis validates that the cell maintains a reliable performance under PVT variations. The aforementioned merits are achieved for area footprint of 0.584 μm^2. This is 27.8% greater than 6TC cell but is

significantly smaller than other counterparts studied. The area for 6TS, 7TL, 7TY, 8TS, 9TC, and 10TK is greater by 5.87%, 1.31%, 50.8%, 66.67%, 57.36%, and 137.63% respectively as compared to 7TP. Thus, it is confirmed that the single port, single bit-line enabled proposed cell is viable when compared to other bit cells.

Data availability Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

References

1. Gavaskar, K., Narayanan, M. S., Nachammal, M. S., & Vignesh, K. (2021). Design and comparative analysis of SRAM array using low leakage controlled transistor techniques with improved delay. *Journal of Ambient Intelligence and Humanized Computing*. https://doi.org/10.1007/s12652-021-03353-z
2. Prasad, G., & Anand, A. (2015). Statistical analysis of low-power SRAM cell structure. *Analog Integrated Circuits and Signal Processing, 82*, 349–358.
3. Mukherjee, R., Saha, P., Chakrabarti, I., Dutta, P. K., & Ray, A. K. (2018). Fast adaptive motion estimation algorithm and its efficient VLSI system for high definition videos. *Expert Systems with Applications, 101*, 159–175.
4. Gavaskar, K., Ragupathy, U. S., & Malini, V. (2019). Design of novel SRAM cell using hybrid VLSI techniques for low leakage and high speed in embedded memories. *Wireless Personal Communications, 108*(4), 2311–2339.
5. Gavaskar, K., & Ragupathy, U. S. (2014). An efficient design and comparative analysis of low power memory cell structures. In *2014 International Conference on Green Computing Communication and Electrical Engineering (ICGCCEE)* (pp. 1–5). IEEE.
6. Lorenzo, R., & Pailly, R. (2020). Single bit-line 11T SRAM cell for low power and improved stability. *IET Computers and Digital Techniques, 14*(3), 114–121.
7. Rathod, A. P. S., Lakhera, P., Baliga, A. K., Mittal, P., Kumar, B. (2015). Performance comparison of pass transistor and CMOS logic configuration based de-multiplexers. In *International Conference on Computing, Communication & Automation* (pp. 1433–1437)
8. Lorenzo, R., & Chaudhury, S. (2017). A novel 9T SRAM architecture for low leakage and high performance. *Analog Integrated Circuits and Signal Processing, 92*, 315–325.
9. Boley, J., Wang, J., & Calhoun, B. H. (2012). Analyzing subthreshold bitcell topologies and the effects of assist methods on SRAM VMIN. *Journal of Low Power Electronics and Applications, 2*, 143–154.
10. Nayak, D., Acharya, D. P., & Mahapatra, K. (2016). An improved energy efficient SRAM cell for access over a wide frequency range. *Solid State Electronics, 126*, 14–22.
11. Kranti, A., & Armstrong, G. A. (2008). 6-T SRAM cell design with nanoscale double-gate SOI MOSFETs: impact of source/drain engineering and circuit topology. *Semiconductor Science and Technology, 23*, 075049.
12. Mittal, P., Negi, Y. S., & Singh, R. K. (2016). A depth analysis for different structures of organic thin film transistors: modeling of performance limiting issues. *Microelectronics Engineering, 150*, 7.
13. Gavaskar, K., Ragupathy, U. S., & Malini, V. (2019). Proposed design of 1 kb memory array structure for cache memories. *Wireless Personal Communications, 109*(2), 823–847.
14. Raj, K., Kumar, B., & Mittal, P. (2009). FPGA implementation of mask level CMOS layout design of redundant binary signed digit comparator. *International Journal of Computer Science and Network Security, 9*(9), 107–115.
15. Kumar, T. S., & Tripathi, S. L. (2021). Process evaluation in FinFET based 7T SRAM cell. *Analog Integrated Circuits and Signal Processing, 109*, 545–551.
16. Lorenzo, R., & Chaudhury, S. (2017). A novel SRAM cell design with a body-bias controller circuit for low leakage, high speed and improved stability. *Wireless Personal Communications, 94*, 3513–3529.
17. Tu, M. H., Lin, J. Y., Tsai, M. C., Jou, S. J., & Chuang, C. T. (2011). Single-ended subthreshold SRAM with asymmetrical write/read assist. *IEEE Transactions on Circuits Systems I: Regular Papers, 57*(2011), 3039–3047.
18. Kumar, B., Kaushik, B. K., Negi, Y. S., Mittal, P., & Mandal, A. (2011). Organic thin film transistors characteristics parameters, structures and their applications. *IEEE Recent Advances in Intelligent Computational Systems, 2011*, 706–711.
19. Anitha, D., Manjunathachari, K., Kumar, P. S., & Prasad, G. (2017). Design of low leakage process tolerant SRAM cell. *Analog Integrated Circuits and Signal Processing, 93*, 531–538.
20. Mittal, P., & Kumar, N. (2020). Comparative analysis of 90nm MOSFET and 18nm FinFET based different multiplexers for low power digital circuits. *International Journal of Advanced Science and Technology, 29*, 4089–4096.
21. Surana, N., & Mekie, J. (2019). Energy efficient single-ended 6-T SRAM for multimedia applications. *IEEE Transactions on Circuits and Systems II: Express Briefs, 66*, 1023.
22. Liu, J. S., Clavel, M. B., & Hudait, M. K. (2017). An energy-efficient tensile strained Ge/InGaAs TFET 7T SRAM cell architecture for ultra-low voltage applications. *IEEE Transactions on Electron Devices, 64*, 2193–2200.
23. Yang, Y., Jeong, H., Song, S. C., Wang, J., Yeap, G., & Jung, S. O. (2016). Single bit-line 7T SRAM cell for near-threshold voltage operation with enhanced performance and energy in 14 nm FinFET technology. *IEEE Transactions on Circuits and Systems I: Regular Papers, 63*, 1023.
24. Singh, P., & Vishvakarma, S. K. (2017). Ultra-low power high stability 8T SRAM for application in object tracking system. *IEEE Access, 6*, 2279–2290.
25. Cho, K., Park, J., Oh, T. W., & Jung, O. K. (2020). One sided Schmitt-Trigger Based 9T SRAM cell for near threshold operation. *IEEE Transactions on Circuits and Systems I: Regular Papers, 67*, 1551.
26. Kulkarni, J. P., & Roy, K. (2012). Ultralow-voltage process-variation-tolerant schmitt-trigger-based SRAM design. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems, 20*, 319.
27. Kulkarni, J. P., Kim, K., & Roy, K. (2007). A 160 mV robust schmitt trigger based subthreshold SRAM. *IEEE Journal of Solid-State Circuits, 42*, 2303–2313.
28. Negi, S., Rana, A., Baliga, A.K., Mittal, P., Kumar, B. (2015). Performance analysis of dual gate organic thin film transistor and organic SR latch application. In: *International Conference on Computing, Communication & Automation*, pp. 1427–1432.
29. Rawat, B., & Mittal, P. (2020). Analysis of varied architectural configuration for 7T SRAM bit cell. In: *4th International Conference on Recent Trend in Communication & Electronics*.
30. Kang, S., Leblebici, Y., & Kim, C. (2015). *CMOS digital integrated circuits*, 3rd ed. New York; McGraw-Hill, pp.204, 422.
31. 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*, 153–163.

32. Mann, R. W., Wang, J., Nalam, S., Khanna, S., Braceras, G., Pilo, H., & Calhoun, B. H. (2010). Impact of circuit assist method on margin and performance in 6T SRAM. *Solid State Electronics, 54*, 1398–1407.
33. Reddy, B. K. N., Ramalingaswamy, C., Nagulapalli, R., & Ramesh, D. (2019). A novel 8T SRAM with improved cell density. *Analog Integrated Circuits and Signal Processing, 98*, 357–366.
34. Rawat, B., & Mittal, P. (2021). A 32 nm single ended single port 7T SRAM for low power utilization. *Semiconductor Science and Technology, 36*, 095006.
35. Singh, J., Mohanty, S. P., & Pradhan, D. K. (2013). *Robust SRAM designs and analysis*. Springer.
36. Ahmad, S., Gupta, M. K., Alam, N., & Hasan, M. (2016). Single-ended Schmitt trigger based robust low power SRAM cell. *IEEE Transaction on Very Large Scale Integration (VLSI) Systems, 20*, 2634.
37. Islam, A., & Hasan, M. (2012). Leakage characterization of 10T SRAM cell. *IEEE Transactions on Electron Devices, 59*, 631–638.
38. Islam, A., & Hasan, M. (2011). Single-ended 6T SRAM cell to improve dynamic power dissipation by decreasing activity factor. *The Mediterranain Journal of Electronics and Communication, 7*, 172–181.
39. Gavaskar, K., & Ragupathy, U. S. (2019). Low power self-controllable voltage level and low swing logic based 11T SRAM cell for high speed CMOS circuits. *Analog Integrated Circuits and Signal Processing, 100*(1), 61–77.
40. Lin, S., Kim, Y.B., & Lombardi, F. (2008). A highly-stable nanometer memory for low power design. *Proc. IEEE International Workshop Design Test of Nano Devices, Circuits Systems* 17.
41. Doorn, T., Maten, E.T., Croon, J., Bucchianico, J.D., & Wittich, O. (2008). Importance sampling Monte Carlo simulations for accurate estimation of SRAM yield. In *ESSCIRC 2008 - 34th European Solid-State Circuits Conference*.
42. Ding, J., & Asenov, A. (2020). The analysis of static random access memory stability under the influence of statistical variability and bias temperature instability-induced ageing. *Semiconductor Science and Technology, 36*, 025008.

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.

Bhawna Rawat was born in Delhi, India in 1993. She received her B. Tech Degree in Electronics and Communication Engineering from Shiv Nadar University, Greater Noida, India in 2015. She received M.Tech degree in VLSI Design from Indira Gandhi Delhi Technical University for Women, Delhi, India. She is presently pursuing Ph.D. at Delhi Technological University, New Delhi, India. Her research interests include low power circuit design, memory circuits in emerging technologies.

Poornima Mittal (Ph.D, M.Tech(H), B.Tech), Member IEEE has published 140 + research papers in international journals and conferences of repute. Her research interest includes design/modeling of flexible electronic devices, memory and low power VLSI circuits. She has published two patents on novel TFT structure and memory design. Also, she has published a Text Book titled "Organic Thin-Film-Transistor Applications: Materials to Circuits" by CRC Press, Taylor & Francis in 2016. She is the reviewer of IEEE transactions and other reputed international journals of IEEE, IET, Elsevier, IOP, Wiley and Taylor & Francis. She has received the research awards in 2012 and 2015 for her dedicated research at Graphic Era University, Dehradun. Also, she has received Commendable Research Awards in 2019, 2020, 2021 and 2022 at Delhi Technological University, Delhi. She is the recipient of Innovator of the Year Award at Uttarakhand State Science and Technology Congress in 2016. She has delivered many expert talks and chaired sessions in the reputed international conferences. She is the life member of many professional societies. She has more than 16 years of academic and research experience. Presently, she is working as Professor in the Department of Electronics and Communication Engineering at Delhi Technological University, Delhi, India..

This article has been accepted for publication in IEEE Transactions on Device and Materials Reliability. 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/TDMR.2023.3247809

A Modified Bypass Circuit for Improved Reliability of PV Module Validated with Real-Time Data

Kashika Baranwal, Prem Prakash and Vinod Kumar Yadav, *Senior Member, IEEE*

***Abstract*—Solar Power is a promising way of electricity generation but it comes with a drawback of low conversion efficiency which is caused by several factors, the most prominent being partial shading. Partial shading leads to mismatch of PV cells, resulting in formation of hot spot. This leads to power dissipation and physical damage to PV cells. Consequently, panel's credibility is called into question. A bypass diode reduces the amount of power dissipated by a shaded cell, but it doesn't entirely prevent hot spot formation. Moreover, bypass diode doesn't bypass every small shade as it has a shaded area threshold at and above which it begins to bypass the shaded module. Whereas, the rise in temperature may also be triggered even by a small shaded area, like the area of falling leaves, which further creates reliability issues. A few hot spot mitigation strategies have been devised; however, it still has room for improvement. This paper presents a modified bypass circuit arrangement for reliability enhancement of PV systems and it is evaluated for a 3×1 PV string. A comparative analysis with bypass diode is carried out in an experimental study. The efficacy of the proposed methodology is verified based on four performance metrics, namely reverse voltage, power loss under partial shading, the sensitivity of the bypass circuit, and the temperature of the hot spotted module. The proposed technique ensures sensitivity enhancement of the bypass circuit towards shade effect; thus, enhancing the reliability of PV system.**

***Index Terms*—Partial Shading, Mismatch Losses, Hot Spot, Bypass Circuit, PV system reliability.**

NOMENCLATURE

The mathematical symbols and their abbreviations used in this paper are as follows:

Variables

P_{Diss}	Dissipated Power
V_{SHADE}	Shaded cell Reverse Voltage
I_{SHADE}	Current generated by the shaded cell
V_D	Diode Forward Voltage
n	No. of series connected cells in a PV Module
i	No. of shaded cells in a PV Module
V_{CE1}	Collector-Emitter Voltage of Series IGBT
V_{CE2}	Collector-Emitter Voltage of Parallel IGBT
V_o	Output Voltage of 3x1 PV array
V_{Mod1}	Voltage across 1st module
V_{Mod2}	Voltage across 2nd module
V_{MPP}	Voltage at maximum power point
I_{MPP}	Current at maximum power point
P_{MAX}	Power at maximum power point

Parameters

V_{OC}	Open-circuit voltage.
I_{SC}	Short-circuit current.
A_S	Percentage area of shaded cell in a PV module

I. INTRODUCTION

IN THE RECENT DECADE, the Solar PV market has grown at an exponential rate. The continuous decrease in the cost of solar PV generation plays a major part in this expansion. These prices can be further reduced by improving the conversion efficiency of the PV system to transform solar energy into electricity. However, due to frequent variation in solar irradiance, the output power of a PV system can decline disproportionally as they are significantly dependent on environmental conditions. It is therefore important to best utilize a PV system to harness solar energy and generate electricity for a sustainable future.

In order to produce a high amount of power at low current, a PV module is made up of 'n' series connected PV cells. In normal conditions, all solar cells are fully illuminated and hence produce power by working as a source. However, if one of the cells is shaded, mismatch occurs in the electrical characteristics of these cells [1]–[3]. This can happen as a result of module failures (such as defective modules, diode failures, junction-box heating, ethylene-vinyl acetate fogging, yellowing, antireflective coating deterioration, severe soiling, etc.) or due to partial shading of the solar cells [4] [5]. The mismatch leads to reduction in shaded cell current and hence the module current. Consequently, the shaded cell becomes reverse biased and begins to behave as a sink. In other words, it starts functioning as a load resistor and heats up owing to power dissipation. This phenomenon is Hot Spot generation. Quite often, the hot spot temperature rises much above the safe working temperature limit, resulting in the cell(s) collapsing and rendering the module inoperable. Thus, the mismatch condition not only raises concerns about module performance but also casts doubt on the module's reliability, ranging from premature cell failure to permanent degradation.

Kashika Baranwal is with the Electrical Engineering Department, Delhi Technological University, Delhi 110042, India (e-mail: kashikabaranwal_2k21phdee10@dtu.ac.in).

Prem Prakash, is with the Electrical Engineering Department, Delhi Technological University, Delhi 110042, India (e-mail: premprakash@dtu.ac.in).

Vinod Kumar Yadav, Senior Member, IEEE, is with the Electrical Engineering Department, Delhi Technological University, Delhi 110042, India (e-mail: vinodkumar@dtu.ac.in).

© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 01,2023 at 07:23:47 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in IEEE Transactions on Device and Materials Reliability. 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/TDMR.2023.3247809

Several researchers have advocated solutions based on static or dynamic reconfiguration of PV arrays to enhance power output in mismatch conditions [6]-[9]. In the recent past, PV voltage equalizer method has also gained popularity as a solution to the problems of mismatch conditions [10]-[13]. This method greatly enhances the power output of shaded module under partial shading condition (PSC) but it is aimed at lowering mismatch losses and thereby increasing the power output under PSC only. And, it did not address the concern of local hot spot creation in the modules due to the mismatch that arises within, due to defective cells, non-uniform deposition of aerosols and PSC arising due to fallen leaves, bird droppings etc.

Hot spot in PV systems is a reliability and safety issue. A PV panel is normally rated to withstand temperatures up to 85 ^{o}C, but hot spots can cause the cell temperature to rise well above the permitted range. The encapsulant and isolative material surrounding the cells may be destroyed if cell surface temperatures exceed 150 ^{o}C [14]. This can create a thermal breakdown [15]–[17], which results in permanent damage to the PV module. Furthermore, it should be noted that frequent hot spot incidents might cause rapid aging [18], increasing the risk of a fault. Hence, the need for an efficient hot spot mitigation circuit escalated. Different strategies have been proposed to solve this problem. The passive approach and the active method are the two basic methods for reducing hot spots in PV panels. In the subsequent paragraphs, the advancements in hot spot mitigation practice over the years are briefly discussed.

The passive method is the conventional method of connecting an anti-parallel bypass diode across a PV module. Though standard bypass diode (SBD) is the most widely utilized technology in the industry to prevent hot spot formation, it alleviates the problem only by limiting the dissipated power and cannot effectively prevent the emergence of a hot spot [14], posing a serious reliability risk for solar PV systems. When a cell is shaded, it is imposed with reverse voltage which helps in activating the SBD. The magnitude of reverse voltage imposed on the shaded cell can be minimized by decreasing the number of PV cells ('n') connected across each SBD. Thus, bypass diodes are more effective in mitigating hot spots for short string lengths. Apparently, lowering ('n') or increasing the number of SBDs in a solar panel is the simplest solution. To obtain the highest possible reduction in the shaded cell's reverse voltage, Meyer and Dyk [19] recommended connecting SBD across each PV cell, while Yoshika et. al. [20] proposed integrating SBD manufacturing with each PV cell. However, the greater the number of SBDs, the greater will be the conduction losses [21], and the greater will be the cost of PV systems. Consequently, the preceding alternatives did not gain prominence.

However, when an SBD is turned on, it generates a forward voltage that increases the reverse voltage imposed on a shaded cell while dissipating more power. To reduce this imposed voltage, Acciari et al. in [22] and d'Alessandro et al. in [23] respectively proposed the application of Cool Bypass Switch and PNP transistor as an active bypass path. An active switch method minimizes the resultant power loss of the PV systems when in bypass mode. It is therefore marginally better than SBD for hot spot mitigation. This method is called the active bypass method.

The techniques outlined above are aimed at lowering mismatch losses and hence increasing power output in mismatch situations; however, they did not raise the concern of local hot spot creation. Only a few researchers have noticed the rising magnitude of shaded cell reverse voltage, which may exceed the diode's breakdown voltage. Daliento et al. in [24] and Ghosh et al. in [25] respectively advocated using a series MOSFET and a series IGBT with a prevalent bypass diode (SIWPD). The purpose of connecting an active switch in series with a PV module was to lower the magnitude of reverse voltage. When a cell is shaded, the forward voltage of the SBD adds to the reverse voltage while being in conduction mode. Whereas series-connected active switch exhibits a significant voltage drop which is subtracted from the reverse voltage. This method is called active mitigation strategy.

Though researchers have experimentally verified that the low drop active devices are a better alternative to passive diodes, they have only been successful in lowering the intensity of the problem and have not entirely eliminated it. The cutting-edge solution for this problem evolved by combining the advantages of both active bypass and active mitigation strategy of hot spotted module. This method is called active hot spot detection and mitigation. This method utilizes two active switches working in tandem, wherein one switch is connected in series with the PV module and the other switch is connected across the PV module. Kim and Krein [26] proposed a MOSFET-based dual switch device based on this technology. Dhimish et al. [27] investigated this approach proposed by putting it into practice. In both the examinations, it was found that the temperature of the shaded cell lowered when the series MOSFET was switched OFF and the bypass MOSFET was switched ON simultaneously. However, Guerriero and Daliento in [28] proposed, the application of a pair of MOSFETs controlled by an oscillator. This lowers the amount of current flowing through the shaded cell by adjusting the bypass current through diode while being in conduction mode.

Solutions devised so far to abate the reverse voltage of the shaded cell and alleviate the hot spot problem have not been able to entirely eliminate the problem. Tang et al. [29] accomplished this by using a switch-controlled resistor divider in series with the PV module with prevalent SBD. In normal condition, parallel-connected switch turns on to short circuit the divider resistor. This puts the diode in the cut-off state because of reverse biasing. Under shaded conditions, switch is turned OFF and current bypasses via diode. To make the reverse voltage of the shaded cell zero, a significantly large value resistance is required. This generates a significant amount of power loss as well.

SBD has a threshold of the shaded area at and above which it starts to conduct in forward direction. In other words, it doesn't bypass every small shade which would lead to even severe hot spot creation. The latest solutions offered so far has implemented SBD in the bypass path for reducing the influence

© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 01,2023 at 07:23:47 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in IEEE Transactions on Device and Materials Reliability. 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/TDMR.2023.3247809

of local hot spot. However, none of them have worked for the sensitivity enhancement of bypass circuit to bypass the hot spot generated by smaller shades. For this small shaded area, the magnitude of imposed reverse voltage on the shaded cell is so high that the cells connected to SBD undergoes avalanche breakdown and subsequently gives rise to hot spot.

Therefore, it becomes important to develop a more efficient bypass circuit, which is not only capable of reducing the magnitude of reverse voltage but also enhances its sensitivity towards a smaller shade. This study re-examines the hot spot problem in PV systems and improves its reliability by evolving a more sensitive bypass circuit path in the event of shade. In this study, a potential modified bypass circuit arrangement (MBCA) is presented, which not only reduces the reverse voltage and the temperature of the hot spotted cell but also reduces the power loss of the bypassed PV module under shaded conditions.

Following the Introductory Section, the remaining paper is organized as follows. In Section II, the mechanism of PV systems is described under shaded condition, followed by reverse breakdown conditions. Section III discusses the proposed hot spot mitigation circuit. The experimental examination of hardware setup and obtained observations are illustrated in Section IV. Section V focuses on the comparative discussion of the acquired results with respect to SBD and SIWPD. Conclusions are then drawn in Section VI.

II. PV System Mechanism Under Shaded Condition

For one cell under PSC, the available power of all the $'n'$ cells get reduced significantly, if not protected with SBD. This is because there is no path for the currents generating from $(n-1)$ illuminated PV cells. Furthermore, this current applies the reverse voltage across the n^{th} shaded cell and hence dissipates power, resulting in hot spots and PV cell failure. The amount of power dissipation can be quantified as (1).

$$P_{Diss} = V_{SHADE} \times I_{SHADE} \tag{1}$$

When protected with SBD, the generated power of $'n'$ PV cells can be improved as compared to previous case. This is because when the photocurrent of n^{th} shaded cell becomes smaller, the excess current coming from $(n-1)$ illuminated cells is forced to flow through SBD so that the shaded cell does not act as a load. So, the reverse voltage of the n^{th} shaded cell is dominated by the SBD's forward voltage V_D, which usually ranges from 0.5 to 0.7 V for silicon diodes. Thus, the power dissipation is reduced significantly. Therefore, the application of in-built SBD with PV modules has been accepted as an industry standard. The amount of power dissipation with SBD can be quantified as (2).

$$P_{Diss} = V_D \times I_{SHADE} \tag{2}$$

A shaded cell has a reduced value of short circuit current I_{SC}. To bypass the module current, the shaded cell will operate in reverse bias, thereby making SBD active. Consequently, the voltage across the whole module matches with the forward voltage drop V_D of the diode D (which may vary from 0.35-0.7V for Schottky Diodes). The total module voltage is the sum of the various operating voltages of each cell including the reverse-biased cell. If the total module voltage is negative, the SBD turns on and the shaded cell will operate with reverse voltage. From Fig. 1, the magnitude of the shaded cell reverse voltage can be calculated by applying KVL on the close loop formed by $'n'$ series-connected cells with SBD as in (3):

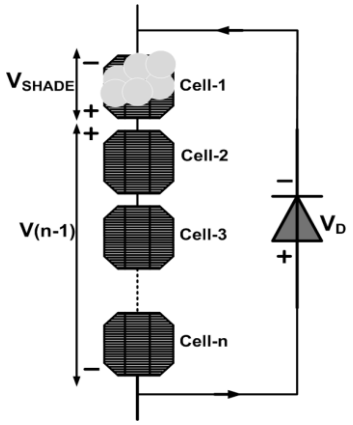


Fig. 1. PV protection scheme with SBD.

$$\left|V_{SHADE}\right| = \sum_{k=1}^{n-i} V_{OC,n} + V_D \tag{3}$$

Considering the cut-in voltage of SBD 0.5 V (i.e., for Schottky diode) and open-circuit voltage of each cell 0.6 V, the reverse voltage becomes 21.5 V for 36 cells module. PV cells appear to be safe at these voltage levels, as the reverse breakdown of simple p-n junction usually occurs in the -100 to -50 V range [30]. PV cells, on the other hand, have different reverse breakdown characteristics than simple p-n junctions due to the additional manufacturing progresses they go through to become a PV cell. The avalanche breakdown occurs between 15 and 20V for alkaline texturized PV cells [31], but only at 13V for acidic texturized PV cells [32]. The reverse breakdown characteristic of silicon PV cells was experimentally examined and categorized into three types [33]. The early breakdown which is caused due to Al contamination, occurs below −5V is designated as Type-1 breakdown. This results in the linear rise of current and is not fatal for PV cells. Type-2 breakdown is caused due to defect-induced by FeSi2 or other precipitates lying in crystallographic grain boundaries. It occurs at voltage between −11 to −7V which results in weak exponential rise of current but does not impact the cell fatally. Type-3 breakdown is the most severe avalanche breakdown, which is caused by etch pits or other sharp kinks in the p-n junction plane. It occurs at voltage between (−18 to −13V) which results in strong exponential rise of current and does irreversible damage to the PV cells [33].

Global breakdown voltage of PV cells reduces even further for higher doping concentration. The lower value of reverse breakdown voltage may lead to hot spot generation which could be detrimental for a PV cell. Hence, the requirement to reduce reverse voltage grew. Since this hot spot can be generated even through a small shaded area, the solution should be effective

© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 01,2023 at 07:23:47 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in IEEE Transactions on Device and Materials Reliability. 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/TDMR.2023.3247809

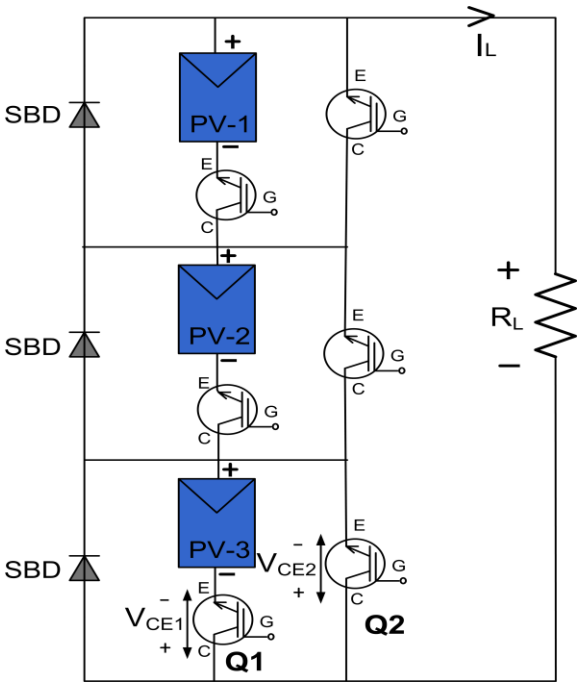


Fig. 2. Schematic of the proposed hot spot detection and protection method.

enough to bypass even the little shade. To alleviate this problem and boost the PV system reliability, a modified bypass circuit is proposed in this paper.

III. Proposed Hot Spot Mitigation Circuit

The circuit proposed in this paper is evolved from the circuit presented in [25] which was primarily concerned with minimizing the reverse voltage of the shaded cell. To emphasize the necessity of a more sensitive bypass circuit and reverse voltage reduction, a dual IGBT-based hot spot mitigation circuit is adopted in this paper instead of a single IGBT-based bypass circuit. The proposed circuit not only reduces the reverse voltage of the shaded cell but also enhances both the sensitivity of the bypass circuit and the peak power output of the bypassed module under shaded conditions.

The experiment was carried out on a 3x1 PV string, wherein each PV module has 36 cells connected in series and is individually protected by an SBD. In all, three modules of the suggested hot spot detection and protection system were developed, one for each PV module. The proposed system was connected to each PV module, as shown in Fig. 2. Two identical STGF10NB60SD IGBT switches were employed in the controller circuit to implement the proposed MBCA. Series-connected IGBT (Q1) was connected to PV modules in series and was normally closed (NC), whereas parallel-connected IGBT (Q2) was connected to PV modules in parallel and was normally open (NO). The controller circuit works on an inversion mechanism for two employed IGBTs (i.e.; when one is in ON-state, the other will be OFF and vice-versa). To compare the two bypass paths (i.e., SBD and Q2) in the same PV module, a separate switch was used in the controller circuit for turning ON and OFF the bypass path Q2 as and when needed.

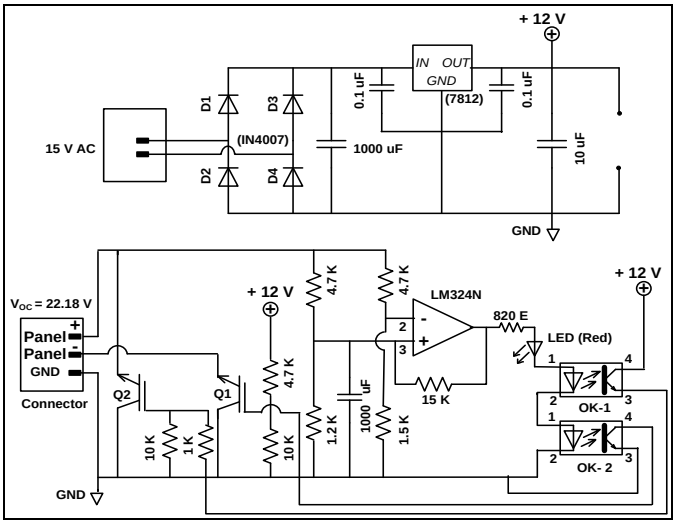


Fig. 3. The hot spot detection and protection circuit.

The complete controller circuit is shown in Fig. 3, which has separate drive circuit for both Q1 and Q2. The driving circuits of Q1 and Q2 are controlled via Optocoupler 2 (i.e., OK-2) and Optocoupler 1 (i.e., OK-1) respectively. The OK-1 and hence OK-2 is further driven by the output of the comparator IC. Under normal condition, the potential at pin-2 and 3 of the comparator IC are equal, which results in low differential output. Due to the low comparator output, OK-1 and hence OK-2 remain OFF in normal condition. When partial shading occurs, the potential at pin-2 and 3 drops. However, because a capacitor is employed to store the previous value of pin-3, it would hold onto that value for a period of time equal to the capacitor's time constant. However, the potential at pin-2 will change instantly due to the absence of capacitor. This would create a significant difference between the two input pins of comparator to make its output high. This high output of comparator will trigger to activate OK-1 and hence OK-2 via LED (red). The pin-3 of OK-1 is connected to drive circuit of Q2, providing it a conduction path and pin-3 of OK-2 is connected to drive circuit of Q2 providing current an alternative short circuit path to flow. So, the IGBT Q2 will now be switched-on and IGBT Q1 will be switched-Off. A red LED has been added to indicate the switching of Q1 and Q2 in the event of PSC. The 12 V dc supply required for driving circuits may be given by the 3×1 string voltage or externally. In the given circuit, the 12 V dc supply was generated from an external ac supply, followed by a rectifier and a filter circuit as shown in the upper part of Fig. 3. Different transformer tappings were used to supply the controller of each module.

IV. Experimental Validation Of Proposed MBCA

The Dual IGBT-based bypass circuit was tested on a 3×1 PV string, wherein all three PV modules were put under different shading states to experimentally test the efficacy of the proposed bypass circuit. The tests were carried out under real world conditions, with an average natural irradiance of 850-950 W/m^2 and temperatures ranging from 25 °C to 35 °C. Fig. 4 depicts the entire hardware setup. The string was connected to

© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 01,2023 at 07:23:47 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in IEEE Transactions on Device and Materials Reliability. 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/TDMR.2023.3247809

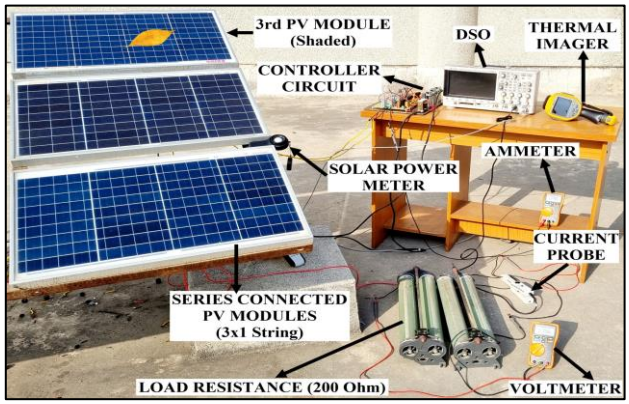


Fig. 4. Hardware setup under partial shading condition.

TABLE I
SPECIFICATIONS OF A PV MODULE UNDER STC (at 1000 W/m^2, 25 °C)

Parameters	Values
n	36 cells
V_{MPP}	18 V
I_{MPP}	2.22 A
P_{MAX}	40 W
V_{OC}	22.18 V
I_{SC}	2.4 A

a variable resistive load of 200 ohms, which was tuned close to the open circuit by keeping the resistance at its maximum value. White paper sheets were used to partially shade the cells. The electrical characteristics of each PV module employed in the setup is shown in Table I.

A. Reverse Voltage Reduction

In normal operation, series-connected Q1 is in the ON state and thus behaves like a short circuit path (as the on-state resistance of IGBT and hence V_{CE1} is very small), while the parallel-connected Q2 is OFF. In the event of partial shading, Q1 becomes OFF, while Q2 becomes ON for providing a bypass path to the shaded cell module. The collector current flowing through series-connected Q1 becomes close to zero, and hence V_{CE1} value becomes higher at this operating point, which is deducted from the reverse voltage of the shaded cell. In mathematical terms, the shaded cell reverse voltage deduced to equation (4), derived from Fig. 5(a).

$$\left|V_{SHADE}\right| = \sum_{k=1}^{n-i} V_{OC,n} - V_{CE1} + V_{CE2} \qquad (4)$$

Here, the value of V_{CE1} is larger than the value of V_{CE2}. Therefore, a reduction in the magnitude of the reverse voltage of the shaded cell is obtained when compared to equation 3. The purpose of Q1 is to share a portion of the voltage distributed by the illuminated cells during bypass conditions and, thus lower the reverse voltage across the shaded cell while only slightly affecting the solar panel's operation (under normal conditions). Whereas, Q2 is employed across the PV system to offer a bypass path in the event of partial shading. The purpose of this IGBT is to allow a bypass path to current when one of the modules is under shaded conditions and thus, preventing the module from hot spot generation. With SBD in the bypass path, the shaded cell reverse voltage deduced to equation (5), derived from Fig. 5(b).

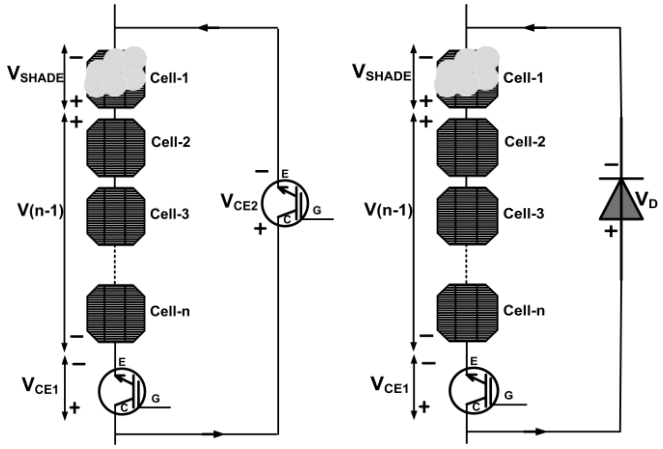


Fig. 5. PV protection scheme with (a) MBCA (b) SIWPD.

$$\left|V_{SHADE}\right| = \sum_{k=1}^{n-i} V_{OC,n} - V_{CE1} + V_D \qquad (5)$$

When comparing equation 4 (Q2 bypass path of MBCA) with respect to equation 5 (Diode bypass path of SIWPD), the value of potential V_{CE2} was found lesser than the diode's forward potential V_D ($i.e., V_{CE2}$ = 0.589 Volts, V_D = 0.625 Volts). This is due to the presence of low drop active device in former instead of passive device in latter bypass path. Consequently, the shaded cell's reverse voltage would be slightly lower in the first case. Thus, MBCA performs better than both SBD and SIWPD in reducing the magnitude of reverse voltage.

It is understood from (4) and (5) that for a PV module, the magnitude of reverse voltage would be lesser if the number of shaded cells (i) is larger. This will likewise impact the temperature of the shaded cell in the same manner, as can be seen in Section IV (D).

B. Sensitivity Enhancement

The hardware setup included a Digital Storage Oscilloscope (DSO) for comparing and validating the sensitivity of the SBD and MBCA. For the sensitivity comparison, two shading scenarios were studied - masking two cells and masking one cell respectively. The string was linked to a variable resistive load of 200 ohms, which was kept close to open circuit condition by keeping the resistance at its maximum value. On the third PV module, different operating phases were examined at a particular shaded area to assess the sensitivity of the two bypass circuits. The respective variation in electrical entities were recorded in DSO, wherein the scale of voltage, power and current are mentioned at the bottom of Fig. 6 and 7.

To begin, two cells in the third module were masked, and the SBD, in conjunction with the MBCA, was found equally effective in bypassing the shading effect created by the two cells, as shown in Fig. 6. This was performed in four phases. In the first phase, third module's parameters were recorded in an unshaded situation (i.e., Normal Condition), with the MBCA switched OFF. The third module was then partially shaded in the second phase, with the area encompassing less than two cells. This is because SBD could not bypass this shaded area, leading to hotspot formation. Then in the third phase, shading

© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 01,2023 at 07:23:47 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in IEEE Transactions on Device and Materials Reliability. 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/TDMR.2023.3247809

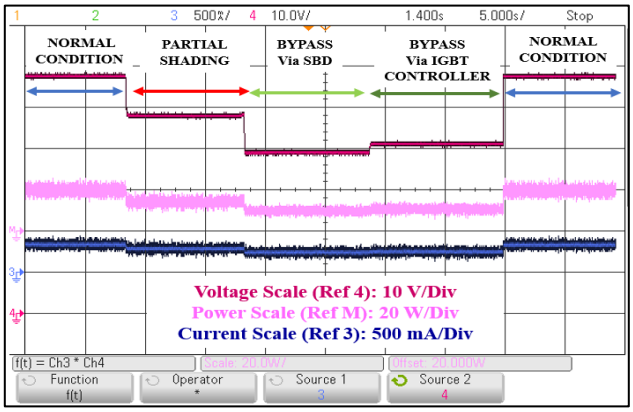


Fig. 6. Electrical entities (Voltage, Power, Current) with two shaded cells.

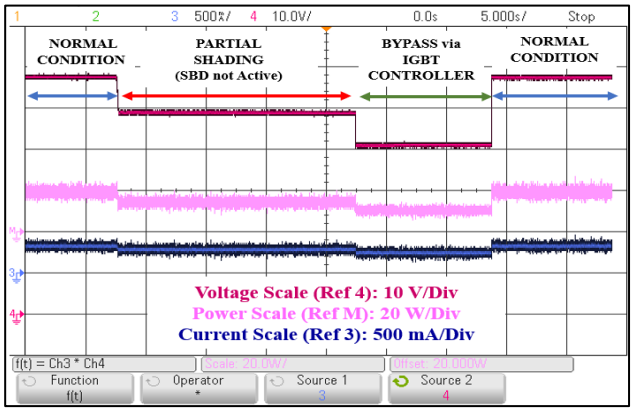


Fig. 7. Electrical entities (Voltage, Power, Current) with single shaded cell.

TABLE II
VARYING ELECTRICAL PROFILE WITH TWO SHADED CELLS

State of 3rd Module	Controller Switch	String Voltage (Volts)	Current (Amps)	Power Generated (Watts)
NC	-	57	0.276	15.73
PSC	OFF	48.1	0.233	11.21
Bypass via Diode	OFF	39.96	0.189	7.55
Bypass via Q2	ON	40.18	0.191	7.67
Restoring NC	-	57	0.276	15.73

TABLE III
VARYING ELECTRICAL PROFILE WITH SINGLE SHADED CELL

State of 3rd Module	Controller Switch	String Voltage (Volts)	Current (Amps)	Power Generated (Watts)
NC	-	57	0.276	15.73
PSC	OFF	48.8	0.233	11.37
No Bypass via SBD	OFF	48.8	0.233	11.37
Bypass via Q2	ON	40.17	0.191	7.67
Restoring NC	-	57	0.276	15.73

was done at SBD's working threshold of two cells to activate it, however, the MBCA switch had been turned OFF during these first three phases. Finally in the fourth phase, the MBCA switch was turned ON to bypass the shading effect of two cells via Q2. Later, the shade effect was removed, allowing the third module to resume normal operation and the controller to reset. The variation in electrical properties of the entire string (i.e., Voltage, Power, and Current measurements) during these phases obtained with 2 shaded cells is shown in Table II.

Following that, single cell of the third module was masked instead of two. This again, was performed in the same four phases as performed for two cells and it was found that the SBD was ineffective in bypassing the shading effect created by single cell, as shown in Fig. 7. In this case, the third module was partially shaded with the area encompassing one cell which could not be bypassed by SBD. The MBCA was then switched ON to test the efficiency of Q2 against this area, and it worked effectively to bypass even this small shaded region. The variation in electrical properties of the entire string (i.e., Voltage, Power, and Current measurements) during different states obtained with 1 shaded cell is shown in Table III. By comparing Fig. 6 and 7, it can be inferred that the SBD failed to bypass the shade of a single cell, whereas the Q2 succeeded in doing the same. In fact, the sensitivity of both the circuits was tested for even smaller shaded areas and it was found that the bypass via Q2 was successful even when 80 % of one cell was masked whereas bypass via SBD happened only when 2 or more cells of the same module were masked. Thus, it can be said that the sensitivity of MBCA is more than double (approximately 150 % more) the sensitivity of SBD.

To validate this further, the characteristic of the SBD has been plotted by varying a load resistance of 200 ohms from its maximum to minimum value, shown in Fig. 8. Initially, one cell was masked and it was observed that shading causes nearly no

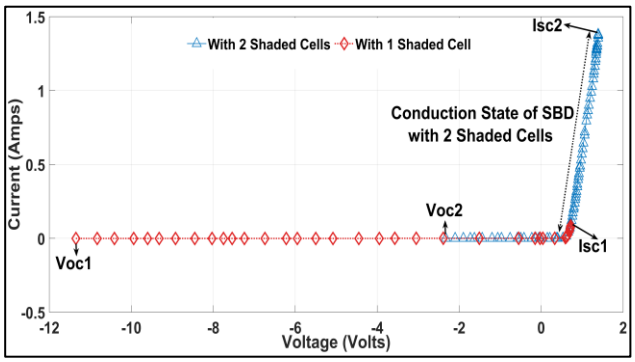


Fig. 8. SBD characteristic with single and two shaded cells.

current to flow through the SBD. Following that, two cells were masked and it was found that a significant amount of current began to flow, indicating the diode's conduction state unlike with one masked cell. V_{OC1} and V_{OC2} are the module's open-circuit voltages recorded during one and two cell shading conditions respectively, and; I_{SC1} and I_{SC2} are the module's short circuit currents recorded under the same shading scenarios.

C. Power Loss Reduction

The power curves of the 3x1 PV array were drawn using real time experimental data with different operating states under partial shading, shown in Fig. 9. During shading, PV array generates multiple maximum power points in the P-V Curve (i.e., local peaks and global peaks) because of the difference in the received irradiance by the modules. For a 40 W_P PV module with one SBD connected across each module, various shaded operating states were created and examined. Following are the four states that were being considered.

State 1 (NC): All three modules were fully illuminated.

© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 01,2023 at 07:23:47 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in IEEE Transactions on Device and Materials Reliability. 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/TDMR.2023.3247809

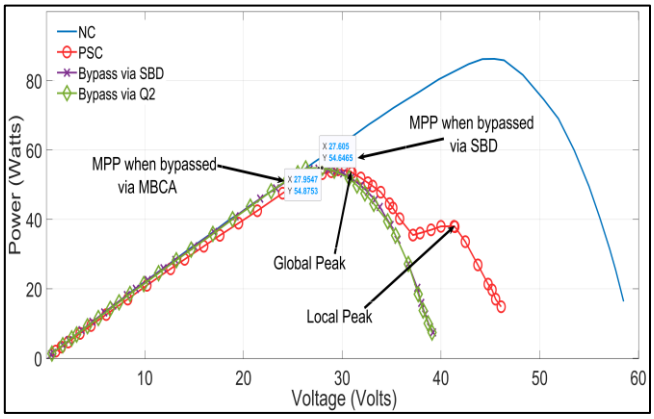


Fig. 9. P-V curve of 3x1 PV string under states 1, 2, 3, and 4.

State 2 (PSC): With MBCA switched OFF, the first and second module were exposed 100 percent to solar radiation, while one cell of the third module (*i.e.*, 2.77 percent) was exposed to shading. (HOT SPOT Generation Mode).
State 3 (Bypass via SBD): The first and second modules were exposed 100 percent to solar radiation, while two cells of the third module (*i.e.*, 5.55 percent) were exposed to shading. MBCA switch was turned OFF in this situation, thus bypassing the third Module via SBD.
State 4 (Bypass via Q2): The first and second modules were exposed 100 percent to solar radiation, while two cells of the third module (*i.e.*, 5.55 percent) were exposed to shading. MBCA switch was turned ON in this situation, thus bypassing the third module via an IGBT-based bypass circuit.

The output voltage of the PV string when the third module is shaded and bypassed would remain same irrespective of the shading pattern on the module. This can be calculated as the sum of the voltage of the first module, second module, and the voltage drop of the bypass path Q2, when bypassed via MBCA, shown in equation (6).

$$V_o = V_{Mod1} + V_{Mod2} - V_{CE2} \quad (6)$$

If SBD had been used in the bypass path instead of Q2, equation (6) would have modified into equation (7).

$$V_o = V_{Mod1} + V_{Mod2} - V_D \quad (7)$$

Here, the value of potential V_{CE2} is slightly lesser than the diode's forward potential V_D. Consequently, the output power is higher with Q2 bypass path. As evident from Fig. 9., the peak power generation is 54.88 W and 54.65 W when bypassed via MBCA and SBD respectively. Relative peak power enhancement when bypassed via MBCA instead of SBD is 0.42 %. Thus, it can be inferred that the MBCA is capable to reduce the conduction loss by 0.419 % when compared with SBD.

D. Thermal Analysis

Thermographic images of the shaded cell module were obtained by changing its shaded area as well as its shading pattern, as shown in Fig. 10. The different shading patterns chosen are influenced by natural shading processes that are mainly caused by fallen leaves, shade of the adjacent building, or a nearby tree, dust deposition etc. Different patterns were created on different PV modules to check the effectiveness of proposed MBCA on the whole PV string. As shown in Fig. 11, the image of radiant LED in each controller circuit was captured when its corresponding module was shaded and bypassed.

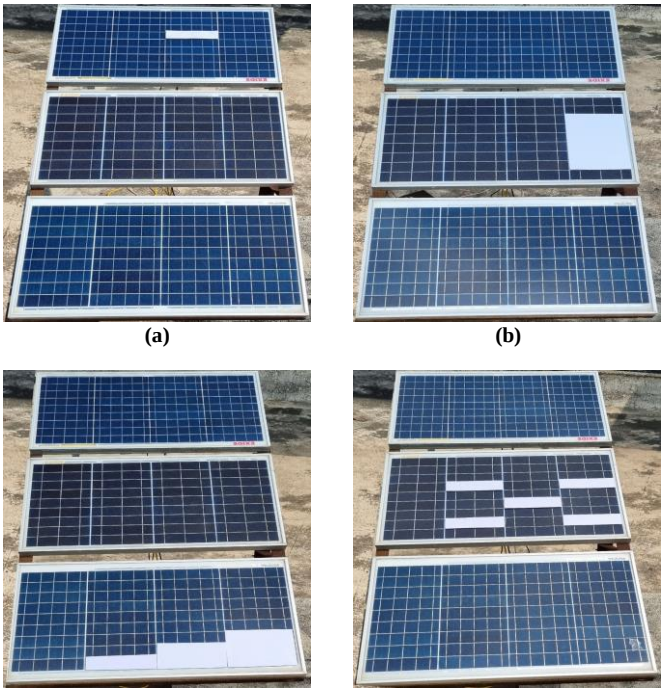


Fig. 10. Different shading patterns considered (a) shading from fallen leaf (b) shading from adjacent pole or building (c) shading from a tree branch (d) random pattern due to non-uniform dust deposition or soiling.

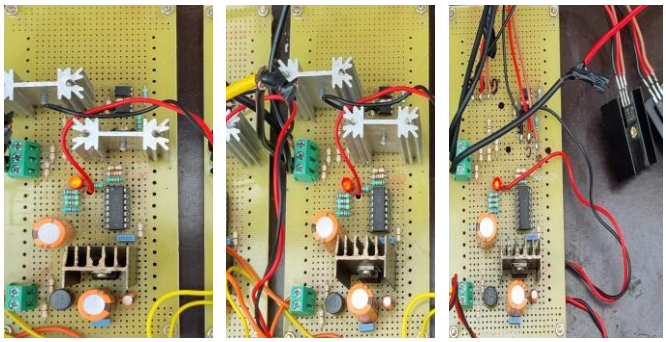


Fig. 11. Active controller module (LED glowing) corresponding to each PV module for different shading patterns.

TABLE IV
PERFORMANCE COMPARISON OF SBD & MBCA

Shading Pattern	Status of SBD	Status of MBCA	Remarks on Shaded Cell Temperature
Single Cell	No Bypass, leads to hotspot formation	Bypass, lowers the temperature up to 54.9 ºC	MBCA lowers down the temperature by 19.03 %
Column Shading	Bypass, lowers the temperature up to 43.9 ºC	Bypass, lowers the temperature up to 40.9 ºC	MBCA lowers down the temperature further by 6.83 %.
Ladder Shading	Bypass, lowers the temperature up to 38.3 ºC	Bypass, lowers the temperature up to 38.1 ºC	MBCA lowers down the temperature further by 0.52 %.
Random Shading	Bypass, lowers the temperature up to 38.7 ºC	Bypass, lowers the temperature up to 37.8 ºC	MBCA lowers down the temperature further by 2.32 %.

© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 01,2023 at 07:23:47 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in IEEE Transactions on Device and Materials Reliability. 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/TDMR.2023.3247809

Fig. 12(a) Hot spot generation (with single shaded cell).

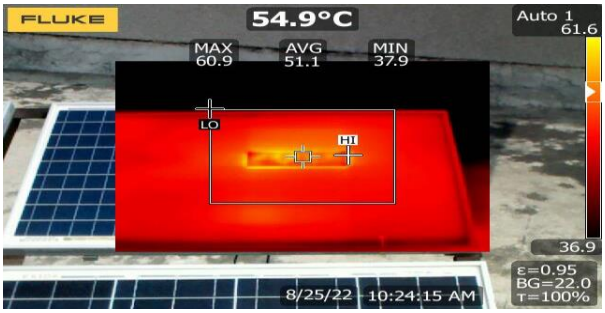


Fig. 12(b) Bypassed via MBCA with single shaded cell.

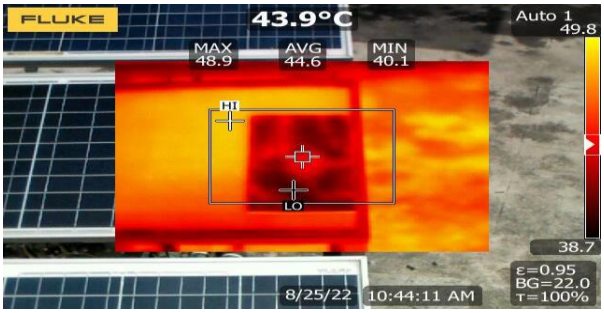


Fig. 12(c) Bypassed via SBD (Column Shading).

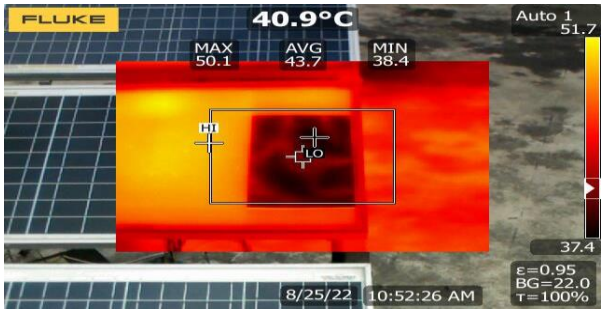


Fig. 12(d) Bypassed via MBCA (Column Shading).

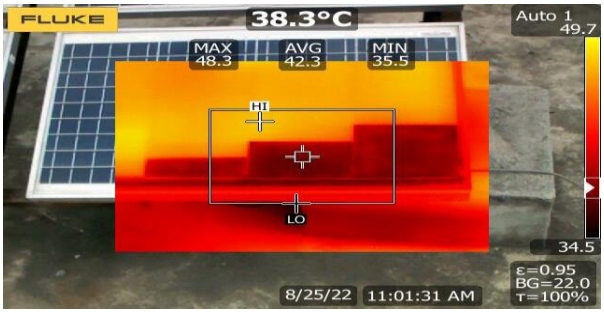


Fig. 12(e) Bypassed via SBD (Ladder Shading).

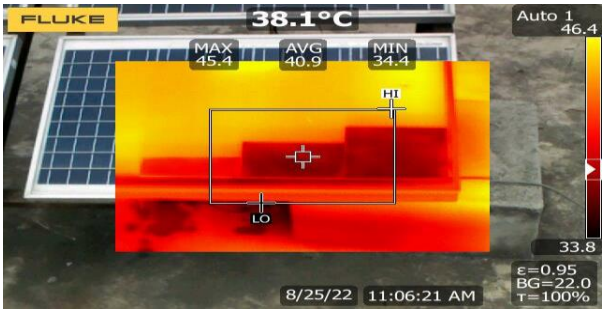


Fig. 12(f) Bypassed via MBCA (Ladder Shading).

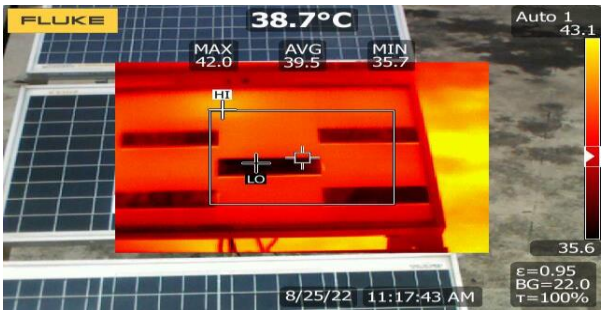


Fig. 12(g) Bypassed via SBD (Random Shading).

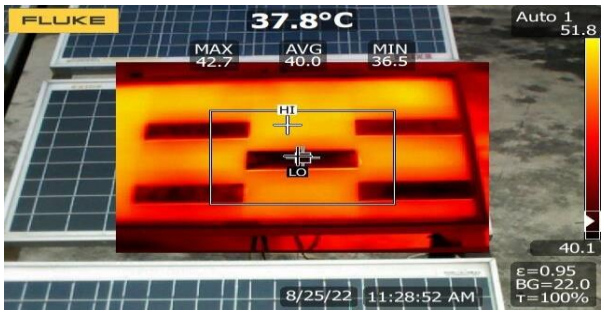


Fig. 12(h) Bypassed via MBCA (Random Shading).

Initially, a single cell of the SBD integrated module was covered and the controller was switched OFF. It was observed that the temperature of the shaded cell increased up to 67.8 ^{o}C in 35 minutes (with a maximum temperature of 83.4 ^{o}C), shown in Fig. 12(a). This shows the inefficiency of SBD to bypass the effect of a small shaded area, resulting in the formation of hot spot. The module temperature would have crossed the safe operating limit if the module had been left in this condition for a prolonged period. As soon as the MBCA is switched ON, the hot-spot temperature starts to decrease steeply, as shown in Fig. 12(b). This temperature was lowered up to 54.9 ^{o}C in approximately 15 minutes, making it efficient to bypass even the small shade. Then, the performance of both bypass circuits SBD and MBCA were compared by switching-OFF and ON the controller, for column, ladder and random shading patterns, (see Table IV). The thermal results obtained by SBD and MBCA are shown in Figs. 12 (c)-(d) for column shading, in Figs. 12 (e)-(f) for ladder shading, in Figs.12 (g)-(h) for random shading, respectively. As evident from the thermographic images, MBCA reduced the temperature better as compared to SBD. Since MBCA also succeeded in bypassing the shading impact of a small shaded region, this makes it more reliable than SBD. Thus, it is deduced that utilizing an IGBT in the bypass path improves the reliability of the PV system even better.

© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 01,2023 at 07:23:47 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in IEEE Transactions on Device and Materials Reliability. 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/TDMR.2023.3247809

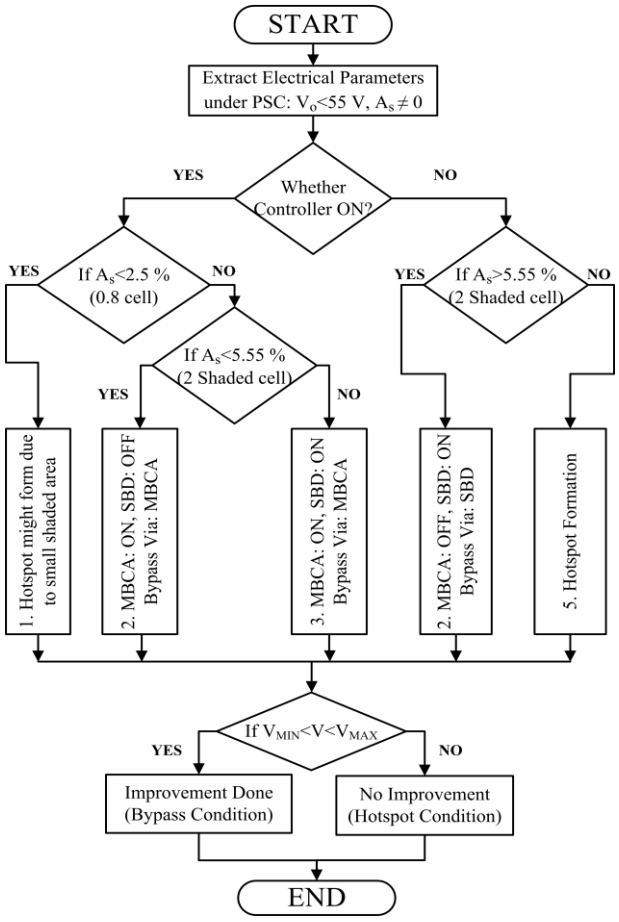


Fig. 13. Flow chart of the proposed bypass mechanism.

E. Process Flow

The working operation of the 3x1 PV string under different shading states is summarized with the help of a flow chart in Fig. 13. Under normal conditions, A_S = 0 and the value of V_o is 57 V or more at good illumination. This value of V_o varies from 55 V to 57 V under PSC when shaded with area encompassing lesser than two cells, and not bypassed by SBD. The process involved in the flow chart depicts the health status of the PV system at the end. The string voltage under shaded and bypassed condition varies from a minimum of V_{MIN} = 37 V and maximum of V_{MAX} = 40 V when bypassed via SBD and MBCA. These voltages are therefore chosen for deciding if the PV system is operating under healthy working conditions or not.

V. Result And Discussion

A modified hot spot mitigation strategy is proposed in this paper to enhance its sensitivity for efficient hot spot mitigation which employs two IGBTs, one in series and other in parallel with a PV module, working on inversion mechanism. This makes the PV system responsive even to a small shaded area. To check the effectiveness of the proposed MBCA, all three PV module was put to test with different shading patterns and is compared with both SBD (Fig. 1) and SIWPD (Fig. 5(b)). To perform this comparison on same PV module, an additional switch was added to MBCA. A quick review of comparison with the two circuits is shown in Table V.

TABLE V
Comparative Analysis of MBCA With SBD And SIWPD

Performance Indices	Percentage change in parameters obtained by using MBCA with respect to:		Remarks about MBCA	Discussed in
	SBD	SIWPD		
Reverse Voltage	Reduced	Reduced	Better than both the circuits	Section-IV (A)
Sensitivity Towards Shade	Increased by 150 %	Increased by 150 %	Better than both the circuits	Section-IV (B)
Peak Power	Increased by 0.42 %	Increased by 0.42 %	Better than both the circuits	Section-IV (C)
Temperature of the Shaded Cell Module	Reduced	Reduced	Better than both the circuits	Section-IV (D)
Reliability	**More**	**More**	**Better**	**Result**

Since peak power and sensitivity towards shade are the parameters that are determined by the nature of bypass element, therefore, SIWPD would show the same behavior as that of SBD for these two parameters. As evident from Table V, the peak power generation of the shaded module and sensitivity of the bypass path towards shade were enhanced by 0.42 % and 150 % respectively when comparing MBCA with SBD. Therefore, the MBCA outperforms SIWPD in peak power enhancement of shaded PV module and sensitivity enhancement of bypass path as well. Also, the magnitude of reverse voltage and hence, the temperature of hot spotted module reduction done by MBCA is more than that of both SBD and SIWPD. The thermal analysis performed in Section-IV (D) shows that SBD failed to bypass the impact of severe hot spot generated by the area covered by single shaded cell. Moreover, the temperature of the hot spotted module was reduced further when connected to MBCA instead of SBD. Thus, it can be inferred that MBCA performs better in all the four-performance metrics when compared with both the circuits.

However, it is worth contemplating that the series IGBT (Q1) used in the proposed circuit experiences a marginal conduction loss under normal working condition. This IGBT is rated at 16 A, 600 V, the $V_{CE(Sat)}$ value is also low (i.e., 0.876 V for a load of 200 Ω and string current of 0.27 A). So, the power loss is quantified as:

$$P_L = (V_{CE})_{Sat} \times I_{String} = 0.236W \qquad (8)$$

Hence, the loss in the IGBT accounts for only 0.59 % power of a 40 W_P Solar Panel. Moreover, this IGBT is suitable even for larger PV array of rating up to 10 A, 600 V. And hence in the real-world applications, the percentage of power loss would be negligible.

VI. Future Scope

As future work, the marginal conduction loss of the Q1 during normal condition can be minimized using parallel circuits across it. The parallel path will remain active in normal

© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 01,2023 at 07:23:47 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in IEEE Transactions on Device and Materials Reliability. 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/TDMR.2023.3247809

condition to bypass the current flowing through Q1 and will remain OFF in partial shading condition. This will further improve the performance of the system under normal condition.

VII. Conclusion

The efficacy of the proposed methodology is verified based on four performance metrics, namely reverse voltage, power loss under partial shading, the sensitivity of the bypass circuit, and the temperature of the hot spotted module. The proposed technique ensures sensitivity enhancement of the bypass circuit towards shade effect; thus, enhancing the reliability of PV system. In experiments performed, the sensitivity enhancement of MBCA has been validated by DSO as well as the diode's forward characteristic plot. Therefore, this solution will potentially increase the sensitivity and hence the reliability of PV modules. Moreover, IGBT is an excellent semiconductor device as it combines the high voltage and current capabilities of a BJT with the voltage control properties of a MOSFET. Its lower on-state resistance and conduction losses, as well as its high voltage and current ratings, makes it ideal for the intended application over other active solid-state devices. As a result, the proposed paradigm would be completely scalable for larger PV systems.

References

[1] D. Giaffreda et al., "Local Shunting in Multicrystalline Silicon Solar Cells: Distributed Electrical Simulations and Experiments," in *IEEE Journal of Photovoltaics*, vol. 4, no. 1, pp. 40-47, Jan. 2014.

[2] A. M. Imtiaz, F. H. Khan and H. Kamath, "All-in-One Photovoltaic Power System: Features and Challenges Involved in Cell-Level Power Conversion in ac Solar Cells," in *IEEE Industry Applications Magazine*, vol. 19, no. 4, pp. 12-23, July-Aug. 2013.

[3] P. Manganiello, M. Balato and M. Vitelli, "A Survey on Mismatching and Aging of PV Modules: The Closed Loop," in *IEEE Transactions on Industrial Electronics*, vol. 62, no. 11, pp. 7276-7286, Nov. 2015.

[4] K. A. Kim, G. -S. Seo, B. -H. Cho and P. T. Krein, "Photovoltaic Hot-Spot Detection for Solar Panel Substrings Using AC Parameter Characterization," in *IEEE Transactions on Power Electronics*, vol. 31, no. 2, pp. 1121-1130, Feb. 2016.

[5] M. C. C. de Oliveira, A. S. A. D. Cardoso, M. M. Viana, and V. D. F. C. Lins, "The causes and effects of degradation of encapsulant ethylene vinyl acetate copolymer (EVA) in crystalline silicon photovoltaic modules: A review," *Renewable Sustain. Energy Rev.*, vol. 81, pp. 2299–2317, 2017.

[6] S. G. Krishna and T. Moger, "Optimal SuDoKu Reconfiguration Technique for Total-Cross-Tied PV Array to Increase Power Output Under Non-Uniform Irradiance," in *IEEE Transactions on Energy Conversion*, vol. 34, no. 4, pp. 1973-1984, Dec. 2019,

[7] R. Venkateswari and N. Rajasekar, ''Power enhancement of PV system via physical array reconfiguration-based Lo Shu technique,'' *Energy Conversion Manage*., vol. 215, pp. 1–22, Jul. 2020.

[8] G Harish Kumar Varma, Barry Venugopal Reddy and Rohit Kumar Jain, "An MMTES algorithm for dynamic photovoltaic array reconfiguration to enhance power output under partial shading conditions", *IET Renewable Power Generation*, vol. 15, no. 4, pp. 809-820, June 2021.

[9] M. Matam and V. R. Barry, "Improved performance of dynamic photovoltaic array under repeating shade conditions", *Energy Convers. Manage*., vol. 168, pp. 639-650, 2018.

[10] M. Gokdaga, M. Akbabab, and O. Gulbudakc, "Switched-capacitor converter for PV modules under partial shading and mismatch conditions," *Sol. Energy*, vol. 170, pp. 723–731, Aug. 2018.

[11] I. Shams, S. Mekhilef and K. S. Tey, "Modular Voltage Equalizer Topology with Reduced Number of Switch Count for Enhancing the Energy Yield During Partial Shading Conditions For PV System," in *2021 IEEE 12th Energy Conversion Congress & Exposition - Asia* (ECCE-Asia), 2021, pp. 1063-1067.

[12] M. Tahmasbi-Fard, M. Tarafdar-Hagh, S. Pourpayam and A. -A. Haghrah, "A Voltage Equalizer Circuit to Reduce Partial Shading Effect in Photovoltaic String," in *IEEE Journal of Photovoltaics*, vol. 8, no. 4, pp. 1102-1109, July 2018.

[13] K. Kato and H. Koizumi, "A study on effect of blocking and bypass diodes on partial shaded PV string with compensating circuit using voltage equalizer," 2015 *IEEE International Symposium on Circuits and Systems* (ISCAS), 2015, pp. 241-244.

[14] S. Pennisi, F. Pulvirenti, and A. La-Scala, "Low-power cool bypass switch for hot spot prevention in photovoltaic panels," *ETRI J.*, vol. 33, no. 6, pp. 880–886, Dec. 2011.

[15] S. Ghosh, V. K. Yadav and V. Mukherjee, "Improvement of partial shading resilience of PV array though modified bypass arrangement", *Renew. Energy*, vol. 143, pp. 1079-1093, Dec. 2019.

[16] F. Reil et al., "Determination of arcing risks in PV modules with derivation of risk minimization measures," in *Conference Record of the IEEE Photovoltaic Specialists Conference*, pp. 3159–3164, 2013.

[17] S. Wendlandt, A. Drobisch, T. Buseth, S. Krauter, and P. Grunow, "Hotspot risk analysis on silicon cell modules," in *Proc. Eur. Photovoltaic Sol.Energy Conf. Exhib.*, pp. 4002–4006, 2010.

[18] C. Olalla, M. N. Hasan, C. Deline, and D. Maksimovic, "Mitigation of hot-spots in photovoltaic systems using distributed power electronics," *Energies*, vol. 11, no. 4, pp. 1–16, Mar. 2018, Art. no. 726.

[19] E. L. Meyer and E. E. van Dyk, "Assessing the reliability and degradation of photovoltaic module performance parameters," in *IEEE Transactions on Reliability*, vol. 53, no. 1, pp. 83-92, March 2004.

[20] H. Yoshioka et al., "Non hot-spot PV module using solar cells with bypass diode function," in *Proc. IEEE Photovolt. Spec. Conf.*, pp. 1271–1274, 1996.

[21] S. Daliento, L. Mele, P. Spirito, R. Carta and L. Merlin, "Experimental Study on Power Consumption in Lifetime Engineered Power Diodes," in *IEEE Transactions on Electron Devices*, vol. 56, no. 11, pp. 2819-2824, Nov. 2009.

[22] G. Acciari, D. Graci and A. La Scala, "Higher PV Module Efficiency by a Novel CBS Bypass," in *IEEE Transactions on Power Electronics*, vol. 26, no. 5, pp. 1333-1336, May 2011.

[23] V. d'Alessandro, P. Guerriero and S. Daliento, "A Simple Bipolar Transistor-Based Bypass Approach for Photovoltaic Modules," in *IEEE Journal of Photovoltaics*, vol. 4, no. 1, pp. 405-413, Jan. 2014.

[24] S. Daliento, F. D. Napoli, P. Guerriero, and V. D'Alessandro, ''A modified bypass circuit for improved hot spot reliability of solar panels subject to partial shading,'' *Sol. Energy*, vol. 134, pp. 211–218, Sep. 2016.

[25] S. Ghosh, V. K. Yadav and V. Mukherjee, "A Novel Hot Spot Mitigation Circuit for Improved Reliability of PV Module," in *IEEE Transactions on Device and Materials Reliability*, vol. 20, no. 1, pp. 191-198, March 2020,

[26] K. A. Kim and P. T. Krein, "Reexamination of Photovoltaic Hot Spotting to Show Inadequacy of the Bypass Diode," in *IEEE Journal of Photovoltaics*, vol. 5, no. 5, pp. 1435-1441, Sept. 2015.

[27] M. Dhimish, V. Holmes, B. Mehrdadi, M. Dales and P. Mather, "Output-Power Enhancement for Hot Spotted Polycrystalline Photovoltaic Solar Cells," in *IEEE Transactions on Device and Materials Reliability*, vol. 18, no. 1, pp. 37-45, March 2018.

[28] P. Guerriero and S. Daliento, "Toward a Hot Spot Free PV Module," in *IEEE Journal of Photovoltaics*, vol. 9, no. 3, pp. 796-802, May 2019.

[29] S. Tang, Y. Xing, L. Chen, X. Song, and F. Yao, ''Review and a novel strategy for mitigating hot spot of PV panels,'' *Sol. Energy*, vol. 214, pp. 51–61, Jan. 2021.

[30] O. Breitenstein, "The Role of Inhomogeneities for Understanding Current–Voltage Characteristics of Solar Cells," in *IEEE Journal of Photovoltaics*, vol. 8, no. 6, pp. 1429-1435, Nov. 2018.

[31] D. Lausch et al., "Correlation of pre-breakdown sites and bulk defects in multicrystalline silicon solar cells," *Phys. Status Solidi RRL*, vol. 3, pp. 70–72, 2009.

[32] J. Bauer *et al.*, "Hot spots in multicrystalline silicon solar cells: Avalanche breakdown due to etch pits," *Phys. Status Solidi RRL*, vol. 3, nos. 2–3, pp. 40–42, 2009.

[33] O. Breitenstein et al., "Understanding junction breakdown in multicrystalline solar cells," *J. Appl. Phys.*, vol. 109, no. 7, Apr. 2011, Art. no. 071101.

© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 01,2023 at 07:23:47 UTC from IEEE Xplore. Restrictions apply.

Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A Novel Cost-efficient Parallel Binary Adder without Crossover in QCA with Energy Dissipation Analysis

Hemanshi Chugh (hemanshichugh_2k20phdec508@dtu.ac.in)
Delhi Technological University

Sonal Singh (sonalsingh@dtu.ac.in)
Delhi Technological University

Research Article

Keywords:

DOI: https://doi.org/

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

A Novel Cost-efficient Parallel Binary Adder without Crossover in QCA with Energy Dissipation Analysis

Hemanshi Chugh[1*] and Sonal Singh[1]

[1]Department of Electronics and Communication, Delhi Technological University, Bawana Road, New Delhi, 110042, Delhi, India.

*Corresponding author(s). E-mail(s): hemanshichugh_2k20phdec508@dtu.ac.in;
Contributing authors: sonalsingh@dtu.ac.in;

Abstract

Quantum dot cellular automata(QCA) is a computational paradigm which uses tiny semiconductor crystals as the basic units of information in computer circuits that can be used to process information similar to transistors. This work introduces a novel half-adder design in single-layer using 90° cells in a single clock zone. Moreover, a full adder circuit using 37 cells in one layer with 0.5 latency is proposed which further forms the basis of a nano-scale parallel binary adder(RCA) design in QCADesigner v2.0.3. The proposed 4-bit design uses 174 cells without the use of any cross-wiring with a latency of only 1 clock. The detailed energy dissipation study of the circuits is also investigated for different tunneling energies using the QCAPro tool. The comparison results of the adder circuits in respect of quantum cost, latency, cell count and total average energy dissipation show significant improvement as compared to the latest QCA-based adder designs.

Keywords: QCA, Full adder, Parallel binary adder, Energy dissipation, QCAPro

1 Introduction

Moore's law has driven the development of modern processors, which now include 16 billion transistors in a chip with a feature size of 5 nanometers, up from 67 transistors in 1965 [1]. Nevertheless, technical challenges, high fabrication costs, and undesirable quantum effects occur if the device density is increased further as electrons can tunnel through gates smaller than 2 nanometers. Also as transistor size gets smaller, they require less power to operate generating less amount of heat, making it more difficult to eliminate. Therefore, alternative technologies are being explored extensively in order to continue the scaling process.

QCA technology has been developed as a potential substitute for building nano-scale electronic devices [2]. Standard semiconductor fabrication methods can be used to create QCA circuits, which are more energy-efficient and densely packed than conventional transistors. The QCA quantum dots are arranged in a regular grid and are used to depict the binary digits 0 and 1. A state of a QCA based circuit is ascertained by the alignment of the quantum dots, which can be manipulated using electrical or magnetic fields. QCA has several advantages over traditional silicon-based transistor technology. For example, QCA circuits can be much smaller and more energy efficient than their transistor counterparts. They have the potential to revolutionize computing and other fields by enabling the development of very small, more stable, and high-performance devices. Adder circuits are an integral part of digital circuits as they can be used to build devices that can perform complex mathematical calculations, process and analyze data, and perform a wide range of computational tasks. A new implementation of co-planar QCA-based HA, FA, and n-bit RCA designs is suggested in this work in order to minimize the cell count, and reduce the cost and complexity with decreased energy and power dissipation.

The article flow is arranged as follows. Some basic fundamentals of QCA technology have been described in Section 2. Further, a brief literature survey of the adder designs in QCA is reviewed in Section 3. Section 4 outlines the improved QCA-based adder layouts in detail whereas Section 5 analyses the simulation results and power consumption of proposed design in detail. An elaborate comparison and evaluation of the performance parameters of the QCA designs with the energy dissipation study is offered in Section 6. Section 7 concludes the presented work.

2 QCA Basic Concepts

QCA is a four-dot square device that consists of 4 quantum dots electrically connected to one another as possible vacancies for the two free electrons. Due to the electron-electron repulsion, there is two possible polarization of the cell depending on the two diagonal positions of the electrons as shown in Fig 1. The logic values 0 and 1 in binary is represented by the polarization values of -1 and +1 respectively in QCA cells [3].

There are three fundamental components that form the building blocks of any

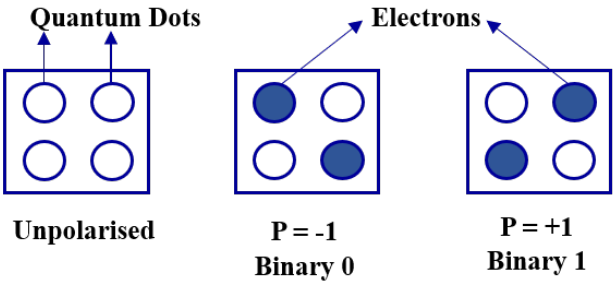


Fig. 1 Polarisation States of QCA cell

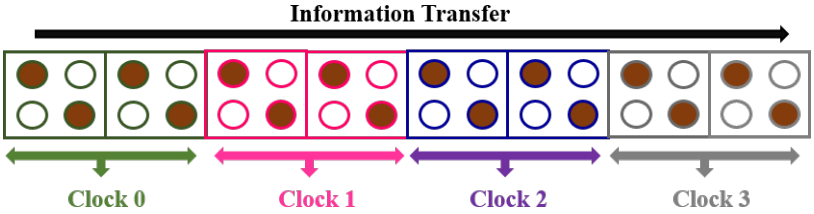


Fig. 2 Information Transfer using QCA Wire

circuit in QCA namely the Inverter, Wire, and Majority Voter(MV) Gate. Information flow is possible in QCA circuits by placing quantum cells in series. Due to the coulumbic interactions between the cells, the neighboring cell tends to take the same polarisation [4]. Hence, the polarisation is propagated through the 'wire' of quantum cells. It is worth noting that Fig. 2 also illustrates the four clock zones: clock 0,1,2 and 3 in QCA that are responsible for perfect data transfer in a circuit. Each clock zone is 90° out of phase with the prior clock. The four phases in QCA clocking [5, 6] (as illustrated in Fig. 3), namely switch, hold, release, and relax permit the electron movement inside the cell so as to synchronize the operation of the circuit. At a time, among these clock phases, the valid data will only be in one clock which would be in the hold state. The potential barrier between quantum dots is lower in the first phase, also known as the switch phase, and the cell is in an unpolarised state. The barrier rises at this clock, and the cell starts to polarise in response to its own input. The barrier is held high during this phase, often known as the Hold phase. The third phase, also known as the release state, is characterized by the lowering of the barrier and allowing the cells to enter an internal polarisation stage. Lastly in the fourth phase, known as the relax state, the potential barrier is low and the QCA cell retains its unpolarised condition [7].

One of the main designs employed in QCA is the NOT gate. The cells positioned diagonally to one another exhibit coulombic repulsion, which causes reverse polarisation. Utilizing this quality, an inverter is created (INV). The

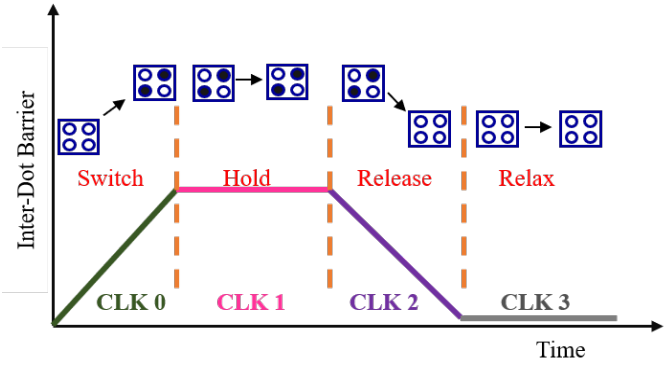


Fig. 3 Clock zones in QCA(adapted from [7])

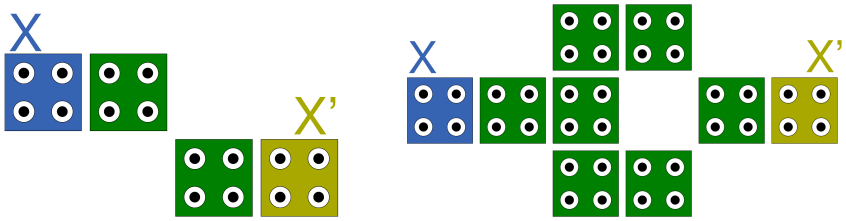


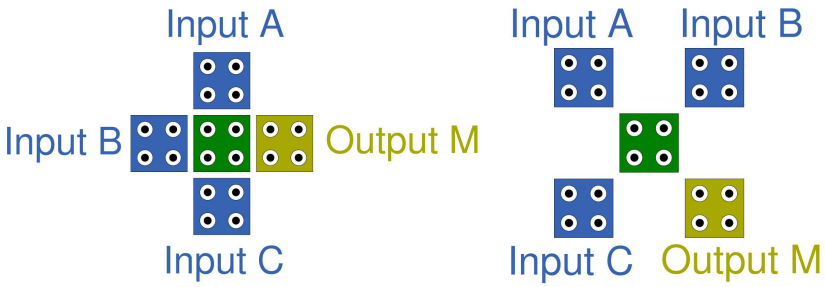
Fig. 4 NOT gates in QCA

inverter gate with two alternative QCA layouts has been shown in Fig. 4. The blue cells signify the input cell (X) and the yellow cells denote the output cell(X').

Furthermore, the majority voter gate(MV) is another dominating gate in QCA circuits [8]. MV gate is of utmost importance for QCA layout implementations. The gate typically has 3 inputs, 1 device cell, and 1 output. The output of the MV gate depends on the maximum polarisation states of the inputs. Eq. 1 provides the Boolean equation for the 3 input MV. Fig. 5 presents two forms of 3-input majority voters.

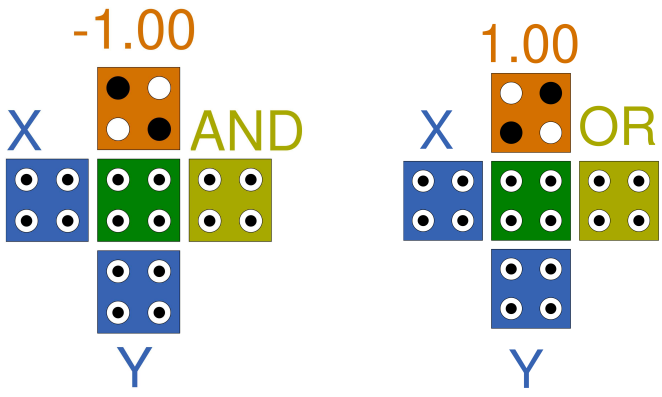
$$MG = AB + BC + AC \tag{1}$$

The two-bit logical AND and OR operations can be designed by simply adjusting the polarisation of any one input of MV gate to -1 (binary '0') or 1 (binary '1') as illustrated in Fig. 6 [7]. The yellow cell indicates the fixed polarisation cell.



(a) (b)

Fig. 5 Majority Voter Gate in QCA



(a) (b)

Fig. 6 Logical AND and OR operation in QCA

3 Related Works

This segment plans to survey past research related to QCA based adder designs and determine their principal highlights and shortcomings.

Numerous QCA half-adder designs have been put forth in the literature. The initial design of HA was a multi-layer layout with 77 cells in 0.08 μm^2 area. The earlier designs by Lakshmi et al [9], Santra et al [10] and Ajitha et al [11] were based on linear logic modeling with a majority voter-based approach, thereby, resulting in high complexity and area, However, using 37 cells in a single-layered architecture led to profound cell optimization with an area of only 0.05 μm^2. However, employing an improved XOR gate design, an efficient HA circuit was implemented with only 21 cells in 0.02 μm^2 but with 0.5 clock delay as it employs 2 clock cycles in the critical path.

To construct a FA circuit, two effective approaches are used in QCA technology:

- using XOR gates and AND-OR logic function realization
- employing majority voter gates to realise adder function

However, MV-based FA circuits use more cells than the XOR-based FA realisation.

The first QCA-based coplanar FA design was implemented by Tougaw et al. [3] using 5 3-input MG, 3 INV and 9 crossovers. Further, Sasamal et al. [12] presented a more efficient design of 55 QCA cells(including 5-input MG) dissipating 98.57 meV energy. However, when optimised XOR gate designs are used to implement FA designs, appreciable efficiency of design metrics is seen such as cell count has drastically reduced from 292 to 40, 43, 38,39, 37 and so on[13–15]. Trailokya and Ashutosh's design [16] of a one-bit FA using 3-input XOR gate and MV gate incorporates 40 cells with increased area and delay of 5 clock phases. Various multilayer architectures have also been proposed [17–19] with further reduction in the cells but an increase in fabrication cost. One example of the ultra-dense structure of FA was by Sarmadi et al. [20]. They have proposed a revised six active layer QCA-based one-bit FA. The one-bit FA that is being presented can have an impact on all of the earlier ideas in terms of cell numbers and space usage. The recommended adder size is kept equal to a 5- input MV gate due to the inventive use of six active layers thereby, considerably minimizing the area occupied. The downsides of the suggested technique are the increased expenses of fabrication due to complexity in the upper layer.

Various types of adder circuits can further be implemented from the FA circuits such as RCA, CSA(Carry Save Adder), CSeA(Carry Select Adder), and so on. Mohsen Vahabi et al [21] implemented two designs of RCA using co-planar crossovers. The said circuit is stable and more robust due to the decrease in a number of INV cells, rotational cells, and crossovers. The first implementation of the adder uses 233 cells in an area of 0.28 μm^2 with a 1.25 clock zone delay. The second implementation of the adder uses 209 cells with an area of 0.25 μm^2 with the same latency.

Saeid Seyedi et al [22] have designed a nano-scale RCA in order to minimize the cell number and delay thereby improving the complexity of the circuit. The proposed co-planar design uses 89 cells in 0.12 μm^2 but giving a delay of 3.75 clock cycles which results in an increase in the quantum cost.

In order to implement one-bit FA with reduced delay and complexity, Balali et al. [23] presented a 3-input EXOR gate layout based on cell interaction. As a result, a four-bit single-layer RCA is recommended by using 269 cells in 0.37 μm^2 with an increased delay of 3.5 clock cycles.

Hashemi and Navi [24] have suggested a reliable QCA-based FA based on a five-input MV gate. The RCA design simulation was carried out using a reliable crossover method. It utilizes both engines of the QCADesigner simulation program, the results show 442 cells in a space of 1 μm^2 with a latency of 2 clock cycles.

Additionally, utilising 494 cells in a 0.68 μm^2 region, Angizi et al. [25] demonstrated a single layer implementation of a 4-bit RCA. The circuit shows an

increased latency of 4.25 clock cycles.
To decrease the number of QCA cells and the amount of space used, Abedi, et al. [26] have introduced a FA design based on coplanar crossover technique in QCA. Two crossing wires were used in the crossover in nonadjacent clock zones. A recent coplanar 4-bit RCA design is put forward by Jeyalakshmi and Santhi [27] utilizing only 70 QCA cells but with an increased delay of 5 clock cycles, thereby increasing the quantum cost of the design. Finally, a multi-layer four-bit RCA was designed by Mohammadi et al. [19] utilizing a FA with 38 cells in 0.02 μm^2 and a delay of 3 clock cycles. Furthermore, RCA with 237 cells in a 0.24 μm^2 area is designed using this adder. The 4-bit RCA circuit has an increased delay of 6 clock phases. Additionally, higher-order RCA circuits have been developed.

4 Proposed Adder Circuits

Among the fundamental computational circuits in digital logic, the most commonly used digital arithmetic circuit, are adders. Due to its widespread application in all arithmetic and computational circuits, the HA and FA circuits are the most significant and fundamental adder circuits. In this section, we first design the layout of a novel HA circuit followed by the realisation of the FA circuit by strategic placement of cells. Finally, we implement the 4-bit parallel adder circuit employing the same suggested FA. Compared to the best earlier models, these designs are intended to have decreased cost function and are implemented in a single layer without the usage of rotational cells(45° cells) or crossovers.

4.1 Half Adder

Two binary variables A and B can be added using the combinational circuit; half adder producing the outputs as sum and carry defined in the equations 2 and 3 as :

$$Sum = A \oplus B \tag{2}$$

$$Carry = A \,.\, B \tag{3}$$

Here, a new design of a single-layer HA circuit is presented in QCA technology. The key features to be noted about the design are that it uses one INV gate and one MV gate in a single clock phase without the use of any crossovers. The suggested layout consists of 23 cells. There are four labeled cells wherein P and Q are the inputs to the HA and X and CY represent the outputs: sum and carry respectively. The presented design shows a latency of 0.25 clock zones. The representation is simple with reduced cell count and requires no cross-wiring thereby, decreasing the fabrication cost. The layout of the circuit is shown in Fig. 7 and the operation can be represented by:

$$X = XOR(P\,,\,Q) \tag{4}$$

$$CY = MV(P\,,\,Q) \tag{5}$$

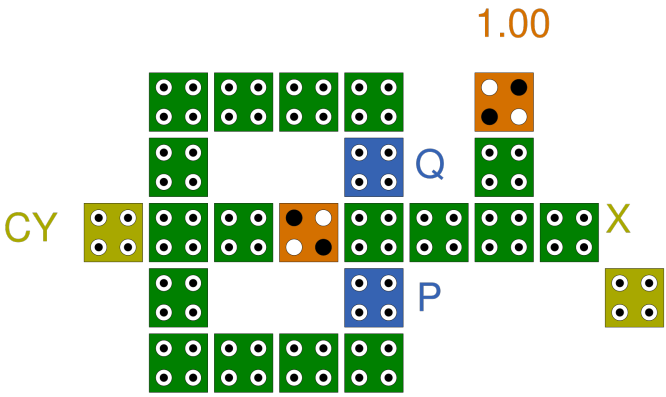


Fig. 7 Proposed Half Adder cell design in QCA

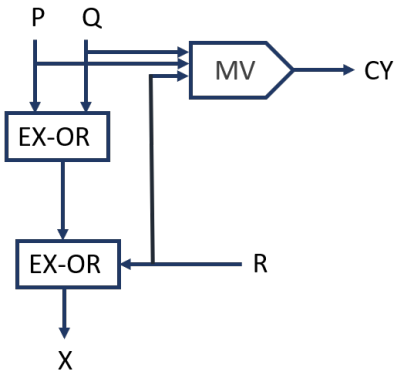


Fig. 8 Full Adder logic diagram

4.2 Full Adder

Three input bits are added arithmetically by a combinational circuit known as a 1-bit FA having two significant bits: A, B, and a preceding carry: C_{in}. There are two outputs namely, sum and carry defined as per the equations:

$$Sum = A \oplus B \oplus C_{in} \tag{6}$$

$$Carry - out = AB + BC_{in} + AC_{in} \tag{7}$$

In this, a novel QCA-based FA is introduced with P, Q, and R as inputs and outputs X and CY as sum and carry-out respectively. The QCA FA design approach is as illustrated in Fig. 8. The schematic design as shown in Fig. 9 illustrates the best-organised structure of its cells. The layout of FA is based

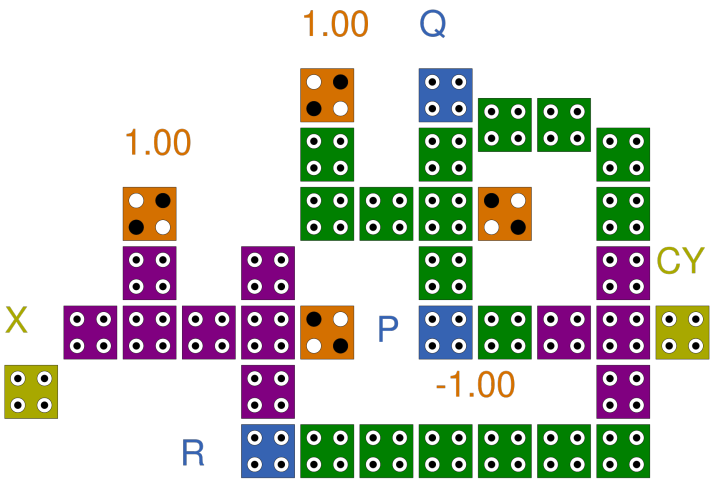


Fig. 9 Proposed Full Adder cell design in QCA

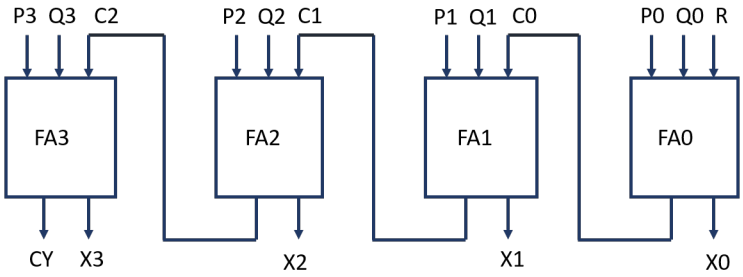


Fig. 10 4-bit parallel binary adder

on cell interaction and uses two XOR gates to generate X and one 3-input MV gate to generate the CY. As shown in Fig. 9, the proposed design does not use any type of coplanar cross-wiring and uses only 90° cells in a single layer. The efficacy of the design can be validated as it uses only 37 cells for the single-layer implementation. The latency of the said design is 0.5 clock cycles.

4.3 n-bit Parallel Binary Adder

An n-bit parallel binary adder is a digital circuit used to perform the addition of n-bit numbers. For n-bit addition, n FA are cascaded in series. In this work, a 4-bit parallel binary adder is designed using 4 FAs(as illustrated in Fig. 10) introduced above. Consider two 4-bit binary numbers P and Q such that P = P3 P2 P1 P0 and Q = Q3 Q2 Q1 Q0. Also, R is the input carry to the first FA. The 5-bit output X is defined as CY X3 X2 X1 X0. As it can be seen in Fig. 11 that carry ripples from one stage to another, this adder is otherwise named ripple carry adder (RCA). The layout is implemented in one layer to

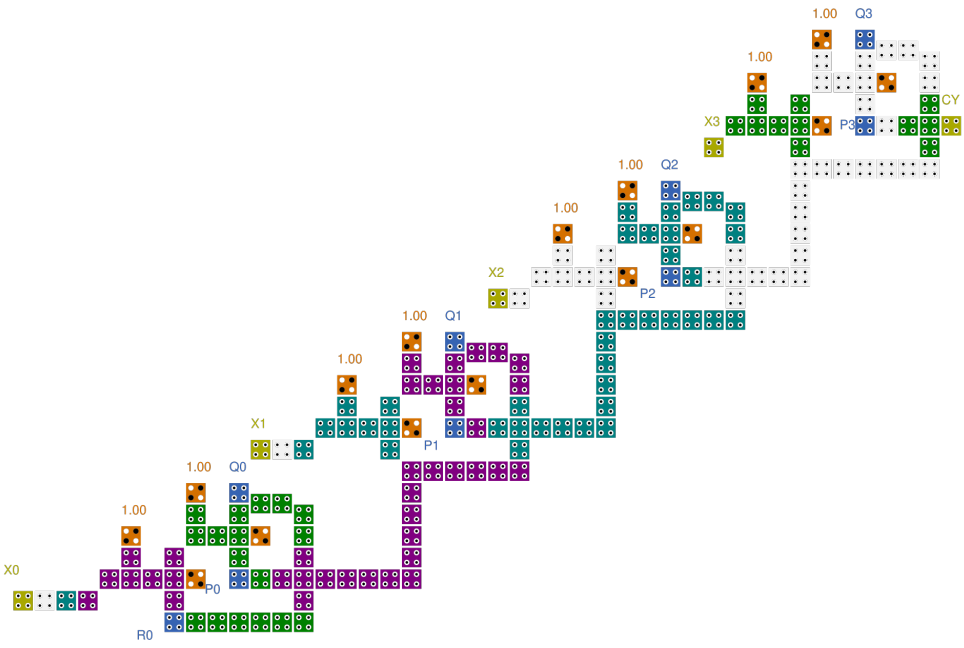


Fig. 11 QCA layout of 4-bit parallel binary adder

make it cost efficient as multi-layer designs increase the fabrication cost. The developed 4-bit parallel adder incorporates 174 cells in 0.53 μm^2 area with a minimum delay of 1 clock. Further, the QCA layout implementation for an 8-bit parallel adder has also been shown in Fig. 12. The adder has three inputs P[7:0], Q[7:0], R0; and two outputs X[7:0] and COUT. The layout uses 371 cells only with a minimum latency of 2 clock cycles.

5 Experimental Results and Analysis

5.1 Simulation Results

This section outlines the simulation results for the proposed adder circuits. The above structures have been tested for their functionality by performing simulation on QCADesigner Tool v2.0.3. The GaAs cells are considered. Table 1 describes the initialization of the basic parameters used for simulation. The accuracy of the HA circuit can be verified from Fig. 13. The truth table of HA is displayed using a bistable approximation engine.

Further, Fig. 14 illustrates the input-output waveform of the FA. The entire truth table has been verified for its correctness using bi stable and coherence vector simulation engines. However, bi stable simulation engine outputs have been displayed.

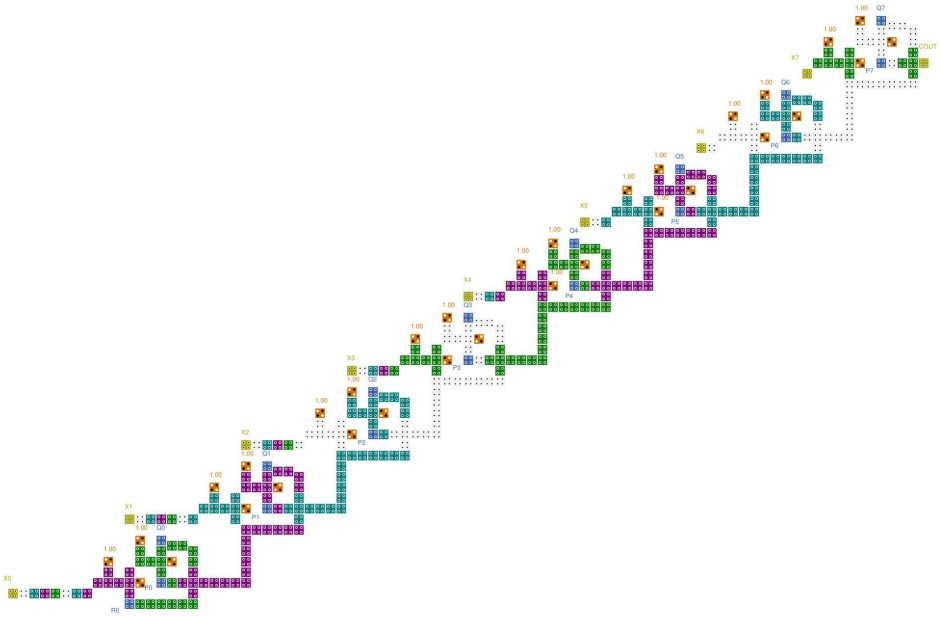


Fig. 12 QCA layout of 8-bit parallel binary adder

Table 1 Simulation parameters

Parameter	Value
Cell size	18*18 nm
Radius of effect	65 nm
Cell seperation	2 nm
Dot diameter	5 nm
Simulation engine	Bistable approximation
Relative permittivity	12.9000000
Number of samples	12800
Convergence tolerance	0.001000
Clock amplitude factor	2.000000
Clock High	9.8000000e-022
Clock shift	0.000000e+000
Clock Low	3.8000000e-023
Max Iterations per sample	100
Layer separation	11.500000

Because of the numerous entries in the test of the RCA, we evaluated some vector sets using the vector table as shown in Fig. 15. The results can be verified from the Fig. 16 as the 4-bit parallel adder runs proficiently with a latency of 1 clock cycle.

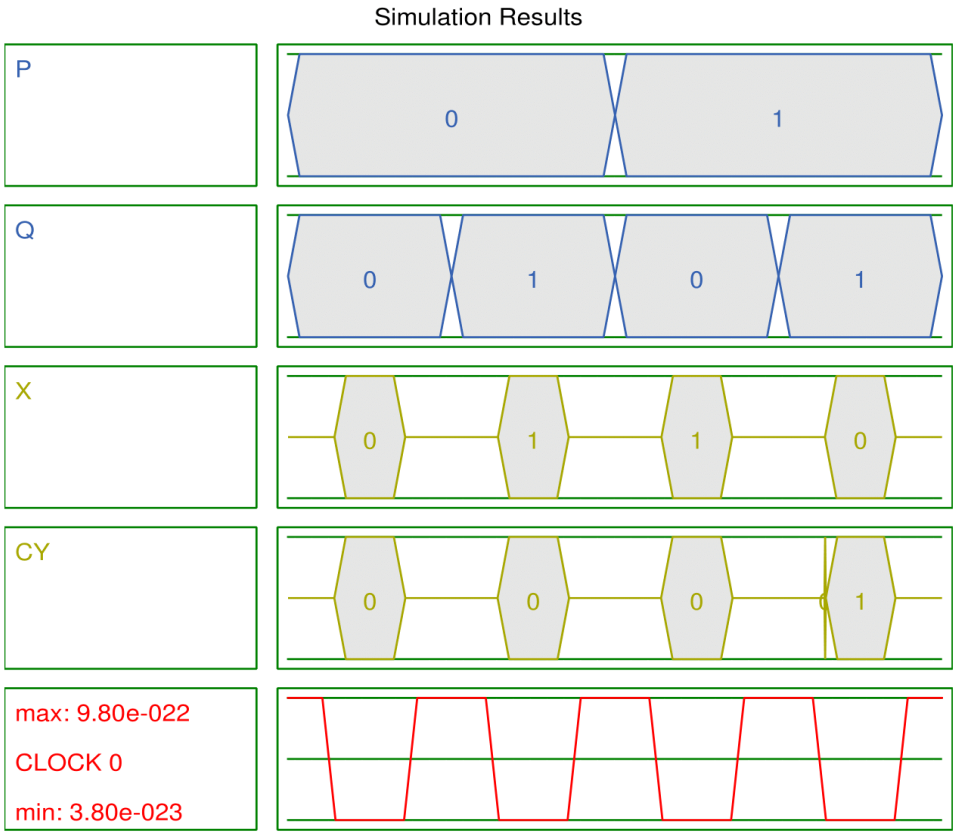


Fig. 13 Verified truth table of proposed HA

5.2 Power Dissipation Analysis

QCA Pro tool is a probabilistic modeling tool to estimate to analyze polarization error and non-adiabatic switching power loss as it uses fast approximation. Figure 17 and 18 display the the power dissipation maps for the proposed HA and FA designs respectively at different tunnelling energies of $0.5E_k$, $1.0E_k$, $1.5E_k$.

6 Detailed Performance Evaluation

This section compares the various performance attributes related to the proposed adder designs with their previous counterparts. The cost value is computed using the following equation:

$$Cost = Area * Latency \tag{8}$$

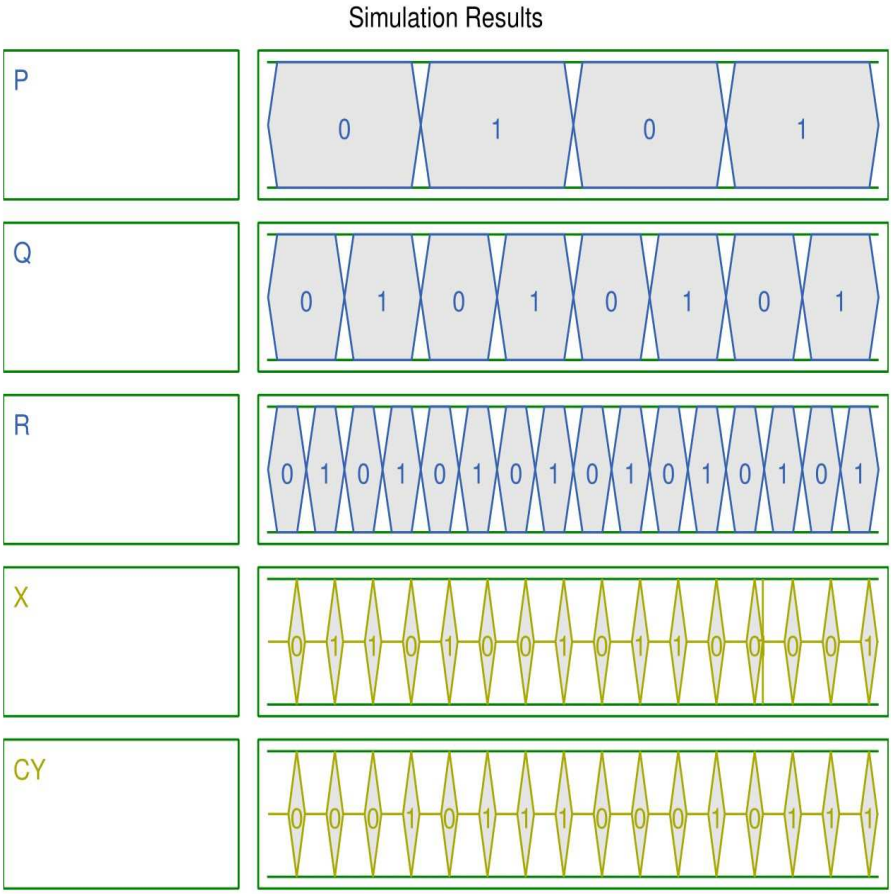


Fig. 14 Verified truth table of proposed FA

Where the area is displayed in units of μm^2 and the Latency represents the number of clock cycles required.

Energy dissipation is one of the most important parameters for the evaluation of their performance and efficacy of the designed circuit. The analysis has been carried out using the QCAPro tool [49]. In these calculations, it is assumed that all QCA cells have a constant power consumption. At three distinct tunnelling energies (0.5 E_k, 1 E_k, and 1.5 E_k) at a temperature of 2K, the designed HA and FA circuits' detailed energy dissipation data are shown in Table 5 and 6 respectively.

7 Conclusion and Future Directions

In this paper, a novel HA layout was designed using only 23 QCA cells followed by the FA layout design with 37 QCA cells in a single layer without any

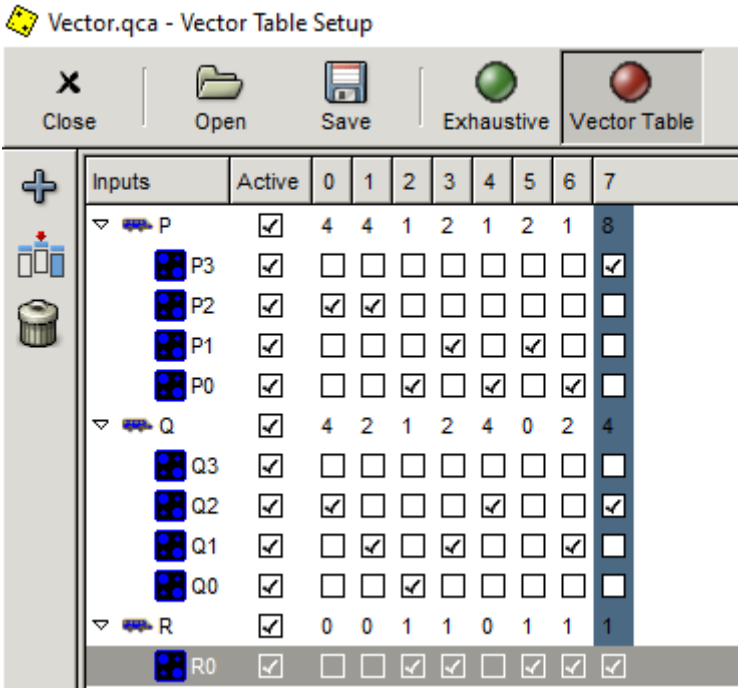


Fig. 15 Vector Table Setup for RCA

Table 2 Comparison of proposed HA with existing HA designs

Ref	Cell Count	Area(μm^2)	Crossover	Latency	Cost
[9]	77	0.08	Multilayer	1	0.08
[10]	65	-	Coplanar	1.25	-
[14]	62	0.08	NR	2	0.0375
[28]	61	-	NR	0.75	-
[11]	34	0.05	NR	0.75	0.05
[29]	31	0.05	NR	1	0.0135
[30]	26	0.027	NR	0.5	0.0135
[31]	21	0.02	Multilayer	0.5	0.01
Proposed	**23**	**0.024**	**NR**	**0.25**	**0.006**

NR: Not Required

need for crossovers in the design. In order to evaluate the proficiency of FA in larger adder circuits, a 4-bit and 8-bit binary parallel adder was designed. The structures were simulated and verified using QCADesigner v2.0.3. The relative comparison of the proposed adder designs with the state of the art adder circuits demonstrates that the developed structures are more cost-effective than the previous structures. QCAPro tool was used to analyze power consumption and get a detailed insight of energy dissipation for different tunneling energies for the proposed co-planar QCA structures. In the past, various RCA designs have been implemented using multi-layers or coplanar crossovers with adjacent

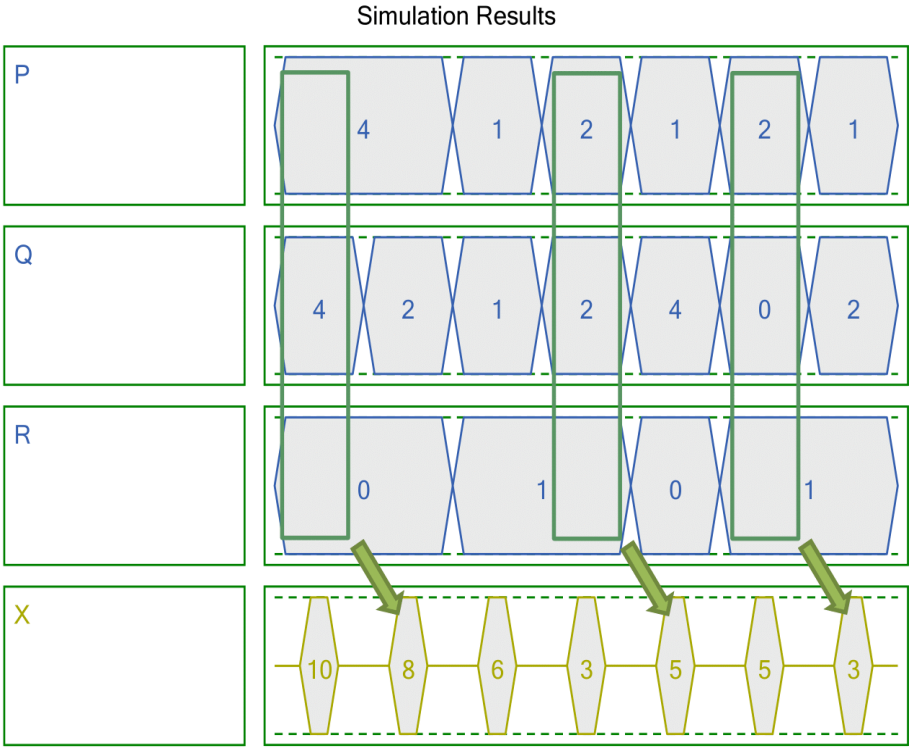


Fig. 16 Simulation outputs of proposed 4-bit parallel adder

clock zones leading to increased latency and cost. Due to the fact that it is a single-layer design with no crossovers, our design has shown to be effective. Hence, the implemented circuit is significantly superior in terms of parameter metrics as it is more stable and robust than previous designs. In order to enhance and optimize composite and complicated circuits, these circuits can be used effectively.

8 Declarations

Ethical Approval Not Applicable

Competing Interests The authors declare that they have no known competing financial and personal interests.

Author's Contribution Hemanshi Chugh and Sonal Singh conceptualised the idea of the paper. Hemanshi Chugh developed the theory and performed the layout with its computations. Sonal Singh verified the analytical methods. All authors discussed the results and contributed to the final manuscript.

Funding This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Availability of data and materials The paper includes all data.

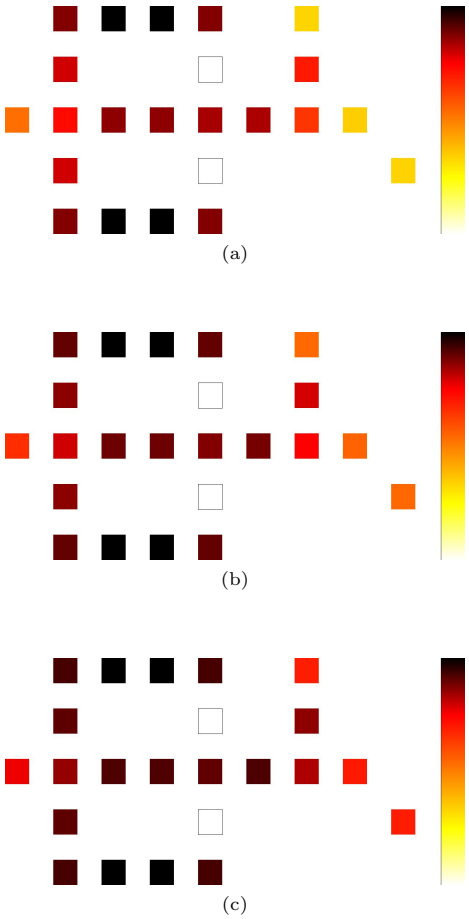
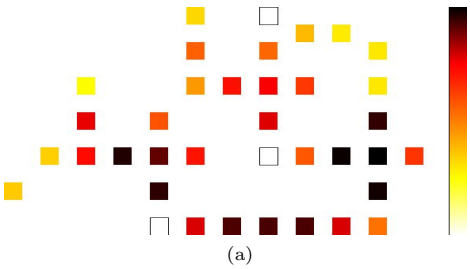
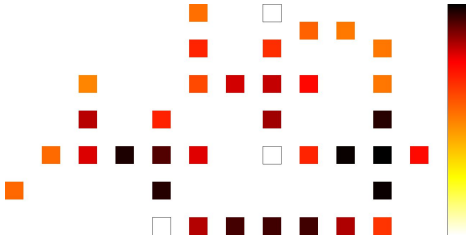


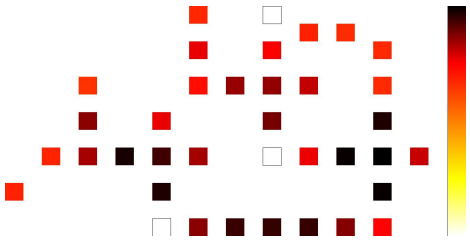
Fig. 17 Thermal Hotspots for HA at $T = 2K$ temperature with tunneling energy (a)$E_k = 0.5$, (b) $E_k = 1.0$ and (c) $E_k = 1.5$



(a)



(b)



(c)

Fig. 18 Thermal Hotspots for FA at $T = 2K$ temperature with tunneling energy(a)$E_k = 0.5$, (b) $E_k = 1.0$ and (c) $E_k = 1.5$

Table 3 Relative evaluation of the proposed FA circuit with the state of the art

Ref.	Cell Count	Area(μm^2)	Crossover	Latency	Cost
[32]	228	0.28	NR	1.75	0.49
[10]	150	0.28	Co-planar	2.25	0.63
[9]	122	0.114	Co-planar	1	0.114
[33]	120	0.15	Co-planar	2.5	0.375
[34]	114	0.23	Co-planar	1.25	0.2875
[25]	95	0.09	Co-planar	1.25	0.1125
[17]	78	0.09	Co-planar	0.75	0.0675
[35]	71	0.06	Co-planar	1.25	0.075
[36]	69	0.07	Co-planar	1	0.07
[15]	61	0.076	Co-planar	0.5	0.038
[2]	60	0.057	NR	1	0.057
[37]	56	0.047	Co-planar	1	0.047
[38]	52	0.038	NR	0.75	0.0285
[30]	47	0.04	Co-planar	0.75	0.03
[38]	46	0.04	NR	0.75	0.03
[39]	45	0.043	NR	0.75	0.0322
[40]	44	0.043	NR	1.5	0.0645
[39]	43	0.037	NR	0.75	0.02775
[41]	41	0.04	Co-planar	0.5	0.02
[42]	41	0.04	Coplanar	0.5	0.02
[43]	37	0.0245	Co-planar	1	0.0245
Proposed	**37**	**0.039**	**NR**	**0.5**	**0.0195**

Table 4 Comparative analysis of 4-bit parallel binary adder

Ref.	Cell Count	Area(μm^2)	Crossover	Latency	Cost
[44]	1246	2.5	Co-planar	3.25	8.125
[13]	558	0.85	-	5	-
[25]	494	0.68	Co-planar	4.25	2.89
[35]	442	1	Co-planar	2	2
[45]	314	0.32	Multi-layer	1.5	0.48
[24]	308	0.29	Multi-layer	2	0.58
[46]	295	0.3	Co-planar	1.5	0.45
[23]	269	0.37	Co-planar	3.5	1.295
[26]	262	0.208	Co-planar	1.75	0.364
[19]	237	0.24	Multi-layer	1.5	0.36
[41]	209	0.3	Co-planar	1.25	0.375
[47]	187	0.2	Co-planar	4	0.8
[48]	184	0.1	Multi-layer	2	0.121
[27]	70	0.18	Co-planar	5	0.90
Proposed	**174**	**0.53**	**NR**	**1**	**0.53**

Table 5 Analysis of energy dissipation of the suggested HA design versus previous designs at three distinct tunnelling energy levels

	Leakage energy dissipation (meV)			Switching energy dissipation (meV)			Total energy dissipation (meV)		
HA Circuit	0.5 E_k	1 E_k	1.5 E_k	0.5 E_k	1 E_k	1.5 E_k	0.5 E_k	1 E_k	1.5 E_k
[50]	17.86	55.56	100.56	97.36	84.65	72.06	115.22	140.21	172.62
[10]	35.47	98.54	168.28	94.90	79.07	65.62	130.37	177.61	233.90
[33]	13.81	42.29	75.56	67.73	58.38	49.43	81.54	100.67	124.99
[51]	12.21	37.07	65.81	54.02	46.55	39.51	66.23	83.62	105.32
[30]	8.16	24.08	42.30	29.26	25.00	21.06	37.42	49.08	63.36
Proposed HA	**6.35**	**19.5**	**35.10**	**31.02**	**29.2**	**25.1**	**37.37**	**48.86**	**60.2**

Table 6 Analysis of energy dissipation of the suggested FA design versus previous designs at three distinct tunnelling energy levels

	Leakage energy dissipation (meV)			Switching energy dissipation (meV)			Total energy dissipation (meV)		
FA Circuit	0.5 E_k	1 E_k	1.5 E_k	0.5 E_k	1 E_k	1.5 E_k	0.5 E_k	1 E_k	1.5 E_k
[10]	81.73	229.58	394.42	154.39	129.06	107.19	236.12	358.64	501.61
[25]	26.52	81.39	147.1	159.02	140.40	121.8	185.5	221.7	269.0
[50]	31.46	96.51	171.64	158.97	135.54	114.32	190.43	232.05	285.96
[26]	16.43	51.09	92.6	105.94	92.68	79.61	122.37	143.77	172.21
[30]	13.31	41.51	74.74	76.80	66.00	55.80	90.11	107.51	130.54
[52]	14.05	39.32	68.02	50.61	44.67	38.95	64.66	83.99	106.97
[53]	21.38	62.38	106.89	80.92	67.68	57.9	102.18	130.05	164.79
[15]	29.19	84.83	148.5	127.94	108.82	91.46	157.13	193.65	239.97
[54]	16.93	51.11	91.37	104.91	92.41	80.36	121.84	143.52	171.73
[40]	17.3	45.89	76.92	41.11	34.33	28.86	58.41	80.22	105.78
[2]	15.37	44.87	76.92	61.87	56.03	51.61	77.24	100.9	128.53
[37]	14.95	47.9	87.32	92.15	80.63	69.3	107.1	128.53	156.62
[55]	11.22	32.95	58.43	64.03	56.66	49.46	75.25	89.61	107.89
[31]	15.48	43.11	75.92	42.91	36.72	29.65	58.39	79.83	105.57
[16]	13.96	37.77	63.88	29.55	24.65	20.4	43.51	62.42	84.28
[39]a	13.91	40.87	72.13	55.13	48.1	41.32	69.04	88.97	113.45
[39]b	14.15	43.37	77.91	56.98	46.54	42.47	71.13	89.91	120.38
[21]	11.12	32.92	58.37	49.1	42.82	36.62	60.22	75.74	94.99
Proposed FA	**12.97**	**35.75**	**61.4**	**44.24**	**37.75**	**31.88**	**57.2**	**73.5**	**93.28**

References

[1] Zhang, G.-Q., Roosmalen, A.v.: More than moore. Springer (2009)

[2] Wang, L., Xie, G.: A novel xor/xnor structure for modular design of qca circuits. IEEE Transactions on Circuits and Systems II: Express Briefs **67**(12), 3327–3331 (2020)

[3] Tougaw, P.D., Lent, C.S.: Logical devices implemented using quantum cellular automata. Journal of Applied physics **75**(3), 1818–1825 (1994)

[4] Kim, H.-I., Jeon, J.-C.: Quantum lfsr structure for random number generation using qca multilayered shift register for cryptographic purposes. Sensors **22**(9), 3541 (2022)

[5] Edrisi Arani, I., Rezai, A.: Novel circuit design of serial–parallel multiplier in quantum-dot cellular automata technology. Journal of Computational Electronics **17**(4), 1771–1779 (2018)

[6] Vankamamidi, V., Ottavi, M., Lombardi, F.: Two-dimensional schemes for clocking/timing of qca circuits. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems **27**(1), 34–44 (2007)

[7] Riyaz, S., Kakkar, V.: Quantum-dot cellular automata: An efficient and adroit alternate to cmos technology. In: 2020 International Conference on Emerging Smart Computing and Informatics (ESCI), pp. 282–288 (2020). IEEE

[8] Lent, C.S., Tougaw, P.D., Porod, W.: Quantum cellular automata: the physics of computing with arrays of quantum dot molecules. In: Proceedings Workshop on Physics and Computation. PhysComp'94, pp. 5–13 (1994). IEEE

[9] Lakshmi, S.K., Athisha, G.: Design and analysis of adders using nanotechnology based quantum dot cellular automata. Journal of Computer Science **7**(7), 1072 (2011)

[10] Santra, S., Roy, U.: Design and implementation of quantum cellular automata based novel adder circuits. International Journal of Nuclear and Quantum Engineering **8**(1), 178–183 (2014)

[11] Ajitha, D., Ramanaiah, K.V., Sumalatha, V.: An efficient design of xor gate and its applications using qca. i-Manager's Journal on Electronics Engineering **5**(3), 22 (2015)

[12] Sasamal, T.N., Singh, A.K., Mohan, A.: An optimal design of full adder based on 5-input majority gate in coplanar quantum-dot cellular

automata. Optik **127**(20), 8576–8591 (2016)

[13] Hänninen, I., Takala, J.: Binary adders on quantum-dot cellular automata. Journal of Signal Processing Systems **58**, 87–103 (2010)

[14] Ahmad, P.Z., Ahmad, F., Khan, H.A.: A new f-shaped xor gate and its implementations as novel adder circuits based quantum-dot cellular automata (qca). IOSR Journal of Computer Engineering (IOSR-JCE) **16**(3), 110–117 (2014)

[15] Erniyazov, S., Jeon, J.-C.: Carry save adder and carry look ahead adder using inverter chain based coplanar qca full adder for low energy dissipation. Microelectronic Engineering **211**, 37–43 (2019)

[16] Sasamal, T.N., Mohan, A., Singh, A.K.: Optimal realization of full adder in qca using 5-input majority gate. In: 2020 International Conference on Industry 4.0 Technology (I4Tech), pp. 203–206 (2020). IEEE

[17] Cho, H., Swartzlander, E.E.: Adder and multiplier design in quantum-dot cellular automata. IEEE Transactions on Computers **58**(6), 721–727 (2009)

[18] Pudi, V., Sridharan, K.: Low complexity design of ripple carry and brent–kung adders in qca. IEEE Transactions on nanotechnology **11**(1), 105–119 (2011)

[19] Mohammadi, M., Mohammadi, M., Gorgin, S.: An efficient design of full adder in quantum-dot cellular automata (qca) technology. Microelectronics journal **50**, 35–43 (2016)

[20] Sarmadi, S., Sayedsalehi, S., Fartash, M., Angizi, S.: A structured ultra-dense qca one-bit full-adder cell. Quantum Matter **5**(1), 118–123 (2016)

[21] Vahabi, M., Bahar, A.N., Otsuki, A., Wahid, K.A.: Ultra-low-cost design of ripple carry adder to design nanoelectronics in qca nanotechnology. Electronics **11**(15), 2320 (2022)

[22] Seyedi, S., Pourghebleh, B., Jafari Navimipour, N.: A new coplanar design of a 4-bit ripple carry adder based on quantum-dot cellular automata technology. IET Circuits, Devices & Systems **16**(1), 64–70 (2022)

[23] Balali, M., Rezai, A., Balali, H., Rabiei, F., Emadi, S.: Towards coplanar quantum-dot cellular automata adders based on efficient three-input xor gate. Results in physics **7**, 1389–1395 (2017)

[24] Hashemi, S., Tehrani, M., Navi, K.: An efficient quantum-dot cellular automata full-adder. Scientific Research and Essays **7**(2), 177–189 (2012)

[25] Angizi, S., Alkaldy, E., Bagherzadeh, N., Navi, K.: Novel robust single layer wire crossing approach for exclusive or sum of products logic design with quantum-dot cellular automata. Journal of Low Power Electronics **10**(2), 259–271 (2014)

[26] Abedi, D., Jaberipur, G., Sangsefidi, M.: Coplanar full adder in quantum-dot cellular automata via clock-zone-based crossover. IEEE transactions on nanotechnology **14**(3), 497–504 (2015)

[27] Maharaj, J., Muthurathinam, S.: Effective rca design using quantum dot cellular automata. Microprocessors and Microsystems **73**, 102964 (2020)

[28] Jagarlamudi, H.S., Saha, M., Jagarlamudi, P.K.: Quantum dot cellular automata based effective design of combinational and sequential logical structures. World Academy of Science, Engineering and Technology **60**, 671–675 (2011)

[29] Poorhosseini, M., Hejazi, A.R.: A fault-tolerant and efficient xor structure for modular design of complex qca circuits. journal of Circuits, Systems and Computers **27**(07), 1850115 (2018)

[30] Bahar, A.N., Ahmad, F., Wani, S., Al-Nisa, S., Bhat, G.M.: New modified-majority voter-based efficient qca digital logic design. International Journal of Electronics **106**(3), 333–348 (2019)

[31] Majeed, A.H., Zainal, M.S.B., Alkaldy, E., Nor, D.M.: Full adder circuit design with novel lower complexity xor gate in qca technology. Transactions on Electrical and Electronic Materials **21**, 198–207 (2020)

[32] Taherkhani, E., Moaiyeri, M.H., Angizi, S.: Design of an ultra-efficient reversible full adder-subtractor in quantum-dot cellular automata. Optik **142**, 557–563 (2017)

[33] Ahmad, F., Bhat, G.M., Ahmad, P.Z.: Novel adder circuits based on quantum-dot cellular automata (qca). Circuits and Systems **2014** (2014)

[34] Singhal, R., Perkowski, M.: Comparative analysis of full adder custom design circuit using two regular structures in quantum-dot cellular automata (qca). In: 2019 IEEE 49th International Symposium on Multiple-Valued Logic (ISMVL), pp. 194–199 (2019). IEEE

[35] Hashemi, S., Navi, K.: A novel robust qca full-adder. Procedia Materials Science **11**, 376–380 (2015)

[36] Kianpour, M., Sabbaghi-Nadooshan, R., Navi, K.: A novel design of 8-bit adder/subtractor by quantum-dot cellular automata. Journal of computer and system sciences **80**(7), 1404–1414 (2014)

[37] Safoev, N., Jeon, J.-C.: Design of high-performance qca incrementer/decrementer circuit based on adder/subtractor methodology. Microprocessors and Microsystems **72**, 102927 (2020)

[38] Sarvaghad-Moghaddam, M., Orouji, A.A.: New symmetric and planar designs of reversible full-adders/subtractors in quantum-dot cellular automata. The European Physical Journal D **73**, 1–12 (2019)

[39] Vahabi, M., Lyakhov, P., Bahar, A.N., Wahid, K.A.: Design and implementation of new coplanar fa circuits without not gate and based on quantum-dot cellular automata technology. Applied Sciences **11**(24), 12157 (2021)

[40] Zoka, S., Gholami, M.: A novel efficient full adder–subtractor in qca nanotechnology. International Nano Letters **9**(1), 51–54 (2019)

[41] Balali, M., Rezai, A.: Design of low-complexity and high-speed coplanar four-bit ripple carry adder in qca technology. International Journal of Theoretical Physics **57**(7), 1948–1960 (2018)

[42] Heikalabad, S.R., Asfestani, M.N., Hosseinzadeh, M.: A full adder structure without cross-wiring in quantum-dot cellular automata with energy dissipation analysis. The Journal of Supercomputing **74**, 1994–2005 (2018)

[43] Zahmatkesh, M., Tabrizchi, S., Mohammadyan, S., Navi, K., Bagherzadeh, N.: Robust coplanar full adder based on novel inverter in quantum cellular automata. International Journal of Theoretical Physics **58**, 639–655 (2019)

[44] Chan, S.T.Y., Chau, C.F., bin Ghazali, A.: Design of a 4-bit ripple adder using quantum-dot cellular automata (qca). In: 2013 IEEE International Conference on Circuits and Systems (ICCAS), pp. 33–38 (2013). IEEE

[45] De, D., Das, J.C.: Design of novel carry save adder using quantum dot-cellular automata. Journal of computational science **22**, 54–68 (2017)

[46] Labrado, C., Thapliyal, H.: Design of adder and subtractor circuits in majority logic-based field-coupled qca nanocomputing. Electronics letters **52**(6), 464–466 (2016)

[47] Mokhtari, D., Rezai, A., Rashidi, H., Rabiei, F., Emadi, S., Karimi, A.: Design of novel efficient full adder architecture for quantum-dot cellular automata technology. Facta Universitatis, Series: Electronics and Energetics **31**(2), 279–285 (2018)

[48] Safoev, N., Jeon, J.-C.: Compact rca based on multilayer quantum-dot

cellular automata. In: Information Systems Design and Intelligent Applications: Proceedings of Fourth International Conference INDIA 2017, pp. 515–524 (2018). Springer

[49] Srivastava, S., Asthana, A., Bhanja, S., Sarkar, S.: Qcapro-an error-power estimation tool for qca circuit design. In: 2011 IEEE International Symposium of Circuits and Systems (ISCAS), pp. 2377–2380 (2011). IEEE

[50] Kumar K, S., et al.: Design of area-delay efficient adder based circuits in quantum dot cellular automata. International Journal on Intelligent Electronics Systems **9**(2) (2015)

[51] Rani, T., Karmakar, S., Metri, A.A., Sharan, P.: Comparative study of half adder and subtractor circuits based on quantum-dot cellular automata (qca). In: International Conference on Advances in Computer & Communication Engineering, vol. 3, pp. 61–67 (2015)

[52] Heikalabad, S.R., Asfestani, M.N., Hosseinzadeh, M.: A full adder structure without cross-wiring in quantum-dot cellular automata with energy dissipation analysis. The Journal of Supercomputing **74**, 1994–2005 (2018)

[53] Ahmad, F., Bhat, G.M., Khademolhosseini, H., Azimi, S., Angizi, S., Navi, K.: Towards single layer quantum-dot cellular automata adders based on explicit interaction of cells. Journal of Computational Science **16**, 8–15 (2016)

[54] Sandhu, A., Gupta, S.: Performance evaluation of an efficient five-input majority gate design in qca nanotechnology. Iranian Journal of Science and Technology, Transactions of Electrical Engineering, 1–12 (2019)

[55] Moosavi, N., Navi, K., Aghazarian, V.: Ultra-low cost full adder cell using the nonlinear effect in four-input quantum dot cellular automata majority gate. International Journal of Nonlinear Analysis and Applications **11**(2), 1–16 (2020)

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/368575895

A review of dissimilar aluminum alloys welding via Friction stir welding

Article · February 2023

DOI: 10.36037/IJREI.2022.7102

CITATIONS
0

READS
25

3 authors, including:

Sumit Jain
Delhi Technological University
13 PUBLICATIONS **36** CITATIONS
SEE PROFILE

Dr- Husain Mehdi
Meerut Institute of Engineering & Technology
59 PUBLICATIONS **671** CITATIONS
SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Friction Stir Welding View project

All content following this page was uploaded by Dr- Husain Mehdi on 17 February 2023.

The user has requested enhancement of the downloaded file.

International Journal of Research in Engineering and Innovation (IJREI)
journal home page: http://www.ijrei.com
ISSN (Online): 2456-6934

ORIGINAL ARTICLE

A review of dissimilar aluminum alloys welding via Friction stir welding

Sumit Jain, R.S. Mishra

Department of Mechanical Engineering, Delhi Technological University, Delhi, India

Article Information

Received: 06 Nov 2022
Revised: 21 Jan 2023
Accepted: 12 Feb 2023
Available online: 17 Feb 2023

Keywords:

Friction stir welding
Aluminium alloys
Tensile strength
Microhardness

Abstract

In the present study, a review was carried out on the friction stir welding of dissimilar aluminum alloys. Friction stir welded joints hold promise due to its low welding heat input and capacity to reduce the degree of intermetallic compound (IMC) production in dissimilar metals. The goal of the current paper was to evaluate the effect of various processing parameters like tool rotating speed, welding/transverse speed, tool pin profile, tool tilt angle and shoulder diameter etc. on the weld quality of dissimilar aluminum alloys. From the present study it can be concluded the processing parameters significantly affects the mechanical characteristics of the weld quality of dissimilar aluminium alloys. This study will serve as a foundation for researchers to take FSW into account for a variety of diverse aluminum alloys.
©2023 ijrei.com. All rights reserved

1. Introduction

Friction stir welding (FSW) is a solid-state method, which means that the joining materials do not melt during welding. It is utilized when the original qualities of metal must be preserved to the greatest extent feasible. This is due to the mechanical intermixing of the two metal pieces, which allows them to be fused by softening them with mechanical pressure [1-4]. In FSW, a revolving tool of cylindrical shape is used, which moves along the parting line of two metal plates. Frictional heat is generated from friction between the work piece and the rotating tool, allowing the material to soften without melting under mechanical pressure. The weld is generated when the material is deformed at temperatures less than the melting point. The traverse and rotation speed of the FSW tool generates a specific imbalance between the neighbouring surfaces throughout the welding process. The advancing side (AS) is where the tool rotation aligns welding tool translation, whereas the retreating side is where the two motions, rotation and translation, oppose each other [5-8]. Various investigations have been carried out on the welding of different aluminum alloys using different combinations of processing parameters of FSW like tool rotating speed (TRS), welding speed (WS), tool shoulder diameter, tool tilt angle and tool pin profile etc., the effect of processing parameters on the mechanical characteristics of the weld joint was investigated. Sadeesh et al. investigated the impact of TRS, WS and five-pin profiles with a 2° constant tool pin angle on the tensile strength, hardness and microstructure of weld joint dissimilar AA2041/AA6061 [9]. Ratnam et al. examined the effect of twin pin profiled tools with different TRS and WS to weld dissimilar AA2041/AA6061 [10]. Ying Li et al. studied the solid-state flow during FSW of AA6061 and AA2024 [3]. Babu et al. revealed the effect of the advancing side of position in the case of dissimilar AA2041/AA6061 on ultimate tensile strength (UTS) [11, 12]. Anil Kumar et al. compared the result of similar welding of AA5083 and 6082 with dissimilar weld joints of both and investigated the effect of TRS, WS and tool

Corresponding author: Sumit Jain
Email Address: sumitjain4@gmail.com
https://doi.org/10.36037/IJREI.2022.7102.

tilt angle on the joint efficiency of AA5083 with AA6082 [13]. Dragatogiannis et al. They have added Tic micro and nano-particles to enhance the weld joint's mechanical properties. The effect of TRS, WS and tool tilt angle variation on the joint strength was also investigated [14]. Reinforcement particles were mixed with ethanol for easy filling and managing reinforcement, revealing TiC reinforcement's effect on corrosion resistance [15]. Jain et al. optimized the process parameters of FSW, which were TRS and WS. Shoulder diameter and pin profile using the Taguchi method for single response optimization and grey relation with weight method for multi-response optimization [16]. Tasman et al. studied the effect of the constant ratio of TRS to WS with triangular and pentagonal tool pins on mechanical properties [17]. Various researchers have researched joining aluminum alloy 5083 with aluminum alloy 6061 because both are most widely used in marine industries [18-21]. FSW of AA5083-O with 6061-T6 carried at four different tool rotation speeds of 450,560, 710 & 900 rpm and the joint strength was compared with the joint strength of welding carried by fusion welding, i.e., TIG and MIG welding with a filler rod of AA4043 [18]. Comparison of FSW of AA5083/AA5083, AA6061/AA6061 and AA5083/AA6061 done on rotational tool speed of 630, 1600 rpm and welding speed of 16, 25 and 40 mm/min. Dissimilar welding of AA6061/AA5083 was performed by taking AA 6061 on AS. 13 experiments were performed to make similar and dissimilar joints of AA6061 and AA5083 [19]. Davaiah et al. join AA5083-H321 with AA6061; both have a thickness of 5 mm at four different WS of 40, 63, 80, and 100 mm/min with constant TRS of 1120 rpm, 2.5-degree tool tilt angle and axial force keeping AA6061 on AS and AA5083 on RS in all experiments [20].

2. Tool used

Generally, tool made of H13 steel is used for joining aluminum similar and dissimilar alloys. The difference consisted of pin and shoulder profiles with different sizes per thickness and base material properties. A high-speed tool with nickel coating having a taper square cross-section was used by as shown in Fig. 1 [13]. H13 steel tools of five-pin shapes were used to check the effect on weld quality shown in Fig. 2 [9]. The effect of four pin profiles, i.e., cylindrical threaded, square, cylindrical, and trapezoidal, made up of high carbon steel, was revealed and considered as a process parameter in FSW shown in Fig. 3 [16]. Davaiah used H13 steel with cylindrical threaded taper pin and shoulder of concave scrolled contour having a diameter of 18mm and pin length and diameter of 4.7 mm and 6mm, respectively [20].

Figure 1: Square pin profiled tool [13]

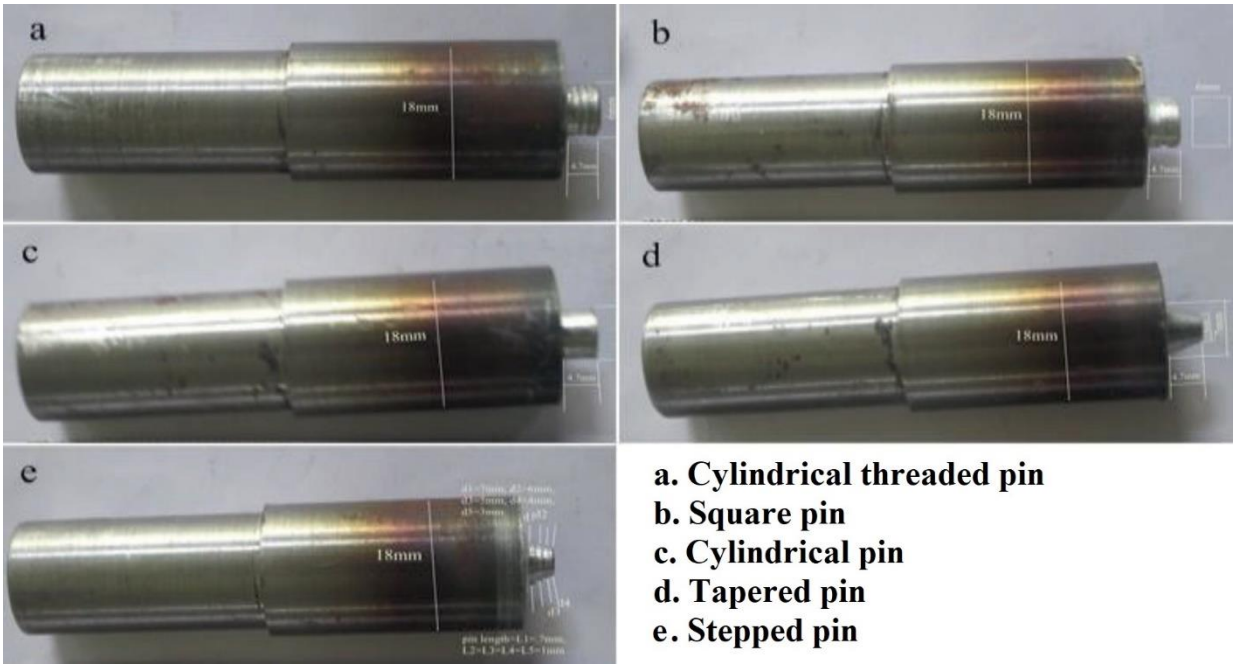


Figure 2: Pictorial view of different tool pins used [9]

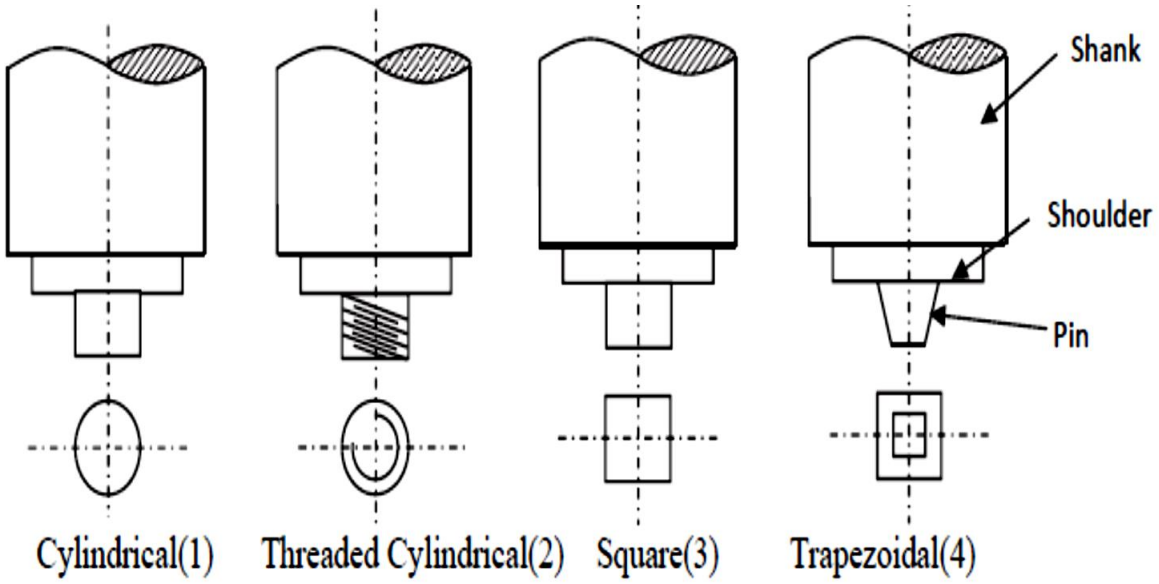


Figure 3. Different tool pin profiles used in FSW of AA6082/AA5083 [4]

Table 1: Literature review of FSW of aluminium alloys

S.No	Authors name	Output responses	Materials used	Conclusions
1	Krishna et al. (2014)	Yield strength, Tensile strength and % elongation	AA6351 and AA5083	Describes the effect of an FSW process, including butt joining of identical AA6351 and AA6351 combinations and dissimilar AA6351 and AA5083 combinations on tensile, hardness, and impact behavior. When compared to different alloy combinations, the tensile, hardness, and impact characteristics of Aluminum alloys of comparable alloy combinations show better results [22-25].
2	Jerome.S. et al. (2012)	Microstructural Observations. Microhardness Details	AA5083 /TiC particles	The hardness profiles of the treated samples were assessed along the top surface and across the cross-section. Compared to the base metal (88Hv), the average hardness along the top surface increased by 27.27 % [26].
3	Klobcar.D. et al. (2012)	Microstructure, Hardness, Tensile properties	AA 5083 (T-4-mm)	The microstructure was prepared for examination under a polarized light source using a light microscope. At an FPR of 0.35 mm/r, a set of ideal welding conditions was identified, allowing excellent welds to be formed with a slight improvement in hardness and a 15% reduction in tensile strength [27].
4	El-Danaf,et al. (2014)	Hardness, Tensile Test	AA5083 and AA1100(T- 6- mm)	Tool geometries substantially impact dissimilar material FSW weld quality. In the case of taper cylindrical tools with the same process parameter, both hardness and tensile strength were greater. In dissimilar friction stir weldments, lower TS combined with more incredible RTS produce the highest tensile characteristics [28].
6	Johannes et al. (2007)	Microstructure examination Tensile testing	AA5083.	By modifying the size and distribution of component particles before cold rolling, FSP assists in the reduction of post-recrystallization grain size. It was discovered that adding the FSP stage refined the recrystallized grain size, decreased flow stresses and enhanced elongations [29].
7	Scialpi. et al. (2006)	Mechanical and Microstructural properties	AA 6082 T6	A study of three shoulder geometries was conducted in this study. The tool analysis was performed on 1.5 mm thick AA 6082 T6 sheets. The tool was rotated at 1810 rpm with a federate of 460 mm/min during the welding operation. The analysis revealed that a shoulder welded the optimum joint for thin sheets with fillet and cavity [30].
8	Amini.S., (2015)	Vertical force and welding force	FSW on AA5083 with dimensions of 120 mm×60	The impact of an offset pin on vertical force and welding force reduction (between 50 to 70 %) is larger than the impact of a concentric pin with the tool shoulder axis on these forces. Tools

			mm×4 mm.	with a half pin and an arching pin exert greater force than tools with an offset pin but less force than tools with a concentric pin [31].
9	Suresha. et al. (2010)	Tensile strength, joint efficiency	AA7075-T6 of 5mm thickness	The welded joints of the conical tool are more efficient than the square tool's. ANOVA was used to determine the percentage contribution of these FSW process parameters, and it was observed that tool rotating speed, when compared to weld traverse speed and plunge depth, has a substantial contribution in both conical and square tools [32].
10	Kandasam y.J, (2012)	Yield strength, Tensile strength % Elongation, Hardness, Distribution	6mm thick AA7075 and AA6061	The experiment's goal is to reduce the temperature difference between the bottom surfaces and the top of the plates as much as possible. The creation of Al2Cu and Al4Cu9 IMCs increases bond strength, according to the analysis of the generated intermetallic compound (IMC) [33].
12	Mishra et al. (1999)	Strain rate, Tensile tests	AA 7075-T6	The current findings show that friction stir processing may create a microstructure in a commercial aluminum alloy that is susceptible to high strain rate superplasticity [34].
14	Bahemma t. et al. (2015)	Ultimate Tensile strength, Elongation, Hardness,	AA7075-O and AA2024- T4	The microstructure analysis aims to see how the pin profile and rotating speed affect grain size. Furthermore, one of the essential aims of the current study is to get high-quality welds with the least amount of money [35].
15	Dongxiao. et al. (2015)	Microstructural characteristics, welding process, Hardness tensile strength	7075-T651 aluminumallo y	Both the existence of precipitates and the displacement of component particles may be blamed for the reduced crack initiation energy in the HAZ. The presence of precipitates in the HAZ was the primary cause of the SSFSW joint's lower tensile strength [36].
16	Babu. et al. (2009)	Tensile strength	AA2219 aluminium alloy	With a 95% confidence level, a mathematical model was built to estimate the tensile strength of FSWed AA2219 joints [37].
17	Nam et al. (2016)	Potential dynamic and EIS studies	6061 aluminum alloy	According to the findings, a high-quality thin coating on the surface with improved properties resulted in a high-corrosion-resistant alloy. This might be owing to the FSW's homogenous impurity distribution in the alloy, on which a homogeneous passive coating was generated to improve corrosion resistance [38].
18	Guillo. et al. (2016)	Inbuilt real-time algorithm	AA5754-H22	According to this article, a robot with an inbuilt real-time algorithm for compensating lateral tool deviation may recreate the same FSW equality as a gantry-type CNC system [39].
19	Zhang et al.. (2016)	tensile strength and hardness	AA2219-T6	The ideal method to increase the mechanical characteristics of FSW2219Al-T6 joints has been demonstrated to be a combination of post-welding artificial aging, and high welding speed, with the highest joint efficiency of 91 percent attained [40].
20	Mehtaa et al. (2016)	defects	AA6061-T651	Welds formed the most erratic and substantial copper particles with triangular pin geometries. In addition, polygonal pin profiles produced holes, tunnels, fractures, and fragmentary faults independent of their static and dynamic constant regions. Also, defects decreased as the number of polygonal edges grew. The macro joint on a cylindrical tool pin profile has been claimed to be devoid of flaws [41-43].
21	Aval et al. (2011)	Temperature distribution, yield & tensile effect.	AA5086	The final microstructures and mechanical characteristics of welded alloys can be significantly influenced by work-hardened and annealed conditions [44].
22	Heurtier, et al. (2005)	Microhardness, estimations of the temperatures strains and strain rates.	AA2024-T351	The semi-analytical model may be employed to calculate stresses, micro-hardness, strain rates, and temperature estimates in different weld zones [45].
23	Raghu Babu, et al. (2008)	tensile strength, hardness and microstructure	AA6082-T6	The joint's tensile strength is less than that of the parent metal. And it's proportional to the speed of travel / welding [46].
24	Zhang et al. (2007)	Material flow	AA 6061 -T6	There appears to be a quasi-linear relationship between the variation of the equivalent plastic strain and the variation of the applied loads on the shoulder. On increasing in the pin's translational and angular velocity, the material flow may be

				accelerated [47].
25	Deplus et al. (2012)	MRR and Surface Finish	AA 2024-T3, AA5754- H111 and AA6082-T6	The static deflection of the wire and vibrational behaviour have been reported to produce inaccuracies [46]. intelligent systems or Expert knowledge have been stated to lessen the inaccuracy due to the static deflection of the wire and vibrational behaviour.
26	Cavalier et al. (2006)	Tensile, fatigue strength, hardness and microstructure	AA2024-7075	The two sheets were successfully welded one after the other, and the welded sheets were tested under strain at room temperature to determine their mechanical reaction in comparison to the parent materials [49].
27	Scialpi et al. (2007)	Tensile strength, microstructure	AA6082	According to the findings, the best joint for thin sheets was welded by a shoulder having cavity and fillet [50-52].
29	Kulekci, et al. (2006)	fatigue strength and microstructure	AA5754	The fatigue strength of joints is reduced when tool rotation is increased for a fixed tool pin diameter. The fatigue strength of joints is reduced when tool pin diameter is increased for a fixed tool rotation [53].
30	Prado et al. (2003)	Tool wear and wear rate	AA6061+20% Al2O3	When weld or traverse speeds are increased, tool wear and the wear rate are found to decrease [54].
31	Fratini,L. (2005)	strain, strain rate and temperature	AA 6082-T6	An inverse-identification technique based on a linear regression methodology was employed to provide the required material characterization. [55].

3. Results and discussion

3.1 Tensile strength

The joint efficiency of dissimilar weld joints of AA2024/AA6061 was found to be 80% with cylindrical threaded tool pins and 81% with square tool pins [9]. Higher welding speed with a twin pin tool induces the maximum UTS. AA2024 on AS gives the maximum UTS compared to AA6061 on AS [10]. UTS of the weld joint of AA5083 with 6061 by FSW was higher than the joint strength by fusion welding. The maximum UTS by FSW obtained was 200 MPa which was higher than the 160 MPa and 192 MPa obtained by TIG and MIG welding, respectively [13]. On performing the tensile test, the joints failed at the Thermo-mechanical heat affected zone and nugget zone, The UTS of the weld joint of AA5083/AA5083 found 57.1% of the UTS of base material and the best result obtained of the joint of AA6061/AA6061 with joint strength of 77.7% compared with base material of AA 6061-T4. The dissimilar sound joints of AA6061-T4 and AA5083-O was obtained at 1600 rpm with joint strength of 71.2% joint efficiency. The decreased heat input on increasing welding speed showed better sound weld quality. At higher welding speeds with a higher cooling rate lacking mixing occurs because the material on the advancing side travels much less into the retreating side and causes defects that decrease UTS [14]. The joint efficiency of dissimilar welding of AA6082/AA5083 with joint efficiency of 85% compared with the joint efficiency of AA5083/AA5083 and AA6082/AA6082 joint efficiency of 56% and 62%, respectively [14]. The UTS yield strength and elongation (El) can also be increased by adding TiC reinforcement in the stir zone (SZ) of FSWed AA6082/AA5083[15]. Process parameters optimization for UTS obtained by Taguchi were 1200 rpm of TRS, 35mm/min of welding and 14mm tool shoulder diameter. The multi-response optimization results were obtained by the GRA-based weight method and were calculated based on different weightage to different responses. The optimum parameters for 0.5 weightage of UTS and elongation were at 1200 rpm TRS, 35 mm/min WS, and 14 mm shoulder diameter with a trapezoidal pin profile. On giving 0.7 weightage to UTS and 0.3 to El optimum parameters obtained were 1200 rpm TRS, 35mm/min WS, and 16mm shoulder diameter with square tool pin; if 0.3 weightage given to UTS and 0.7 to El, then optimum parameters obtained were 1200 rpm TRS, 35mm/min WS and14mm shoulder diameter with trapezoidal pin profile [16]. The triangular pin gives better results on UTS compared with the pentagonal pin profile and increases on increasing TRS and WS [17].

3.2 Microhardness

The hardness of the nugget zone is higher than TMAZ and AA6061 but lower than AA2024 [12]. The hardness of similar weld joints of AA5083/AA5083 and AA6082/AA6082 was compared with the dissimilar FSWed joint of AA 5083/AA6082, and it was found that the hardness of Weld zone (WZ) in the case of FSWed AA5083/AA 6082 was higher than FSWed AA5083/AA5083 and AA6082/AA6082 shown in [13]. Fig. 4 shows the hardness distribution profile of the TIG and TIG+FSP weldments of AA5083 and AA8011 [56]. The microstructure of the welded area may be connected to the microhardness variation. The variances in hardness levels indicate that grain structures are not homogeneous. The dominant component in determining the hardness value may be the grain structure or size, although it is not the only one. The hardness of the welded zone is impacted by temperature distribution, strain rate, and particle density [17]—increased strain hardening results from a drop in temperature. As a result, TIG+FSP welded joints underwent strain hardening due to decreased grain size and low-temperature strain rates. Due to the combination of base metal and ER4043 in the TIG weldment, the HAZ area shows the most inadequate hardness

in TIG welded connections. TIG + FSP welded joints had a hardness distribution pattern comparable to yet distinct from TIG welded joints. The insufficient material flow created by the FSP tool is the cause of the inhomogeneity of microhardness un SZ [17]. The fine-grain structure was shown to have increased hardness at room temperature by the Hall-Petch correlation. TIG weldments showed a substantial variation compared to TIG + FSP welded connections. This is due to intermetallic phases precipitating in TIG joints made of coarse Mg2Si and Al13Fe4. Stronger than the Mg2Si and Al13Fe4 phases, the -Al phase possesses a large volume fraction. If the indenter were positioned near the primary -Al phase rather than the Mg2Si phase in the opposite situation, the hardness value would be reduced; a high hardness value was found.

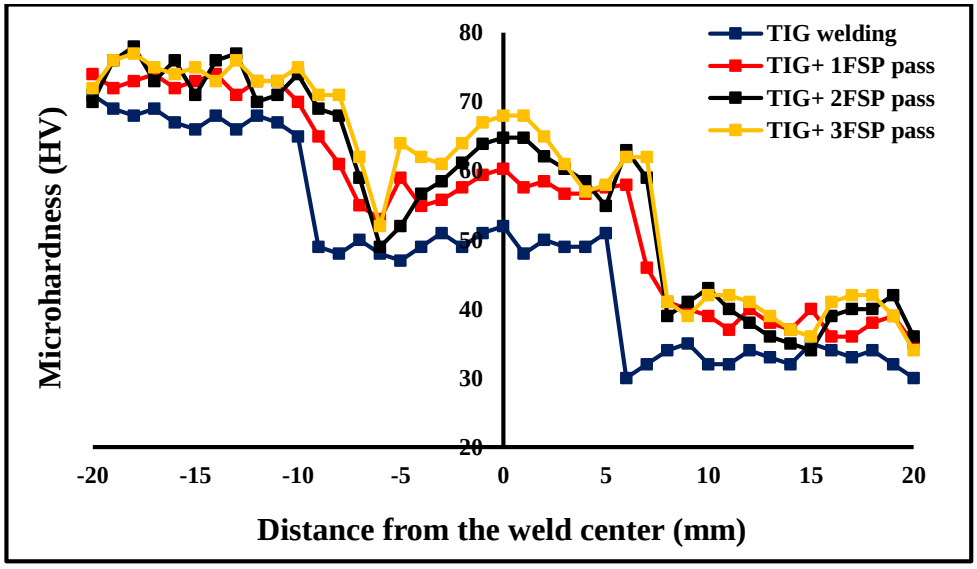


Figure 4: Hardness distribution profile of TIG and TIG+FSP welded joints of AA5083 and AA8011 [56]

3.3 Microstructure

The FSW reconstruction sectional view of the weld joint of AA2024 and AA6061 at 400 rpm is shown in Fig. 5 [11]. The microstructure of the TIG SV + FSP and TIG DV + FSP joints of AA8011 and AA6082 is shown in Fig. 6. Three elements come together to create FSPed joints. The first is a plastic material stirring, which gathers the plasticized multi-layered materials behind the FSP tool and is influenced by a complex interaction between traverse speeds, pin profiles, and RTS. The second occurs when the parent material in the SZ becomes softer because of a rise in temperature in the welded zone [57]. The third involves the FSP tool shoulder hot forging plasticized material. If any incorrect change of these parameters occurred, the flawed FSPed region would be discovered. The microstructure of AA6061 on RS exhibited elongated particles of Mg-Si and Al-Si in a matrix, and AA2024 on AS exhibited elongated particles of Al2Cu and Mg-Si in a matrix. Voids in a matrix were also noticed, as depicted in Figs. 7, 8 [12]. The hardness obtained by the FSW weld joint was lower in the heat-affected zone and higher in WZ [18]. Generally, the hardness decreases from top to bottom in the thickness of WZ, but there is an increase in the hardness from a distance of 1200 µm down from the top of the surface due to rapid cooling on the top surface as compared to the lower surface. And the harness decrease at a distance of about 1200 µm above the bottom surface. The micro-harness was lower on the advancing side because of higher heat input on the advancing side [19].

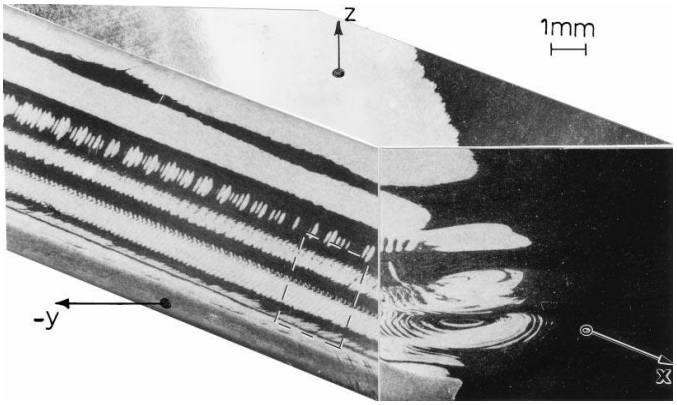


Figure. 5 Reconstruction view of weld joint of AA2024/AA6061 [11]

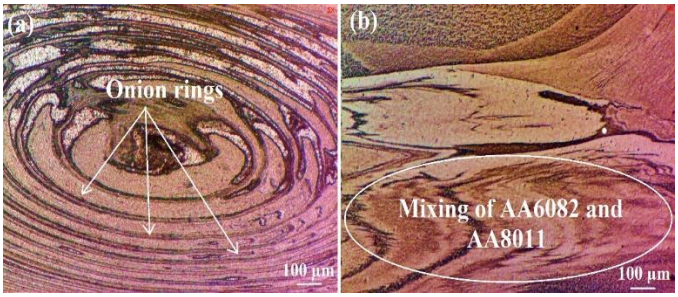


Figure 6: Microstructure of TIG+FSP welded joint, (a) TIG SV+FSP, (b) TIG DV+FSP [57]

The ductility and formability of aluminum alloys are unsatisfactory at room temperature, leading to early failure under challenging stress conditions. The hexagonal close-packed (HCP) crystal lattice's poor symmetry, high basal roughness, and restriction on the number of active slip systems all contribute to this [58-60]. Due to these factors, there is insufficient strength, severe fatigue, and little creep resistance.

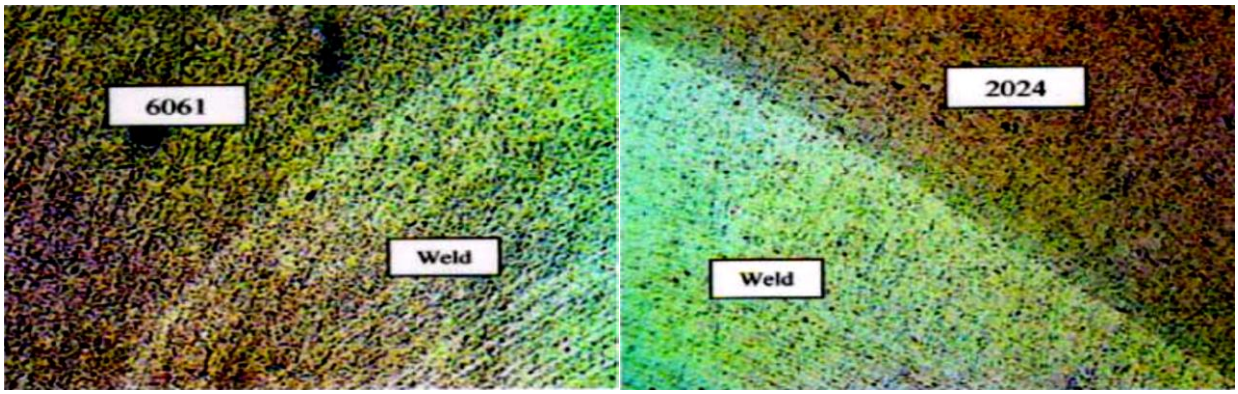


Figure. 7 Microstructure view of AA2024 on advancing side at WS of 35mm/min [12]

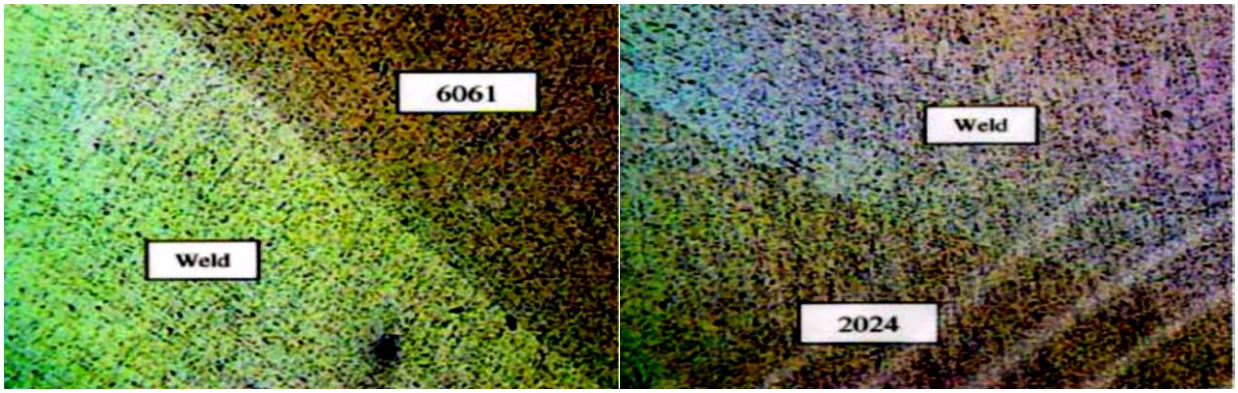


Figure 8: Microstructure view of AA6061 on advancing side at WS of 35 mm/min [12]

4. Conclusions

From the current study following conclusions can be drawn:

- The joint efficiency of dissimilar weld joint of AA2024/AA6061 found 80% with cylindrical threaded tool pin and 81% with square tool pin.
- The sound dissimilar joint of AA6061-T4 and AA5083-O obtained at 1600 rpm with joint strength of 71.2% joint efficiency.
- The triangular pin gives better result on UTS of AA5083/AA6082 on comparing with pentagonal pin profile and UTS increases on increasing TRS and WS.
- Positioning of materials (advancing/retreating side) is an important consideration in case of joining dissimilar aluminum alloys.
- The hardness of weld zone in case of FSWed AA5083/AA6082 was higher than FSWed AA5083/AA5083 and AA6082/AA6082.

References

[1] Husain Mehdi, R.S. Mishra, Mechanical properties and microstructure studies in Friction Stir Welding (FSW) joints of dissimilar alloy- a review, Journal of Achievements in Materials and Manufacturing Engineering, 77 (1), 31-40, (2016).

[2] Husain Mehdi, R.S. Mishra, Investigation of mechanical properties and heat transfer of welded joint of AA6061 and AA7075 using TIG+FSP welding approach, Journal of Advanced Joining Processes, 1, 100003, (2020) https://doi.org/10.1016/j.jajp.2020.100003.

[3] Sumit Jain, R.S. Mishra, Effect of Al2O3 nanoparticles on microstructure and mechanical properties of friction stir-welded dissimilar aluminum alloys AA7075-T6 and AA6061-T6, Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering, Volume 236, Issue 4 (2021).

[4] Husain Mehdi, R.S. Mishra, Study of the influence of friction stir processing on tungsten inert gas welding of different aluminum alloy. SN Applied Science, 1, 712 (2019). https://doi.org/10.1007/s42452-019-0712-0.

[5] Husain Mehdi, R.S. Mishra, Influence of Friction Stir Processing on Weld Temperature Distribution and Mechanical Properties of TIG-Welded Joint of AA6061 and AA7075. Transactions of the Indian Institute of Metals, 73, 1773–1788 (2020). https://doi.org/10.1007/s12666-020-01994-w.

[6] Jain, S., Mishra, R.S. Microstructural and mechanical behavior of micro-sized SiC particles reinforced friction stir processed/welded AA7075 and AA6061. Silicon 14, 10741–10753 (2022). https://doi.org/10.1007/s12633-022-01716-5

[7] Husain Mehdi, R.S. Mishra, Effect of Friction Stir Processing on Mechanical Properties and Wear Resistance of Tungsten Inert Gas Welded Joint of Dissimilar Aluminum Alloys. Journal of Materials Engineering and Performance volume, 30, 1926–1937 (2021). https://doi.org/10.1007/s11665-021-05549-y.

[8] Husain Mehdi, R.S. Mishra, Microstructure and mechanical characterization of tungsten inert gas-welded joint of AA6061 and AA7075 by friction stir processing, Proceedings of the Institution of Mechanical Engineers, Part L: Journal of Materials: Design and Applications, 235 (11), 2531-2546 (2021), https://doi.org/10.1177/14644207211007882.

[9] Sadeesh, P.; Venkatesh Kannan, M.; Rajkumar, V.; Avinash, P.; Arivazhagan, N.; Devendranath Ramkumar, K.; Narayanan, S. (2014). Studies on Friction Stir Welding of AA 2024 and AA 6061 Dissimilar Metals. Procedia Engineering, 75(), 145–149. doi:10.1016/j.proeng.2013.11.031.

[10] CH.Ratnam, B.Sudheer Kumar and K. N.D.Malleswara Rao, Effect of Twin-Pin Profile Tool on the Microstructure and Mechanical Properties of Friction Stir Welded Dissimilar Aa2024 and Aa6061 Aluminium Alloys, International Journal of Mechanical Engineering and Technology, 9(3), 2018, pp. 946–953.

[11] Ying Li; L.E Murr; J.C McClure (1999). Solid-state flow visualization in the friction-stir welding of 2024 Al to 6061 Al. , 40(9), 1041–1046. doi:10.1016/s1359-6462(99)00062-7.

[12] Sipokazi Mabuwa, Velaphi Msomi, Husain Mehdi, Kuldeep Kumar Saxena, Effect of material positioning on Si-rich TIG welded joints of AA6082 and AA8011 by friction stir processing, Journal of Adhesion Science and Technology, 2022. https://doi.org/10.1080/01694243.2022.2142366.

[13] Anil Kumar, H M; Venkata Ramana, V; Pawar, Mayur (2018). Experimental Study on Dissimilar Friction Stir welding of Aluminium Alloys (5083-H111 and 6082-T6) to investigate the mechanical properties. IOP Conference Series: Materials Science and Engineering, 330(), 012076–. doi:10.1088/1757-899X/330/1/012076.

[14] Dragatogiannis, D. A.; Koumoulos, E. P.; Kartsonakis, I.; Pantelis, D. I.; Karakizis, P. N.; Charitidis, C. A. (2015). Dissimilar Friction Stir Welding Between 5083 and 6082 Al Alloys Reinforced with Tic Nanoparticles. Materials and Manufacturing Processes, (), 10426914.2015.1103856–. doi:10.1080/10426914.2015.1103856.

[15] Kartsonakis, I.A.; Dragatogiannis, D.A.; Koumoulos, E.P.; Karantonis, A.; Charitidis, C.A. (2016). Corrosion behaviour of dissimilar friction stir welded aluminium alloys reinforced with nanoadditives. Materials & Design, 102, 2016, 56-67.

[16] Jain, Sumit; Sharma, Neeraj; Gupta, Rajat (2018). Dissimilar alloys (AA6082/AA5083) joining by FSW and parametric optimization using Taguchi, grey relational and weight method. Engineering Solid Mechanics, (), 51–66. doi:10.5267/j.esm.2017.10.003.

[17] Kasman, Sefika; Kahraman, Fatih; Emiralioğlu, Anıl; Kahraman, Haydar (2016). A Case Study for the Welding of Dissimilar EN AW 6082 and EN AW 5083 Aluminum Alloys by Friction Stir Welding. Metals, 7(1), 6–. doi:10.3390/met7010006.

[18] Abdul Wahab Hashmi, Husain Mehdi, Sipokazi Mabuwa, Velaphi Msomi & Prabhujit Mohapatra, Influence of FSP Parameters on Wear and Microstructural Characterization of Dissimilar TIG Welded Joints with Si-rich Filler Metal. Silicon, 14, 11131–11145, 2022. https://doi.org/10.1007/s12633-022-01848-8

[19] Abdul Wahab Hashmi, Husain Mehdi, R. S. Mishra, Prabhujit Mohapatra, Neeraj Kant & Ravi Kumar, Mechanical Properties and Microstructure Evolution of AA6082/Sic Nanocomposite Processed by Multi-Pass FSP. Transactions of the Indian Institute of Metals, (2022). https://doi.org/10.1007/s12666-022-02582-w.

[20] Husain Mehdi, Arshad Mehmood, Ajay Chinchkar, Abdul Wahab Hashmi, Chandrabhanu Malla, Prabhujit Mohapatra, Optimization of process parameters on the mechanical properties of AA6061/Al2O3 nanocomposites fabricated by multi-pass friction stir processing, Materials Today: Proceedings, 56 (4), 1995-2003, 2021, https://doi.org/10.1016/j.matpr.2021.11.333.

[21] Husain Mehdi, R.S. Mishra, Modification of Microstructure and Mechanical Properties of AA6082/ZrB2 Processed by Multipass Friction Stir Processing. Journal of Materials Engineering and Performance (2022). https://doi.org/10.1007/s11665-022-07080-0.

[22] Krishna, G.Gopala & Reddy, P.Ram & Hussain, M Manzoor. (2014). Mechanical Behaviour of Friction Stir Welding Joints of Aluminium alloy of AA6351 with AA6351 and AA6351 with AA5083. International Journal of Engineering Trends and Technology. 10. 161-165. 10.14445/22315381/IJETT-V10P231.

[23] Husain Mehdi, R.S. Mishra, Effect of Friction Stir Processing on Microstructure and Mechanical Properties of TIG Welded Joint of AA6061 and AA7075, Metallography, Microstructure, and Analysis, 9, 403–418 (2020). https://doi.org/10.1007/s13632-020-00640-7

[24] Husain Mehdi, R.S. Mishra, Effect of friction stir processing on mechanical properties and heat transfer of TIG welded joint of AA6061 and AA7075, Defence Technology, 17 (3), 715-727 (2021). https://doi.org/10.1016/j.dt.2020.04.014

[25] Husain Mehdi, R.S. Mishra, Consequence of reinforced SiC particles on microstructural and mechanical properties of AA6061 surface composites by multi-pass FSP, Journal of Adhesion Science and Technology, 36(12), 1279-1298, 2022, https://doi.org/10.1080/01694243.2021.1964846.

[26] Savarimuthu, Jerome & Bhalchandra, S & Babu, Kumaresh & Ravisankar, B. (2012). Influence of Microstructure and Experimental Parameters on Mechanical and Wear Properties of Al- TiC Surface Composite by FSP Route. Journal of Minerals and Materials Characterization and Engineering. 11. 10.4236/jmmce.2012.115035.

[27] Klobčar, Damjan & Kosec, Ladislav & Pietras, Adam & Smolej, Anton. (2012). Friction-stir welding of aluminium alloy 5083. Materiali in Tehnologije. 46. 483-488.

[28] El-Danaf, Ehab & El Rayes, Magdy & Soliman, Mahmoud. (2010). Friction stir processing: An effective technique to refine grain structure and enhance ductility. Materials & Design. 31. 1231-1236. 10.1016/j.matdes.2009.09.025.

[29] Johannes, L.B. & Charit, Indrajit & Mishra, Rajiv & Verma, Ravi. (2007). Enhanced superplasticity through friction stir processing in continuous cast AA5083 aluminum. Materials Science and Engineering A-structural Materials Properties Microstructure and Processing - MATER SCI ENG A-STRUCT MATER. 464. 351- 357. 10.1016/j.msea.2007.02.012.

[30] Scialpi, Agostino & De Filippis, Luigi & Cavaliere, P.. (2007). Influence of shoulder geometry on microstructure and mechanical properties of friction stir welded 6082 aluminium alloy. Materials & Design. 28. 1124-1129. 10.1016/j.matdes.2006.01.031.

[31] Amini, S. & Amiri, Mohsen. (2015). Pin axis effects on forces in friction stir welding process. The International Journal of Advanced Manufacturing Technology. 78. 10.1007/s00170-015-6785-z.

[32] C. N. Suresha, B. M. Rajaprakash & Sarala Upadhya (2011) A Study of the Effect of Tool Pin Profiles on Tensile Strength of Welded Joints Produced Using Friction Stir Welding Process, Materials and Manufacturing Processes, 26:9, 1111-1116.

[33] J. Kandasamy, M. Manzoor Hussain & S. Rajesham (2012) Heterogeneous Friction Stir Welding: Improved Properties in Dissimilar Aluminum Alloy Joints through Insertion of Copper Coupled with External Heating, Materials and Manufacturing Processes, 27:12, 1429- 1436, DOI: 10.1080/10426914.2012.709340.

[34] Mishra, R S, Mahoney, M W, McFaden, S X, Mara, N A, and Mukherjee, A K. High strain rate superplasticity in a friction stir processed 7075 Al alloy. United States: N. p., 1999. Web. doi:10.1016/S1359-6462(99)00329-2.

[35] Bahemmat, Pouya & Haghpanahi, Mohammad & Besharati Givi, Mohammad Kazem & Reshad Seighalani, Kambiz. (2012). Study on dissimilar friction stir butt welding of AA7075-O and AA2024- T4 considering the manufacturing limitation. International Journal of Advanced Manufacturing Technology, 59. 10.1007/s00170-011-3547-4.

[36] Dongxiao Li, Xinqi Yang, Lei Cui, Fangzhou He, Xu Zhang, Investigation of stationary shoulder friction stir welding of aluminum alloy 7075-T651, Journal of Materials Processing Technology, Volume 222, 2015, Pages 391-398.

[37] Rambabu, Gundla & Naik, D. & Ch, Venkata & Rao, K. & Reddy, G.. (2015). Optimization of Friction Stir Welding Parameters for Improved Corrosion Resistance of AA2219 Aluminum Alloy Joints. Defence Technology. 53. 10.1016/j.dt.2015.05.003.

[38] Nam, N.D. & Le, Dai & Mathesh, Motilal & Bian, M.Z. & Thu, Vu. (2016). Role of friction stir welding – Traveling speed in enhancing the corrosion resistance of aluminum alloy. Materials Chemistry and Physics. 173. 10.1016/j.matchemphys.2016.02.004.

[39] Mario Guillo, Laurent Dubourg, Impact & improvement of tool deviation in friction stir welding: Weld quality & real-time compensation on an

industrial robot, Robotics and Computer- Integrated Manufacturing, Volume 39, 2016, Pages 22-31.

[40] Z. Zhang, B.L. Xiao, Z.Y. Ma, Enhancing mechanical properties of friction stir welded 2219Al-T6 joints at high welding speed through water cooling and post-welding artificial ageing, Materials Characterization, Volume 106, 2015, 255-265.

[41] Mehta, Kush. (2016). Effects of Tool Pin Design on Formation of Defects in Dissimilar Friction Stir Welding. Procedia Technology. 23. 513-518. 10.1016/j.protcy.2016.03.057.

[42] A.Nait Salah, Husain Mehdi, Arshad Mehmood, Abdul Wahab Hashmid, Chandrabhanu Malla, Ravi Kumar, Optimization of process parameters of friction stir welded joints of dissimilar aluminum alloys AA3003 and AA6061 by RSM, Materials Today: Proceedings, 56 (4), 1675-1684, 2021, https://doi.org/10.1016/j.matpr.2021.10.288.

[43] Husain Mehdi, R.S. Mishra, R.S. Analysis of Material Flow and Heat Transfer in Reverse Dual Rotation Friction Stir Welding: A Review, International Journal of Steel Structures, 19, 422–434 (2019). https://doi.org/10.1007/s13296-018-0131-x.

[44] Jamshidi Aval, Hamed & Serajzadeh, S. & Kokabi, Amir Hossein. (2011). Theoretical and experimental investigation into friction stir welding of AA 5086. The International Journal of Advanced Manufacturing Technology. 52. 531-544. 10.1007/s00170-010- 2752-x.

[45] Heurtier,H., Jones, M.J., Desrayaud, C., Driver, J.H., Montheillet,F.& Allehaux,D. (2005). Mechanical and thermal modelling of Friction Stir Welding. Journal of Materials Processing Technology, 171, (2006), 348–357.

[46] Raghu Babu, G., Murti, K.G.K., Janardhana, G.R. (2008). An Experimental Study On The Effect Of Welding Parameters On Mechanical And Microstructural Properties Of Aa 6082-T6 Friction Stir Welded Butt Joints. Asian Research Publishing Network (Arpn), 3, 5, 1819-6608.

[47] Zhang, H., Zhang, Z., Chen, J. (2006). 3D modeling of material flow in friction stir welding under different process parameters. Journal of Materials Processing Technology, 183, 62–70.

[48] Deplus, K., Simar, A., Haver, W.V. et al. Residual stresses in aluminium alloy friction stir welds. Int J Adv Manuf Technol 56, 493–504 (2011). https://doi.org/10.1007/s00170-011- 3210-0.

[49] Cavalier,P., Nobile,R., Panella, F.W.& Squillace, A. (2005). Mechanical and microstructural behaviour of 2024–7075 aluminium alloy sheets joined by friction stir welding. International Journal of Machine Tools & Manufacture , 46, (2006), 588–594.

[50] Scialpi,A., De Filippis,L.A.C., Cavaliere,P.(2006). Influence of shoulder geometry on microstructure and mechanical properties of friction stir welded 6082 aluminium alloy. Materials and Design, 28,(2007),1124–1129.

[51] A.Nait Salah, M. Kaddami, Husain Mehdi, Mechanical Properties And Microstructure Characterization of Friction Stir Welded Joint Of Dissimilar Aluminum Alloy AA2024 And AA7050, Turkish Journal of Computer and Mathematics Education, 12(7), 1051-1061 (2021).

[52] Preety Rani, R.S. Mishra, Husain Mehdi, Effect of Nano-sized Al2O3 particles on microstructure and mechanical properties of aluminum matrix composite fabricated by multipass FSW, Part C: Journal of Mechanical Engineering Science (SAGE), 2022. https://doi.org/10.1177/09544062221110822.

[53] Kulekci, M.K.,Sik,A., Kaluc,E. (2006). Effects of tool rotation and pin diameter on fatigue properties of friction stir welded lap joints. Int J Adv Manuf Technol, (2008), 36,877–882.

[54] Prado, R.A., Murr, L.E., Soto, K.F. & McClure, J.C. (2002). Self-optimization in tool wear for friction-stir welding of Al 6061_/20% Al2O3 MMC. Materials Science and Engineering ,A, 349, (2003), 156-/165.

[55] Fratini, L., Buffa, G. (2004). CDRX modelling in friction stir welding of aluminium alloys. International Journal of Machine Tools & Manufacture, 45, (2005), 1188–1194.

[56] Salah, A.N., Mabuwa, S., Husain Mehdi. et al. Effect of Multipass FSP on Si-rich TIG Welded Joint of Dissimilar Aluminum Alloys AA8011-H14 and AA5083-H321: EBSD and Microstructural Evolutions. Silicon 14, 9925–9941 (2022). https://doi.org/10.1007/s12633-022-01717-4

[57] Husain Mehdi, Sipokazi Mabuwa, Velaphi Msomi, Kuldeep kumar Yadav, Influence of Friction Stir Processing on the Mechanical and Microstructure Characterization of Single and Double V-Groove Tungsten Inert Gas Welded Dissimilar Aluminum Joints. J. of Materi Eng and Perform (2022). https://doi.org/10.1007/s11665-022-07659-7.

[58] Husain Mehdi, R.S. Mishra, An experimental analysis and optimization of process parameters of AA6061 and AA7075 welded joint by TIG+FSP welding using RSM, Advances in Materials and Processing Technologies, 8(1), 598-620, 2022. https://doi.org/10.1080/2374068X.2020.1829952.

[59] Husain Mehdi, R.S. Mishra, Effect of multi-pass friction stir processing and SiC nanoparticles on microstructure and mechanical properties of AA6082-T6, Advances in Industrial and Manufacturing Engineering, 3, 100062 (2021). https://doi.org/10.1016/j.aime.2021.100062

[60] Naman Sharma Husain Mehdi, Shivam Sharma, Mohd Anas, The Influences of Variation of Copper Content on the Mechanical Properties of Aluminium Alloy, International Journal of Material Science Innovations, 3(3), 74-86, 2015.

Cite this article as: Sumit Jain, R.S. Mishra, a review of dissimilar aluminum alloys welding via Friction stir welding, International Journal of Research in Engineering and Innovation Vol-7, Issue-1 (2023), 6-14. *https://doi.org/10.36037/IJREI.2022.7102*.

View publication stats

Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 218 (2023) 2191–2200

Procedia Computer Science

www.elsevier.com/locate/procedia

International Conference on Machine Learning and Data Engineering

A Survey on Quantum Computing for Internet of Things Security

Diksha Chawla[1] ,Pawan Singh Mehra[2]

[1]*Delhi Technological University, Rohini, Delhi -110042, India*

[2]*Delhi Technological University, Rohini, Delhi -110042, India*

Abstract

Quantum computing, based on quantum mechanical principle, can potentially provide significant advantages over classical computing. This advantage of quantum computing provides solutions to many previously unsolvable problems in secure communication and finance. IoT is an emerging technology that deals with a large amount of data. The data communicated in IoT need to be secured. The existing security architecture of IoT is based on cryptographic algorithms such as RSA and ECC. Quantum computing had a significant impact on the security of these algorithms. Therefore, our work analyses the security concerns of IoT smart applications and quantum-based solutions. This article provides a survey of quantum computing fundamentals and the impact of quantum computing on IoT security. Thus, this paper aims to provide a wide view of quantum-enabled IoT communication. The main challenges in implementing quantum-enabled communication are also analyzed in our work.

© 2023 The Authors. Published by Elsevier B.V.
This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0)
Peer-review under responsibility of the scientific committee of the International Conference on Machine Learning and Data Engineering

Keywords: Quantum Computing, Quantum Key Distribution (QKD), Internet of Things (IoT), Quantum Entanglement

1. Introduction

The usage of IoT is expanding in many applications such as intelligent environments, cities, smart grids etc. The number of devices connected in IoT communication provides decision-making ability to users. This vast spectrum of IoT-enabled applications [1], as represented in Figure 1, transfers a massive amount of data, which imposes security and privacy challenges. Without authentication and privacy, IoT applications will not be able to reach high demand. It may also create serious security threats to their potential users. IoT has challenges such as privacy, confidentiality and authentication. The existing scenario of IoT applications is based on RSA and ECC-based schemes [1]. However, with the emergence of quantum computers, such encryption primitives will no longer be secure. Quantum computer solves these classically unsolvable problems based on classical cryptographic primitives. Therefore, the security of this IoT communication network is ensured by Quantum computing. Quantum computing is based on the principle of uncertainty[2] and the no-cloning theorem[3]. The main aim of studying quantum computing is to design protocols and algorithms[4] to resolve IoT security issues, which are Quantum resistant. Classical computing manipulates individual bits, whereas quantum computer uses qubits. These qubits, with their associated probability, represent the quantum state. These qubits are based on Quantum mechanics principals such as superposition and entanglement. Superposition allows qubits to be in different possible combinations of values simultaneously. Entanglement creates a strong dependent relation between quantum particles. However, with the advancement of

1877-0509 © 2023 The Authors. Published by Elsevier B.V.
This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0)
Peer-review under responsibility of the scientific committee of the International Conference on Machine Learning and Data Engineering
10.1016/j.procs.2023.01.195

quantum computing, the existing encryption methods are at a significant threat. Quantum key distribution(QKD) [5] is a highly active research area in quantum computing. It enables communicating parties to establish secret keys to communicate securely. Quantum Computing provides many benefits to the futuristic world, such as creating life-saving medicines, advancing artificial intelligence, and creating intelligent infrastructure. The recent advancements in quantum computing impose threats to cyber security algorithms. In the literature, many advantages of quantum computing for securing IoT communication based on the BB84 protocol have already been proposed. Many branches of quantum computing for secure IoT communication, such as QKD[6], Quantum entanglement[7], and Quantum Bit Commitment (QBC) protocols[8], have already been explored.

However, as per the authors' knowledge an in detail, Quantum Computing (QC) analysis for securing the IoT communication model considering different attacks on each layer is still unavailable. The main advantage of quantum computing is utilizing quantum superposition, quantum entanglement, Quantum Signature [9] and QKD to deliver secure data and use smart applications for effective decision-making. Therefore, in our work, we analyzed attacks on IoT layered architecture. We also reviewed the benefits of integrating the Quantum-based layer into the existing IoT layer architecture and its future perspective.

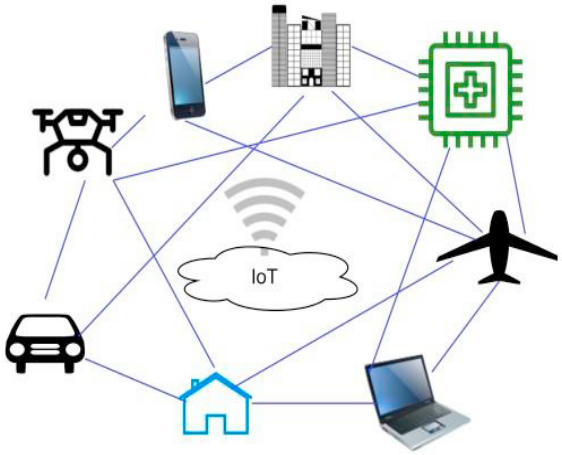


Fig. 1: Futuristic IoT architecture

Our work is organized into sections. Section 2 discusses previous work done on quantum-based communication. Next, in Section 3, IoT security issues and requirements of IoT security are elaborated. The Quantum fundamentals are discussed in Section 4. Quantum key distribution (QKD) is analyzed in section 5. Quantum-based authentication is analyzed in Section 6. In section 7, we discuss quantum-enabled IoT communication. In section 8, we discuss the challenges of IoT communication. Finally, in section 9, we concluded our work. Some of the frequently used abbreviations and symbols are represented in Table 1.

Table 1: List of abbreviations and symbols

Abbrevations/Symbols	Description	Abbrevations/Symbols	Description
IoT	Internet of Things	SARG04	Scarani-Acin-Rigbordy-Gisin 2004
RSA	Ron Rivest, Adi Shamir and Leonard Adleman	E91	Artur Ekert (1991)
ECC	Elliptic Curve Cryptography	Qubit	Quantum Bit
COW	Coherent One-Way	|0>	Ket Zero bit
QKD	Quantum Key Distribution	|1>	Ket One bit
$\emptyset^{\pm}$	Similar Entanglement	WSN	Wireless Sensor Network
$\psi^{\pm}$	Different Entanglement	BBM92	Bennett, Brassard and Mermin

1.1. Contributions

The main contributions of our work are as follows:

1. After understanding the IoT security concerns, a survey on quantum computing for Internet of Things (IoT) smart applications security is analyzed.
2. The main quantum-computing fundamentals and their importance for the security framework are discussed.
3. Thorough analyses of Quantum key security for secure IoT communication are discussed.
4. The main types of Quantum authentication methods and quantum-enabled IoT communication is analyzed.
5. The main challenges for quantum computing-based IoT cryptosystem are provided.

2. Literature survey

In IoT devices, as more devices are connected, enabling security to establish trust among users will become a challenge. Therefore, Hafizul et al.[10]proposed authentication protocol for IoT-enabled Wireless Sensor Networks (WSN). Their work overcomes security threats such as user and sensor node impersonation. IoT devices such as smartphones and wearable are full of sensors and embedded software. All these devices communicated through the internet. Farash et al.[11] proposed authentication protocol where users get data if they pass the authentication and key agreement scheme. Cryptography is used by Mitchell et al.[12] in their work discussed the impact of quantum computers on existing cryptographic primitives. According to Gill et al.[3]quantum advantage can be utilized in different domains such as medicine, cybersecurity, IoT, weather, and national laboratories to solve complex issues such as autonomous cars. Ralegankar et al.[13]work analyse as more data generated and transferred using UAV communication are vulnerable to different attacks such as data modification attacks and snooping attacks. To secure such communication, they proposed the BB84 protocol. Hassija et al.[4] analyses quantum applications for future communication. They also discussed various quantum algorithms such as cryptography, unstructured search, Quantum annealing, and amplitude estimation. Dorothy et al.[14] in their work proposed modified RSA protocol. They utilized BB84 protocols for a secure IoT framework. Rahman et al.[13] introduced quantum layer for adding

security in classical IoT layer architecture. In their work, Gupta et al.[14] analyzed QKD and utilized COW protocol[15] to secure IoT communication. Cheng et al.[16]. Their work discusses the impact of quantum-based encryption over classical encryption methods and the requirement of quantum-resistant encryption methods. Yen et al.[17]in their work discuss the role of entanglement in mutual authentication. In their work, EL-Latif et al. [18]addressed the data protection and information sharing issues in network communications. They proposed the Quantum One-Way Hash Function (QOWHF) to resolve current and future communication frameworks. Most of the work on Quantum computing for securing IoT-enabled communication has been done in literature by many existing authors, as represented in Table: 2

Table 2: Existing work done by authors on quantum-enabled IoT security

Author	Description	Advantages	Disadvantages
Bhatt et al.[1]	Quantum Cryptography for IoT is utilized to overcome drawbacks of classical cryptography.	Resistant to security breaches, side channel attack and data authentication.	Implementation challenges are not discussed.
Routray et al.[2]	Quantum Cryptography for IoT to resolve security issues.	QKD error correction phase	Security concerns are not discussed
Althobaiti et al.[3]	Lattice based quantum computing	Quantum resistant security framework	The proposed model is based on mathematical framework.
Sharbaf et al.[4]	Quantum Key Distribution (QKD) to overcome classical security key issues is analyzed.	BB84 is used to provide Unconditional security.	Security threats are not analyzed.
Lohachab et al.[5]	Quantum Key distribution and ECC for securing IoT Communication	Secure Key Distribution	Impersonation and Man-In Middle attack.
Cheng et al.[6]	Recent developments in Quantum Computing.	Post-Quantum Cryptosystems are analyzed for quantum resistant solutions.	Security threats are not analyzed in their work.
Gyongyosi et al.[7]	In detail analysis of Quantum Computing including quantum gates, quantum memories, Quantum error correction and quantum fundamentals are discussed.	Large-scale quantum computing based implementation is analysed.	Impact of Quantum computing on IoT security is not analyzed.
Angara et al.[8]	Quantum Computing in detail analysis for education purpose is analyzed.	Quantum Computing implementation, theory and hardware is discussed.	Importance of quantum computing for students is discussed.
Mitchell et al.[9]	The impact of quantum computing on real world security is analyzed.	Quantum Computing for 5G mobile communication is analyzed.	Security concerns are not analyzed.
Our Survey	An in detail analysis of Quantum Computing impact on classical cryptographically secured primitives utilized for securing IoT communication.	Analysis of IoT communication layer architecture and security concerns.	-

3. IoT Security Issues

Innovative applications such as intelligent cities, smart agriculture, and innovative health care enabled the

advancement of IoT-enabled communication. The IoT devices that support these applications transmit vast amounts of data in many different environments. The advancement in IoT applications increases cyber-attacks. It also poses threats to user privacy and confidentiality. The main security challenges concerning the IoT environment are authentication, integrity, authorization and trust management. Major security concerns encountered in IoT layered architecture[18], as represented in Figure 2. This section review threats to IoT layer architecture and the benefits of introducing the Quantum layer for IoT security.

3.1. Sensing Layer

Various sensing technologies enable IoT applications such as WSN, RFID, and GPS in this layer. All these technologies deal with IoT sensors and actuators. Sensors are used for perceiving data from surroundings, such as ultrasonic, camera, and temperature detection. Many attacks on this sensing layer are possible such as sensor node capturing, false data code injection, eavesdropping, and sleep deprivation attacks.

3.2. Network Layer

Computational units must process the data received from the lower (sensor) layer. The function of the network layer is to send the information acquired from the sensor layer to processing units. The processed data is required to enable IoT applications. However, due to open internet connectivity, network layers face serious security threats such as access control attacks, Denial of service attacks, attacks during data transient, etc.

3.3. Quantum Layer

The Quantum layer provides security to IoT applications. It includes secure key distribution. Due to quantum mechanics laws, the privacy and security of keys are guaranteed at this layer. However, this layer enabled quantum-based cryptography, which will suffer from security threats such as individual, collective and coherent attacks.

3.4. Application Layer

The application layer is accountable for providing services to the user for decision-making. The critical IoT applications are smart cities, intelligent environments, and competent health care and intelligent grids. IoT heterogeneous applications have severe issues of privacy, confidentiality and authentication of data. Eavesdropping attacks, access control, service interruption attack and malicious code attack are the major issues at the application layer.

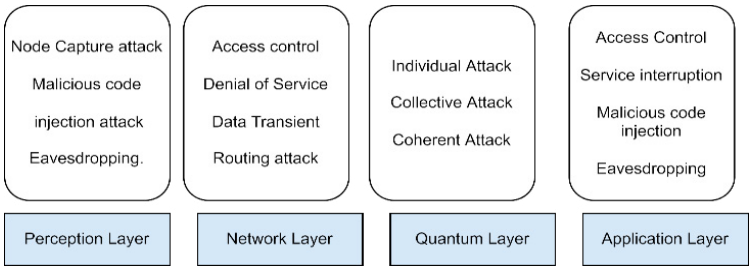


Fig. 2: Security threats on IoT layer architecture.

4. Quantum Fundamentals

The basic building block of Quantum computing is shown in figure 3, consisting of Quantum physical building blocks, Quantum Logic gates and a Quantum programming environment.

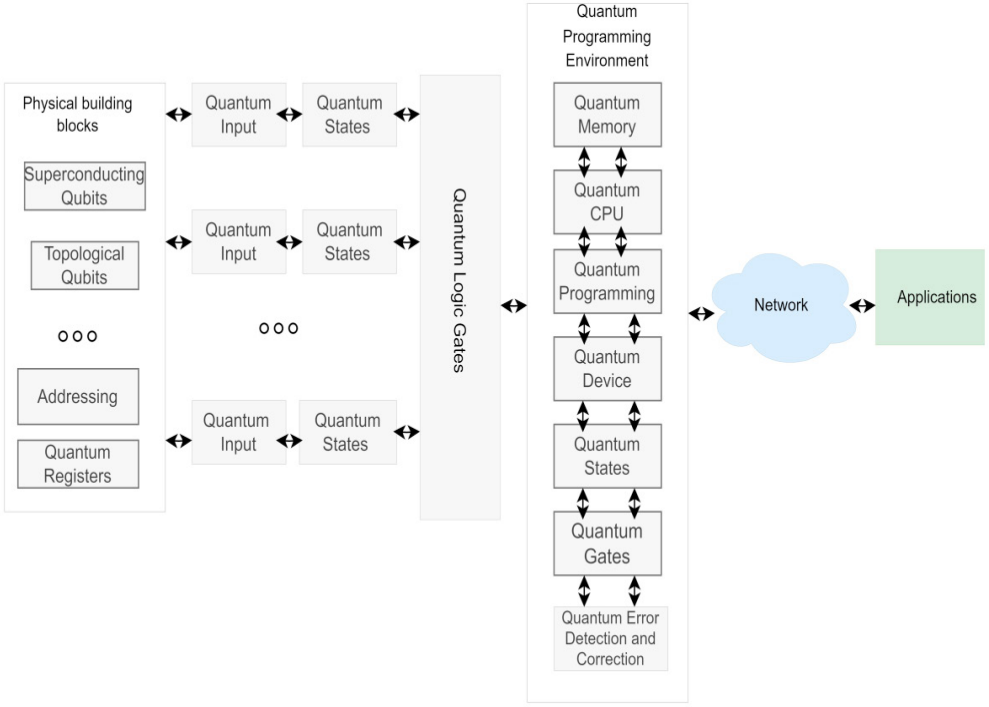


Fig.3: Architecture of Quantum Computing

4.1. From Classical bits to Quantum Bits

In classical computing, we use bits to process information; however, we use Quantum bits known as qubits in quantum computing. These qubits behave according to the laws of quantum mechanics. The quantum particles in spin up and downstate are written as$|0>$ and $|1>$. The group of n qubits store up to 2^n valid values with individual probability measuring each value.

4.2. Quantum Superposition

The qubits are in the linear combination of both $|0>$ and $|1>$, known as quantum superposition. A system can be in superposition[19] of all possible states simultaneously without being confined to one particular state. However, measurement collapses the superposition. It makes qubits to be measured in one of the qubit values$|0>$ or $|1>$.

5. Quantum Key Distribution

Quantum computers bring a huge risk to classical secret key structures. The initial target of quantum computers is asymmetric key structures. Therefore, quantum-based keys need to be utilized for a secure IoT communication framework. The foundation of the quantum key depends on quantum mechanics [16]. The method is based on sharing a secure key between communicating entities. The BB84[14], BBM92[20], SARG04[6] and E91[20] are secure quantum key distribution protocols. The QKD aims to securely distribute and transmit keys using quantum and classical channels. The concept of key distribution is based on sharing random bits using a random basis, which is sent to the receiver. The laws of quantum mechanics reduce the amount of information extracted from the quantum system.

6. Quantum Authentication

The identity authentication for verifying the integrity and confidentiality of messages ensures the security of the message. Mutual authentication between two entities ensures non-repudiation, integrity, and verification of messages. Quantum-based identity authentication can be achieved by using Quantum Signature and Entanglement.

6.1. Quantum Entanglement

The counterintuitive feature of quantum communication is quantum entanglement. When two particles are entangled, their states have to be defined concerning each other. The entangled states are also referred to as Bell states [19]. Four mutually orthonormal entangled states are expressed as in equations (1) and (2)[21]:

$$\emptyset^{\pm} = \frac{1}{\sqrt{2}}(|00> \pm |11>) \quad (1)$$

$$\psi^{\pm} = \frac{1}{\sqrt{2}}(|01> \pm |10>) \quad (2)$$

6.2. Quantum Signature

Quantum signatures are used to verify the authenticated user. Different kinds of Quantum signature protocols exist, such as Arbitrated quantum signature (AQS), Quantum blind signature (QBS) and quantum group signature[22]. A Quantum signature based on the Quantum walk has also been proposed. Quantum signatures are useful for applications such as Quantum Cheque[23], which ensure Non-repudiation and user impersonation attack.

7. Quantum-enabled Internet of Things (IoT)

IoT framework is an interconnection of heterogeneous devices interconnected with diverse technologies such as Wifi, Bluetooth, Zigbee, Bluetooth, and 6LOWPAN. These are IoT enabling technologies. These technologies enable data transfer in IoT applications such as smart cities, innovative medical infrastructure and intelligent farming. These applications require data privacy and confidentiality; therefore, IoT integration with quantum computing plays a significant role. The IoT communication infrastructure is secured by classical cryptography, such as Public and Private-key structures[24]are soon under attack by quantum computers. Quantum base shor's[25] and Grover's [26] algorithm already pose a threat theoretically to Public-key based infrastructure.

Security of sensitive information in IoT communication is required since the devices involved in IoT communication are resource-constrained and secure, lightweight cryptography is essential. In addition to that, the need to handle classical and quantum attacks opens the door toward quantum-resistant cryptography[27]. In future, IoT communication Quantum computing will become essential for secure communication. Figure: 4 show the comparison of existing quantum-based schemes analyzed by researchers.

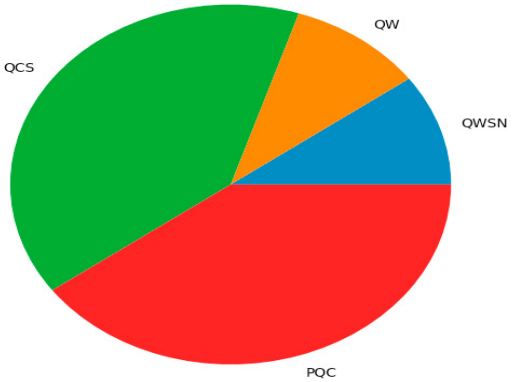


Fig. 4: Comparative analysis of existing quantum-based schemes [Quantum Computing Schemes (QCS), Quantum Walk (QW), Post Quantum Cryptography (PQC), Quantum wireless sensor Networks (QWSN)]

8. Challenges of Quantum based IoT

8.1. Quantum key distribution

It provides a secure way of communicating keys. However, suppose the presence of an eavesdropper was detected on the quantum channel. In that case, the entire process is discarded, and communication will not start again until no eavesdropping is found on the channel.

8.2. Short Distance Communication

The communication of mass users of IoT communication using the quantum channel is challenging due to the short-distance communication of QKD.

8.3. Quantum Reversible Computing

The quantum-based reversible computing[28] also poses a significant threat in the presence of eve.

8.4. Security Attack

The individual attack on quantum-based communication is also possible. In this, the attacker creates a new quantum channel by intercepting a quantum signal transmitted between Alice and Bob.

9. Conclusion

IoT is the interconnection of many heterogeneous devices as all devices are interconnected and provide many benefits to users to facilitate decision-making. Such technology, which contains our daily critical information in health care, intelligent cities and even military applications, must be secure. There are many classical cryptographic primitives, which provide secure communication based on complex mathematical structures. The security based on classical cryptographic structures is no longer secure due to quantum computing attacks. Therefore, IoT-enabled communication requires quantum-based security to resist futuristic quantum attacks. In our survey, we analyzed quantum-based cryptographically secured structures for IoT communication. This article provides a thorough survey of security attacks on IoT-based applications, quantum-resistant solutions for securing IoT communication, the

quantum authentication methods, QKD and challenges in implementing quantum-enabled IoT communication. As a result of such a contribution, this article provides useful guidelines for future IoT researchers to consider more quantum-resistant solutions.

References

[1] M.A. Ferrag, L.A. Maglaras, H. Janicke, J. Jiang, L. Shu, Authentication Protocols for Internet of Things: A Comprehensive Survey, Secur. Commun. Networks. 2017 (2017). https://doi.org/10.1155/2017/6562953.

[2] A. Lohachab, Using Quantum Key Distribution and ECC for Secure Inter-Device Authentication and Communication in IoT Infrastructure, SSRN Electron. J. (2018). https://doi.org/10.2139/ssrn.3166511.

[3] S.S. Gill, A. Kumar, H. Singh, M. Singh, K. Kaur, M. Usman, R. Buyya, Quantum computing: A taxonomy, systematic review and future directions, Softw. - Pract. Exp. 52 (2022) 66–114. https://doi.org/10.1002/spe.3039.

[4] V. Hassija, V. Chamola, A. Goyal, S.S. Kanhere, N. Guizani, Forthcoming applications of quantum computing : peeking into the future, 1 (2020) 35–41. https://doi.org/10.1049/iet-qtc.2020.0026.

[5] S. Krithika, T. Kesavmurthy, Securing IOT network through quantum key distribution, Int. J. Innov. Technol. Explor. Eng. 8 (2019) 693–696. https://doi.org/10.35940/ijitee.F1141.0486S419.

[6] J. Ahmed, A.K. Garg, M. Singh, S. Bansal, M. Amir, Quantum Cryptography Implementation in Wireless Networks, 3 (2014) 129–133.

[7] S.T. Cheng, C.Y. Wang, M.H. Tao, Quantum communication for wireless wide-area networks, IEEE J. Sel. Areas Commun. 23 (2005) 1424–1432. https://doi.org/10.1109/JSAC.2005.851157.

[8] A. Broadbent, C. Schaffner, Quantum cryptography beyond quantum key distribution, Springer US, 2016. https://doi.org/10.1007/s10623-015-0157-4.

[9] S.R. Moulick, P.K. Panigrahi, Quantum cheques, Quantum Inf. Process. 15 (2016) 2475–2486. https://doi.org/10.1007/s11128-016-1273-4.

[10] SK. Hafizul, I. Prosanta, Efficient authentication protocol for secure multimedia communications in IoT-enabled wireless sensor networks, (2017).

[11] M.S. Farash, M. Turkanović, S. Kumari, M. Hölbl, An efficient user authentication and key agreement scheme for heterogeneous wireless sensor network tailored for the Internet of Things environment, Ad Hoc Networks. 36 (2016) 152–176. https://doi.org/10.1016/j.adhoc.2015.05.014.

[12] C.J. Mitchell, The impact of quantum computing on real-world security: A 5G case study, Comput. Secur. 93 (2020) 1–11. https://doi.org/10.1016/j.cose.2020.101825.

[13] V.K. Ralegankar, J. Bagul, B. Thakkar, R. Gupta, S. Tanwar, G. Sharma, I.E. Davidson, Quantum Cryptography-as-a-Service for Secure UAV Communication: Applications, Challenges, and Case Study, IEEE Access. (2021) 1–1. https://doi.org/10.1109/access.2021.3138753.

[14] A.B. Dorothy, S.B.R. Kumar, An approach for IoT security using quantum key distribution, Int. J. Sci. Technol. Res. 8 (2019) 1569–1574.

[15] S. Gupta, C. Dutta, Internet of Things Security Analysis of Networks using Quantum Key Distribution, Indian J. Sci. Technol. 9 (2016). https://doi.org/10.17485/ijst/2016/v9i48/105551.

[16] C. Cheng, R. Lu, A. Petzoldt, T. Takagi, Securing the Internet of Things in a Quantum World, IEEE Commun. Mag. 55 (2017) 116–120. https://doi.org/10.1109/MCOM.2017.1600522CM.

[17] C.A. Yen, S.J. Horng, H.S. Goan, T.W. Kao, Y.H. Chou, Quantum direct communication with mutual authentication, Quantum Inf. Comput. 9 (2009) 376–394. https://doi.org/10.26421/QIC9.5-6-2.

[18] M.S. Rahman, M. Hossam-E-Haider, Quantum IoT: A quantum approach in IoT security maintenance, 1st Int. Conf. Robot. Electr. Signal Process. Tech. ICREST 2019. (2019) 269–272. https://doi.org/10.1109/ICREST.2019.8644342.

[19] D.J. Egger, C. Gambella, J. Marecek, S. McFaddin, M. Mevissen, R. Raymond, A. Simonetto, S. Woerner, E. Yndurain, Quantum Computing for Finance: State-of-the-Art and Future Prospects, IEEE Trans. Quantum Eng. 1 (2021) 1–24. https://doi.org/10.1109/tqe.2020.3030314.

[20] N. Papanikolaou, An introduction to quantum cryptography, XRDS Crossroads, ACM Mag. Students. 11 (2005) 3–3. https://doi.org/10.1145/1144396.1144399.

[21] M.S. Kang, C.H. Hong, J. Heo, J.I. Lim, H.J. Yang, Controlled mutual quantum entity authentication using entanglement swapping, Chinese Phys. B. 24 (2015). https://doi.org/10.1088/1674-1056/24/9/090306.
[22] L. Jian, L. Na, Z. Yu, W. Shuang, D. Wei, C. Wei, M. Wenping, A survey on quantum cryptography, Chinese J. Electron. 27 (2018) 223–228. https://doi.org/10.1049/cje.2018.01.017.
[23] BK. Behera, A. Banerjee, P.K. Panigrahi, Experimental realization of quantum cheque using a five-qubit quantum computer, Quantum Inf. Process. 16 (2017) 1–12. https://doi.org/10.1007/s11128-017-1762-0.
[24] V. Teja, P. Banerjee, N.N. Sharma, R.K. Mittal, Quantum cryptography: State-of-art, challenges and future perspectives, 2007 7th IEEE Int. Conf. Nanotechnol. - IEEE-NANO 2007, Proc. (2007) 1296–1301. https://doi.org/10.1109/NANO.2007.4601420.
[25] A.I. Nurhadi, N.R. Syambas, Quantum Key Distribution (QKD) Protocols: A Survey, Proceeding 2018 4th Int. Conf. Wirel. Telemat. ICWT 2018. (2018) 1–5. https://doi.org/10.1109/ICWT.2018.8527822.
[26] R. Niederhagen, S. Css, P. Michael, W. Fraunhofer, D. Fraunhofer, S. Information, T. Sit, Practical Post-Quantum Cryptography White Paper Practical Post-Quantum Cryptography, (n.d.).
[27] SK. Routray, M.K. Jha, L. Sharma, R. Nyamangoudar, A. Javali, S. Sarkar, Quantum cryptography for IoT: APerspective, IEEE Int. Conf. IoT Its Appl. ICIOT 2017. (2017). https://doi.org/10.1109/ICIOTA.2017.8073638.
[28] AG Aruna, KH Vani, C. Sathya, RS Meena, A Study on Reversible Logic Gates of Quantum Computing, 7 (2016) 427–432.

Multimedia Systems
https://doi.org/10.1007/s00530-023-01060-7

REGULAR PAPER

A two stream face anti-spoofing framework using multi-level deep features and ELBP features

Aashania Antil[1] · **Chhavi Dhiman**[1]

Received: 29 November 2022 / Accepted: 5 February 2023
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2023

Abstract
The recent boom in publicly available face authentication (FA) system has brought forward the susceptibility of FA systems to different attack vectors, especially spoofing/presentation attacks. There has been a noticeable increase in the novelty, variety and complexity of spoofing attempts on FA systems. Recently, numerous strategies have been proposed, employing both traditional as well as deep learning approaches. However, generalized face spoofing detection has always proven itself as a challenging research hotspot. Here, we propose a multi-level ELBP texture and deep features based novel face anti-spoofing framework for face spoofing detection. Extensive experiments including intra-dataset and inter-dataset testing were performed on three challenging face anti-spoofing databases, namely Replay-Attack, CASIA FASD, and MSU-MFSD to validate the effectiveness of our proposed work. A comprehensive ablation study to analyze the effects of different color spaces and multi-level deep feature extraction is also discussed. The proposed approach is effective at resisting photo and video attacks. The same is confirmed by the experimental results, outperforming the other state-of-arts with competitive results for both intra and inter-dataset evaluation protocols. The code and pre-trained model weights are available at https://github.com/aashania/FAS_Framework_multilevel_ELBP.

Keywords Face authentication · Presentation attacks · Face anti-spoofing · ELBP · Xception

1 Introduction

Face authentication and verification technologies have been deployed in wide array of applications, from access control at borders to face unlocking of smart phones and other daily use devices [1]. High accuracy, fast operation and the general convenience associated with these systems has led to face biometrics' wide spread user adoption. The extensive use of this technology has exhibited vulnerabilities and susceptibility to various forms of attacks such as face spoofing, presentation attacks [2, 3] and face manipulation [4]. The ubiquity of social networks has made it trivial for an attacker to obtain images and videos for face spoofing. This has resulted in bringing the attention of research community toward vulnerability of face authentication-based system. Most proposed countermeasures work adequately in controlled environments. The development of a general anti-spoofing framework which can adapt to device/sensor quality and environment variations has been a challenging task. This is evidenced by the reported drop in performance of state-of-art [5]when exposed to real-world variables such as illumination and background variations, etc. Thus, increasing the effectiveness of face spoofing detection in FR is a crucial task. Therefore, face spoofing detection methods for photo [6] and video attacks [7] have evolved in recent decades.

In the early days, global approaches [8, 9] were the main focus of researchers, in which the entire image is considered as a single vector and the dimensionality of feature vectors was reduced by project into a low dimension sub space with suitable methods e.g., Principal Component Analysis (PCA) [10] and Linear Discriminant Analysis (LDA) [11]. However, when subjected to unrestricted circumstances such as position variations, lighting changes, expression changes, occlusion effects, and aging the works failed to exhibit adequately. Lately, the researchers have been exploring transformation functions such as Patterns of Oriented Edge Magnitudes (POEM) [12], Histogram of Gabor Phase Patterns (HGPP) [13], Local Binary Pattern (LBP) [14], and Local

✉ Chhavi Dhiman
chhavi1990delhi@gmail.com

Aashania Antil
aashiantil40@gmail.com

[1] Delhi Technological University, Delhi, India

Published online: 24 February 2023

Gabor Binary Pattern (LGBP) [15] to better encode subregions of face image. These approaches assume that local strategies are more resilient to uncontrolled situations than global strategies. Many works such as [16] have also leveraged the combination of texture features and deep learning approaches to increase the performance and robustness. These works employ traditional features in tandem with deep learning-based methods, providing a promising direction for enhancing generalization. Motivated from these approaches, a novel framework has been proposed based on multi-level texture and deep features representation for face anti-spoofing. In the proposed work, a novel multi-level texture feature representation and transfer learning-based deep face feature representation are combinedly utilized to highlight discriminant face features for face spoofing detection. Xception [17] has better performance than other general image classification models [18, 19] on the benchmark ImageNet dataset. Also in our literature review, we observed that the Xception model has not been explored for the use in FAS. Whereas, Xception and its modifications are generating promising results in other fields such as face manipulation [20], face forgery detection [21] and revealing hidden face [22]. To fully utilize the strength of Xception and its modifications in the field of FAS, we proposed modifications in the Xception baseline ensuring not imparting larger model size while better performance is observed.

In brief, the primary contributions of this work are as mentioned below:

1. A two-stream multi-level face anti-spoofing framework is defined. It extracts salient face anti-spoofing feature representations using multi-level ELBP texture features and modified Xception-based deep features.
2. The proposed face anti-spoofing framework is designed considering the fact excessively large number of trainable parameters result overfitting. Therefore, proposed modified Xception network utilizes the concept of squeeze and excitation [23] while extracting multi-level rich deep features, without increasing the no. of parameters. Also, the proposed multi-level ELBP texture features extraction branch ensures the right selection of levels to extract relevant and effective texture features to maintain a trade-off between model performance and parameters.
3. Extensive ablation study is carried out to exhibit the effect of multi-level texture feature representation in suitable color spaces to derive salient details for face anti-spoofing. Also, experimental analysis supports the efficiency of proposed modified Xception network-based squeeze and excitation enabled deep face features for deep faces representation.
4. For performance evaluation of the proposed framework, experiments were conducted under intra and cross-dataset evaluation protocols on three challenging datasets namely, MSU-MFSD, CASIA FASD and Replay-Attack. By combining the global characteristics retrieved by modified Xception and the local features extracted by ELBP, the proposed approach exhibited improved performance over other state-of-arts.

The rest of the paper is structured as follows: an overview of state-of-arts is provided in Sect. 2. The proposed work is discussed in detail in Sect. 3. Experiment evaluation and supporting analysis is illustrated in Sect. 4. Section 5 presents the conclusion and further discussion.

2 Prior works

Face anti-spoofing has been a focus of research community for the past decade. As a result, the advancements in other applications of computer vision have also found their way into the field of FAS and presentation attack detection. There has been a continuous improvement in both texture and deep learning-based methods. Pivotal works based on these two categories are briefly discussed below:

2.1 Texture features

The texture analysis method for detecting face spoofing has gained popularity, and researchers identify spoofing by evaluating textural details as a discriminative characteristic for face presentation attack detection (PAD). For face spoofing detection, Gragnaniello et al. [24] suggested a co-occurrence of adjacent local binary patterns (CoALBP) based facial texture spoofing detection method. Likewise, Boulkenafet et al. [25] derived a uniform LBP from YCbCr and HSV color spaces and merged them for actual vs. false classification. This approach performs well in intra-dataset conditions but poorly in cross-dataset testing. Following that, Boulkenafet et al. [26] used SURF and Fisher Vector encoding to characterize facial texture. Even with limited training data, the methodology beats most known algorithms and shows very promising generalization capabilities. Inspired from one of the works [25] and CoALBP, a chromatic co-occurrence of local binary pattern was proposed by Peng et al. [27] and investigated the inter-channel information. The combination of CCoLBP feature and intra-channel facial texture achieved good performance. According to the above studies presented, Local binary pattern is frequently utilized in analysis of face texture and has proven to be effective in face spoofing detection [28–30]. However, LBP in its standard form and its simple derivatives take only the difference direction and disregard the magnitude of disparity between the neighbor pixels and center pixel. This issue has been noticed by various scholars and consequently, in recent years

and proposed some modifications, such as completed LBP (CLBP) [31], LBP variance (LBPV) [32], and local ternary patterns (LTP) [33], which have been shown to be efficient in texture classification.

2.2 Deep learning features

In contrast to the traditional handcrafted features, deep features have been frequently used for spoofing detection due to their excellent performance in extracting features. These are defined as automatic learnable feature-based spoof detection algorithms. To leverage this, several researchers treated FAS/PAD problems as binary classification problems and used pre-trained deep models as feature extractor. These models learn mapping functions from input face images to spoof detection. For instance, Lucena et al. [9] and Chen et al. [34] guided the pre-trained VGG16 and ResNet18 models face spoofing detection. Besides binary cross-entropy loss, many works [35–37] also exploit triplet loss, binary focal loss etc., to train the models. Later some literature [38–41] put forward the usage of auxiliary tasks (i.e., depth supervision, binary mask, reflection map, etc.) to provide more fine-grained clues. Liu et al. [38] developed a CNN-based technique for PAD that employs auxiliary supervision. They used a remote photoplethysmography (rPPG) signal for sequence-wise supervision and computed the depth map using CNN-RNN model with pixel-wise supervision. On other hand, Wang et al. [42] estimate face depth information using multiple RGB frames. Li et al. [43] presented a 3D CNN design which increases cross domain generalization by minimizing the MMD (Maximum Mean Discrepancy) distance. Yu et al. [44]developed the CDCN architecture by substituting central difference convolution (CDC) for vanilla convolution in Depth Net to further increase the capability of fine-grained intrinsic feature representation. Furthermore, neural architecture search (NAS) [45] is used to find CDCN++that boosted the performance when combined with multiscale attention fusion module. Besides this, Yu et al. [46] reframed the FAS problem as a material recognition challenge, and combined it with human material perception. Qin et al. [47]suggested a novel MT-FAS method to provide better supervision toward PA detectors.

On the basis of analysis of the methods discussed above, this work aims to design a novel architecture that fuses color texture information with deep learning-based features. The same has been discussed in the following section.

3 Proposed framework

The proposed framework has two branches: Modified Xception and Multi-level ELBP. Figure 1 shows a block diagram of the novel framework that fuses these both branches. The first branch consists of a pre-trained Xception network architecture modified using Squeeze- Excitation [23] module and is given raw RGB images as input data. Other branch first crops the face region from input images and then converts this region of interest (ROI) patch of faces to different color spaces representations- YCbCr and HSV. Multi-level ELBP descriptor is applied on each channel separately to extract more elaborate texture features from respective color spaces. The extracted features are normalized and pass it to a network of fully connected layers. This network fuses the features obtained from both branches and uses them to distinguish between real and spoof. The size of fully connected layers and the color channels used for texture extraction have been chosen after a rigorous experimental analysis. Different stages of the proposed framework are discussed below:

3.1 Pre-processing step

The datasets consist of videos of spoofs and actual persons. In preprocessing, the videos are sampled at every 10th frame, the frames are then organized and labeled as per the

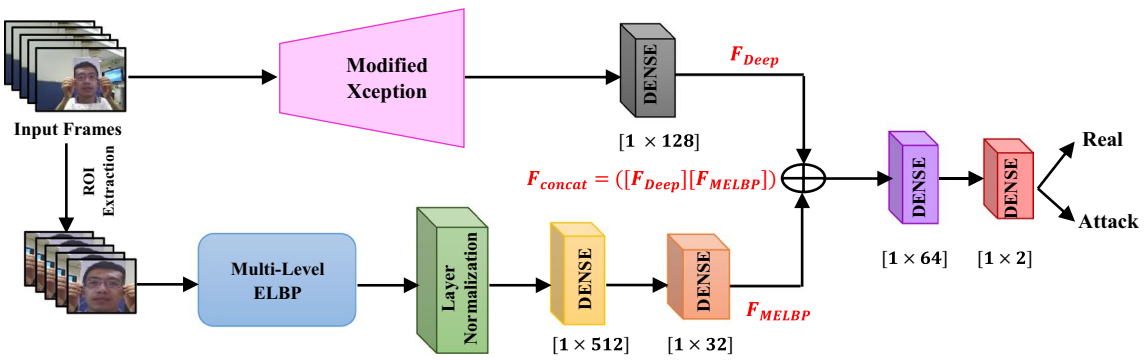


Fig. 1 The proposed two stream face anti-spoofing framework using modified Xception and multi-level ELBP

methodology in corresponding dataset. We resize the frames to [299 × 299] pixels and feed them to our modified Xception Network. Face detection and ROI cropping is done via viola-jones algorithm [48]. However, in 4.3% of sampled frames from Replay Attack, 3.7% from CASIA and 7% from MSU-MFSD, the voila-jones algorithm was unable to detect any face or gave wrong results, hence for such cases, MTCNN [49] is adopted as a fallback for face extraction. The detected face along with some background pixels are extracted and resized to [150 × 150] pixels. Further, we randomly flip the input data (both raw and ROI data) horizontally.

3.2 Color space

For digital applications, RGB is commonly used color space for representing, sensing, and displaying images. Despite the widespread use in image and video codecs, RGB color space alone is not enough to represent all details in an image. According to experts in recent works [29, 50], complementary channels found in the color model hold more fraud details as compared to gray scale texture. These details show up as print defects, noise and moiré effects. Illumination plays a vital role in highlighting imperfection in any surface. Therefore, this research in addition to individual RGB channels, has also adopted the faces in HSV and YCbCr color space for better inclusion of luminance and chrominance information from frames.

3.3 Feature extraction

Texture features and deep features are extracted in two branches independently which are discussed below:

3.3.1 ELBP-based texture features

Zhao et al. [51], remark and signalize eyes and mouth as the preeminent regions of the human face. Human eyes and mouth resemble an elliptical shape. Thus, the Elliptical local binary pattern (ELBP) descriptor is better suited for extraction of micro and gradient facial features from the face image over simple LBP.

In ELBP, for each pixel (X_p, Y_p), neighboring pixels lying on an ellipse with semi major.

axes R_1 and R_2 and (X_p, Y_p) as the center are considered. ELBP of (X_p, Y_p) with N neighbor pixels at distance(R_1, R_2) is determined as:

$$ELBP^{N,R1,R2}(X_p, Y_p) = \sum_{i=1}^{N} F(g_i^{N,R1,R2} - g_p)2^{i-1}. \quad (1)$$

$F(x)$ is defined as:

$$F(x) = \begin{cases} 0, if\ x < 0 \\ 1, if\ x \geq 0 \end{cases}. \quad (2)$$

The i^{th} neighbor pixel coordinates, (X_p, Y_p) are calculated using angular difference along the ellipse as:

$$\Delta\Theta = 2 \times \frac{\pi}{N}, \quad (3)$$

$$x_i = x_p + R1 \times \cos((i-1) \times \Delta\Theta)), \quad (4)$$

$$y_i = y_p - R2 \times \sin((i-1) \times \Delta\Theta)). \quad (5)$$

For generating the ELBP feature vector of the incoming data, first, we use the ELBP operator for generating respective ELBP images. In this process, the elliptical LBP image is first split into non-overlapping sub-regions. Then, the sequence of histograms for each sub-region is calculated individually and concatenated with other sub region histograms for creating the complete ELBP feature vector. If the ELBP operator is denoted as φ then,

$$\varphi = \begin{cases} LBP, & R1 = R2 \\ \varphi_v, & R1 < R2 \\ \varphi_h, & R1 > R2 \end{cases}. \quad (6)$$

Here, φ_h and φ_v represents horizontal and vertical ELBP respectively and values of $(R1, R2)$ in the ELBP operator indicate the semi major and semi minor axes respectively.

As shown in Fig. 2, we have used both ELBP (φ) operators i.e., $\varphi_v = ELBP_v, = ELBP^{8,R1,R2}$ and $\varphi_h = ELBP_h = ELBP^{8,R2,R1}$ to build two ELBP images for an input image (see Fig. 3). Motivated from [52], we choose (5, 3) as $(R1, R2)$ for the horizontal ELBP operator and (3, 5) as $(R1, R2)$ for the vertical ELBP operator to get better performance. Through this, we computed the matching ELBP feature vector for each ELBP image. Finally, the generated horizontal and vertical feature vectors are concatenated for a complete feature vector for the given image.

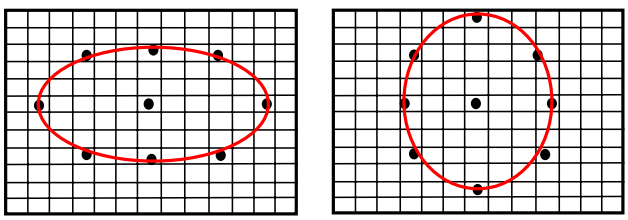


Fig. 2 Horizontal and vertical ELBP Pattern

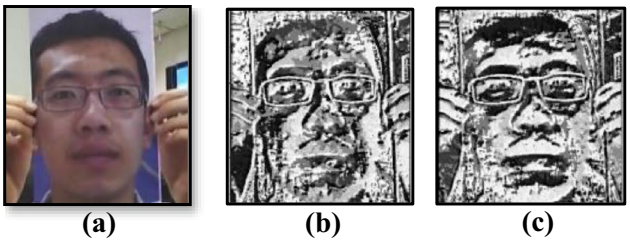


Fig. 3 **a** An original image, **b** $ELBP^{8,5,3}$, **c** $ELBP^{8,3,5}$

To create the feature vector, the frame is segmented into $W \times H$ subregions. In general, the length of ELBP feature for $(8, R_1, R_2)$ neighborhood patterns are $W \times H \times 256$, and the total (vertical and horizontal) feature vector length twice the length $2 \times W \times H \times 256$. If there are no more than two 0–1 or 1–0 bitwise changes LBP's binary representation, that pattern is considered as a uniform pattern [14]. The ELBP feature vector length is decreased by nearly four times when uniform patterns are used (from $W \times H \times 256$ to $W \times H \times 59$ and from $2 \times W \times H \times 256$ to $2 \times W \times H \times 59$). Uniform patterns are used for speeding up the ELBP calculations and reduce the amount of memory needed to store ELBP feature vectors.

For evaluation of ELBP in face anti-spoofing, we also consider fine grid at different levels. We divide every incoming face ROI patch into 4^l sub-regions at each level, where the segmentation level $l \in \{0, 1, 2, 3\}$. For each component channel, ELBP features are extracted from the complete image along with each sub region individually.

For our work, a higher level of decomposition of the image beyond L_2 is not preferred as for L_3, the size of sub region reduces to 1/64th of original image. Also, sub-sections of such small size (18×18) will result in a feature vector that is larger than the image itself. This unwanted increase will act as noise in our data. Moreover, inclusion of L_3 reported a drop in performance with increased computational requirements. Therefore, we consider sub-sectioning only up to level-2 in our method.

Feature histogram for sub-region of size $X \times Y$, the corresponding ELBP texture for pixel (i, j) can be represented as:

$$H(n) = H(n) = \sum_{i=0}^{X} \sum_{j=0}^{Y} \varphi^{N,R_1,R_2}(i,j). \tag{7}$$

The features are extracted for each sub-region and then cascaded, for each channel in color space C, into a vector H, where $I\,(1, 2, k)$ is the channel index, the complete feature vector of face image in color space C is thus defined as:

$$H_c = [H_c^1 H_c^2 .....H_c^k]. \tag{8}$$

ELBP feature extraction process has been explained in algorithm 1 by considering level-0, level-1 and level-2. Figure 4 shows an illustration of extracting of ELBP-based texture Features for Level-0, Level-1 and Level-2.

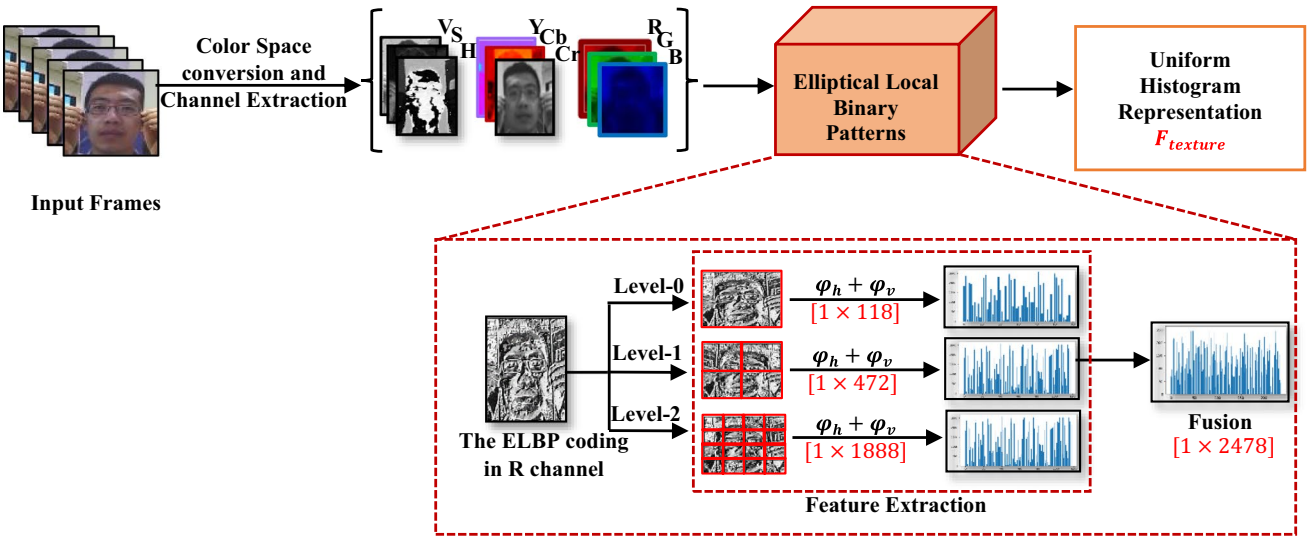


Fig. 4 Multi-Level ELBP-based texture feature extraction

Algorithm 1: Algorithm of multi-level ELBP Feature Vector ($\mathcal{F}$)

Input : Video Frame F, Color Space C, ELBP parameters ($\mathbf{8}, \boldsymbol{R_1}, \boldsymbol{R_2}$)
Output : Feature vector H of Frame F
$\boldsymbol{I}$ ← Face Detection and Cropping ($\boldsymbol{F}$)
$\boldsymbol{I^i}$ ← Color Space Conversion ($\boldsymbol{I}, \boldsymbol{C}$) # $\boldsymbol{N}$ is the number of color channels in $\boldsymbol{C}$
$\boldsymbol{L_0}$ ← 0, $\boldsymbol{L_1}$ ← 1, $\boldsymbol{L_3}$ ←2 # $\boldsymbol{L_0}$,$\boldsymbol{L_1}$ is the first, second and third level of sub-region divisions, respectively
for $\boldsymbol{i}$ ← 1 to N **do**
  $\boldsymbol{I_C^i}$ ← ELBP ($\boldsymbol{I^i}, \mathbf{8}, \boldsymbol{R_1}, \boldsymbol{R_2}$)
  $\boldsymbol{H_C^i}(\boldsymbol{L_0})$ ←uniform histogram ($\boldsymbol{I_C^i}$) #statistic the histogram of $\boldsymbol{I_C^i}$
  for j ← 1 to 4 **do**
    h($\boldsymbol{j}$) ← ELBP ($\boldsymbol{I^i}$ ($\boldsymbol{j}$), $\mathbf{8}, \boldsymbol{R_1}, \boldsymbol{R_2}$) # ELBP of each sub region to histogram
    $\boldsymbol{H_C^i}(\boldsymbol{L_1})$ ← append $h(\boldsymbol{j})$ # append $h(\boldsymbol{j})$ to the end of $\boldsymbol{H_C^i}(\boldsymbol{L_1})$
  end for
  for k ← 1 to 16 **do**
    h($\boldsymbol{k}$) ← ELBP ($\boldsymbol{I^i}$ ($\boldsymbol{k}$), $\mathbf{8}, \boldsymbol{R_1}, \boldsymbol{R_2}$) # ELBP of each sub region to histogram
    $\boldsymbol{H_C^i}(\boldsymbol{L_2})$ ← append $h(\boldsymbol{k})$
  end for
end for
$\boldsymbol{H_C^i}$ = *concatenate* ($\boldsymbol{H_C^i}(\boldsymbol{L_0}), \boldsymbol{H_C^i}(\boldsymbol{L_1})\boldsymbol{H_C^i}(\boldsymbol{L_2})$) # concatenate $\boldsymbol{H_C^i}(\boldsymbol{L_0})$, $\boldsymbol{H_C^i}(\boldsymbol{L_1})\boldsymbol{H_C^i}(\boldsymbol{L_2})$
$\boldsymbol{H_C}$ = append $\boldsymbol{H_C^i}$
end for

3.3.2 Modified Xception-based deep features

In the proposed work, Xception [17] is used as a baseline deep feature extraction network. The basic Xception model consists of entry, middle, and exit flows. It learns channel correlation and spatial correlation separately. Additionally, it is more accurate and has fewer parameters, which fits our demand for anti-spoofing. To increase the robustness of our architecture toward unknown attack types, we have introduced squeeze and excitation [23] modules at different levels of Xception [17] model shown in the Fig. 5a.

Squeeze and excitation block is a lightweight architectural unit that aims to enhance the representational power of a lattice with channel-wise dynamic feature recalibration. Figure 5b shows a schematic representation of the SE block. The 'squeeze' phase consists of a GAP (Global Average Pooling) operation, whereas the 'excitation' phase is composed of two small, fully connected layers followed by a channel-wise scaling operation which is computationally inexpensive.

Consider a transformation $\mathcal{F}_{tr}$ to be a Xception blocks which maps the input $I \epsilon \mathbb{R}^{H' \times W' \times C''}$ to feature maps $M \epsilon \mathbb{R}^{H \times W \times C}$. A SE block, can be built to perform feature calibration on these feature maps. A SE block receives these feature maps input with dimensions $[H \times W \times C]$ and performs feature recalibration on these outputs in two steps:

The features M are first put through a squeeze operation, which aggregates feature maps across their spatial dimensions ($H \times W$) to create a channel descriptor. The squeeze operation can be denoted as:

$$z_C = \mathcal{F}_{sq}(m_C) = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} m_C(i,j). \tag{9}$$

To make use of the information aggregated in the squeeze operation, a second operation 'excitation' is performed to capture channel-wise dependencies. This is achieved with a simple self-gating mechanism using sigmoid activation:

$$s = \mathcal{F}_{ex}(z, W) = \sigma(g(z, W)) = \sigma\big(W_2 \delta\big(W_1 z\big)\big). \tag{10}$$

Here, δ denotes to the ReLU [53] function, $W_1 \in \mathbb{R}^{\frac{C}{r} \times C}$ and $W_2 \in \mathbb{R}^{C \times \frac{C}{r}}$.

The output X of the SE block, is produced by applying these weights to the feature mappings M.

$$X = \mathcal{F}_{\text{scale}}\big(m_C, s_C\big) = s_C m_C. \tag{11}$$

This block can be utilized in any standard architecture by inserting it after the non-linearity operation following each convolutional operation. Consequently, the benefits of feature recalibration can be obtained and offer a good trade-off between improved performance and increased computation cost. We have taken advantage of the Squeeze and Excitation module by integrating it into the pre-trained Xception model at entry, middle and exit flow separately. The feature size extracted from entry and middle level is

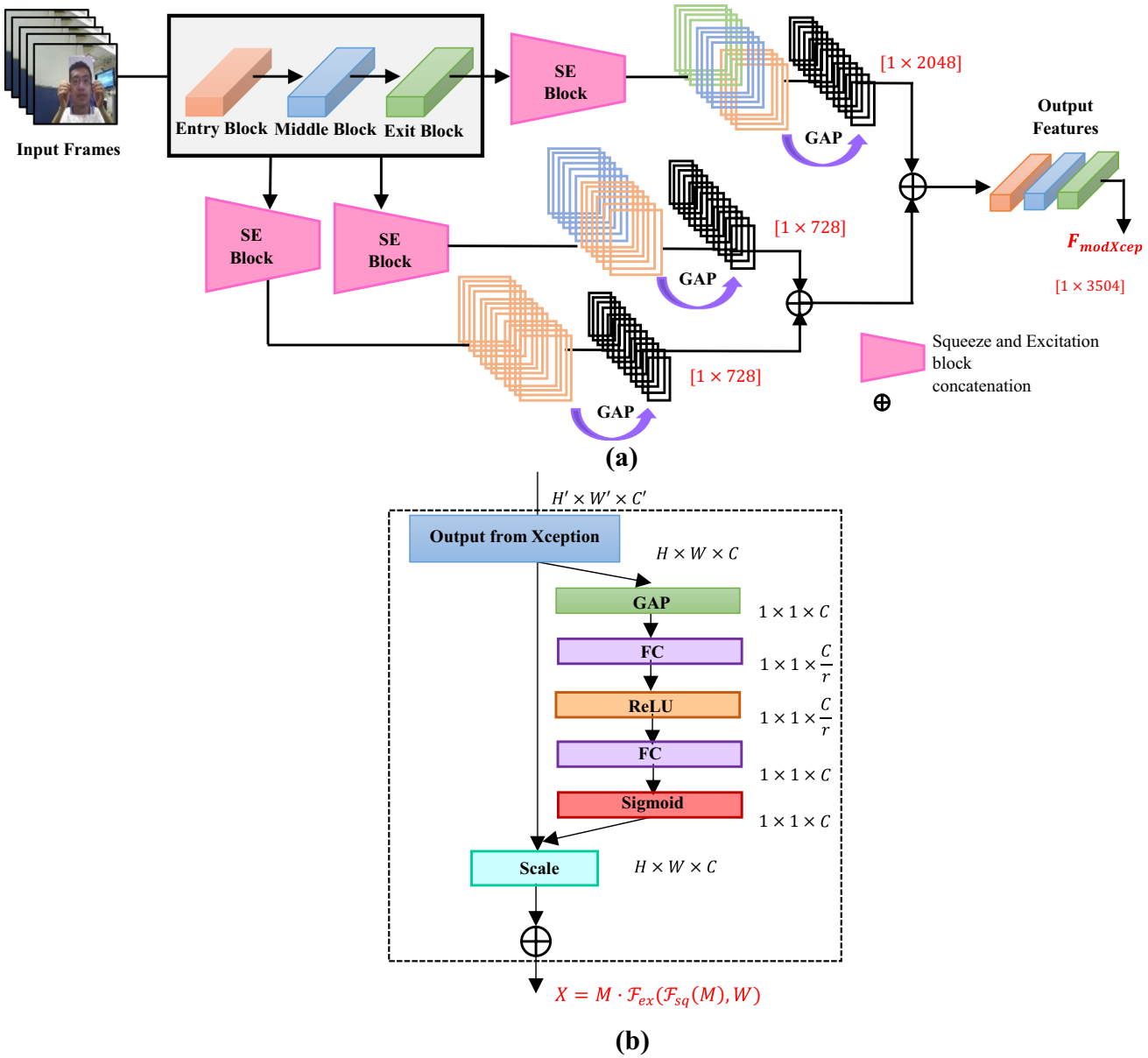


Fig. 5 **a** Modified Xception Network, **b** Squeeze and excitation block [23] applied in Xception [17]

$[19 \times 19 \times 728]$ each whereas the exit flow has a feature output size of $[10 \times 10 \times 2048]$. We apply GAP on the output of 'Squeeze and Excitation' blocks individually and concatenate the reduced features into a single vector.

3.3.3 Classification model

After feature extraction, five dense blocks $D_i, i \in (1, 5)$ are used to learn the cohesive salient patterns among the extracted features to distinguish between real and spoof faces. The dimensions of corresponding dense blocks D_1, D_2, D_3, D_4 and D_5 are $([1 \times 128], [1 \times 512], [1 \times 32], [1 \times 64]$ and $[1 \times 32]$ respectively.

Dense blocks are the learnable feature encodings observed in Modified Xception Deep branch, as F_{Deep}, and Multilevel ELBP branch, as F_{MELBP}, respectively. F_{Deep} represents deep feature interpretation of full-sized input image, whereas F_{MELBP} gives a multi-level texture representations with linear operations, extracted from cropped faces using simple mathematical functions, as described in Algorithm 1. D_2, andD_3 dense blocks introduce the

required non-linearity in the multilevel ELBP branch and convert the texture features into learnable vectors which capture the local inconsistences among patches. Once the dimensions of the salient features obtained from both branches F_{Deep} and F_{MELBP} are reduced with corresponding dense layers, we concatenate both the features as $F_{concat}([F_{Deep}][F_{MELBP}])$. After concatenation, an additional D_4 layer is used to learn from F_{concat} before the final classification.

ReLU [53] is used as a non-linear activation function in all the dense layers and sigmoid in the final classification layer D_5. The size of the dense layers, has been chosen after a rigorous experimental analysis. The details of selection of right no. of neurons in each dense layer is provided in ablation study in experiment section IV.

4 Database and experimental analysis

This section introduces the datasets used for model performance evaluation and the associated experiment methodology in detail. The ablation study of the performance variations of the proposed framework for different model designs is also discussed.

4.1 Benchmark database

To verify the effectiveness and robustness of the proposed framework, we evaluate it on three challenging face anti-spoofing database. Detail description of the database is as follows:

CASIA FASD [6] comprises 600 videos of 50 subjects capture from different imaging devices, and the training set has 240 videos corresponding 20 subjects, the testing set has 360 videos corresponding 30 subjects. It holds three types of image qualities- low, normal, and high quality. To reduce the calculation burden and alleviate savings, high-quality videos are cropped into a resolution of $[1280 \times 720]$ pixels. The database includes three types of fake attacks, namely, cut photo attacks, warped photo attacks and video attacks.

REPLAY-ATTACK [7] comprises 1200 short video clips corresponding to 50 distinct individuals recorded under two different illumination and support conditions. The database also considers two types of holding conditions i.e., hand held and stand fixed. There are 200 actual access videos, 200 printed photo attacks, 400 mobile-based attacks, and 400 high-resolution video attacks. The database comprises four folders, but we only used three folders, namely train, validation, and test data. The three types of attacks in the database are: (i) mobile attacks are attempted using the iPhone 3GS screen. (ii) A high-resolution screen attack is attempted using an iPad (iii) An A4 sheet of paper is used for a hard-copy print attack.

MSU-MFSD [54] There are 280 video clips of attack and genuine faces in this dataset. The dataset is built using 35 people who contributed in the creation of dataset. Each participant has two videos with android phones and laptops respectively for real access. The Canon and iPhone cameras were two types of cameras employed for capturing HD videos for every subjects. The iPhone mobile recordings were replayed to generate the mobile replay attacks while the Canon camera footage were then used for generating the high-definition replay attacks by replaying on the iPad Air screen. Images were printed on A3 paper with a color HP printer for all 35 subjects for generating photo attacks. These 35 individuals' recording videos were split into testing (160 videos for 40 subjects) and training (120 videos for 15 subjects) datasets.

4.2 Evaluation metrics

We report Equal Error Rate (EER) and Half Total Error Rate (HTER) for comparative analysis with other state-of-art. The point on ROC or DET curve where False acceptance rate is equivalent to false rejection rate is termed as EER.

$$\text{EER} = (\text{FAR} = \text{FRR}). \tag{12}$$

HTER is defined as a point where average of False Rejection Rate and False Acceptance Rate is minimal.

$$\text{HTER} = \frac{\text{FRR} + \text{FAR}}{2}. \tag{13}$$

In addition to the above parameters, we also provide the ROC curve and Area Under the Curve (AUC) to evaluate the overall performance of our model.

4.3 Implementation details

The architecture was implemented with Keras, and the experiment is performed on Google Colab pro with an Nvidia P100 GPU with 24 GB RAM. We processed the videos with frame sampling at every 10th frame for image pre-processing and horizontal flip for data augmentation. Adam optimizer is used for network training with learning rate is set at $10e-4$. Batch size of 32 with random shuffling and flipping of images is used. The modified Xception model branch and ELBP branch were trained individually for each dataset and were freezed for adaptive fusion-based classification.

4.4 Comparative analysis with other state-of-arts

To demonstrate the efficiency of our proposed work, performance of the proposed work is compared with other state-of-the-arts. In Tables 1 and 2, the performance comparison of the proposed work with other state-of-the-arts is reported for *CASIA FASD* [6], *REPLAY-ATTACK* [7], *MSU-MFSD*

Table 1 Intra-dataset evaluation on CASIA FASD, Replay-Attack and MSU-MFSD datasets (%)

Methods	CASIA FASD		Replay-Attack		MSU-MFSD	
	EER	HTER	EER	HTER	EER	HTER
YCbCr + HSV-LBP [55]	6.2	–	**0.4**	2.9	–	–
DMD [56]	21.7	–	5.3	3.7	–	–
CNN LBP + TOP [57]	8.02	9.94	3.22	4.70	–	–
Fine-tuned VGG- Face [58]	5.2	–	8.40	4.3	–	–
DPCNN [58]	4.5	–	2.90	6.1	–	–
Color SURF [26]	2.8	–	0.1	2.2	2.2	
Hybrid CNN [59]	**2.2**	–	0.5	1.6	–	–
RI-LBP + Speeded Up Robust Features (SURF) [60]	4.6	–	1.2	4.2	1.5	
Color Texture Markov Feature (CTMF) [50]	8.0	–	4.0	4.4	7.5	
LSCNN [61]	4.44	–	0.33	2.50	–	–
RI-LBP + Deep Features [16]	4.4	–	2.3	2.6	3.1	–
Adversarial [62]	3.2	–	**–**	1.4		
Patch CNN [35]	4.8	–	1.5	–	–	–
Proposed method ($L_0 + L_1$)	6	8.7	0.5	0	0.10	0.04
Proposed method ($L_0 + L_1 + L_2$)	**2.37**	**3.2**	**0**	**0**	**0**	**0.06**

Table 2 Cross-dataset evaluation on CASIA FASD, Replay-Attack and MSU-MFSD in terms of HTER (%)

Train/test datasets methods	CASIA FASD		Replay-attack		MSU-MFSD	
	Replay-attack	MSU-MFSD	CASIA FASD	MSU-MFSD	CASIA FASD	Replay-attack
Colour SURF [26]	26.9	31.8	**23.2**	**19.1**	24.3	29.7
RI-LBP + SURF [60]	**9.6**	19.8	39.2	33.3	29.7	21.4
CTMF [50]	32.3	32.4	45.9	37.7	57.0	42.7
RI-LBP + Deep Features [16]	35.3	**2.1**	38.5	20.6	32.4	35.8
Adversarial [62]	17.5	**9.3**	41.6	30.5	**17.7**	**5.1**
DRL-FAS [63]	28.4	–	33.2	**15.6**	–	29.7
CFSA-FAS [64]	24.3	–	34.0	25.2	–	22.8
GFA-CNN [65]	21.3	–	34.3	23.5	–	25.8
SAPLC [66]	27.31	–	37.50	–	–	–
CDCN + + [44]	6.5	–	29.8	–	–	–
DR-UDA [67]	**15.6**	**9.0**	34.2	29.0	**16.8**	**3.0**
Proposed method ($L_0 + L_1$)	44.71	30.6	36.89	35.37	35.27	37.7
Proposed method ($L_0 + L_1 + L_2$)	**20.05**	**9.5**	**28**	**21.8**	**23.02**	**24.3**

[54] datasets. The experiments are conducted for both intra-dataset testing and inter-dataset testing evaluation strategies. The bold values in Tables [1, 2, 3, 4 and 5] represent best reported results.

4.4.1 Intra-dataset testing

In intra-dataset testing evaluation strategy, we evaluate the performance in terms of both EER and HTER. The proposed work yields superior results on all three datasets. 0% HTER and EER is observed for Replay-Attack. Similarly, for MSU-MFSD our method performs at par with HTER of 0% and EER of 0.06%. This shows the proposed work is able to adapt with the types of spoof cases provided in the dataset. The proposed framework is able to perfectly classify Replay-Attack and MSU-MFSD datasets for an ELBP at L_0 and L_1 levels, while CASIA proved to be challenging. Hence, additional segmentation of ELBP image at level L_2, was performed to extract more fine-grained features. This resulted in a 50% reduction in both HTER and EER.

CASIA FASD dataset thus achieved highly competitive results with HTER of 2.37% and EER of 3.2% which implies our approach has lower FAR and FRR.

The results ensure that the proposed work is able to detect the print and video attacks for all three datasets in intra-dataset testing conditions with high accuracy. However, the intra- dataset testing condition constrains the spoof detection to dataset specific environments.

Therefore, an inter-database analysis is performed to further confirm the robustness of our proposed work against unknown types of spoofs attack in the real world i.e., uncontrolled environment.

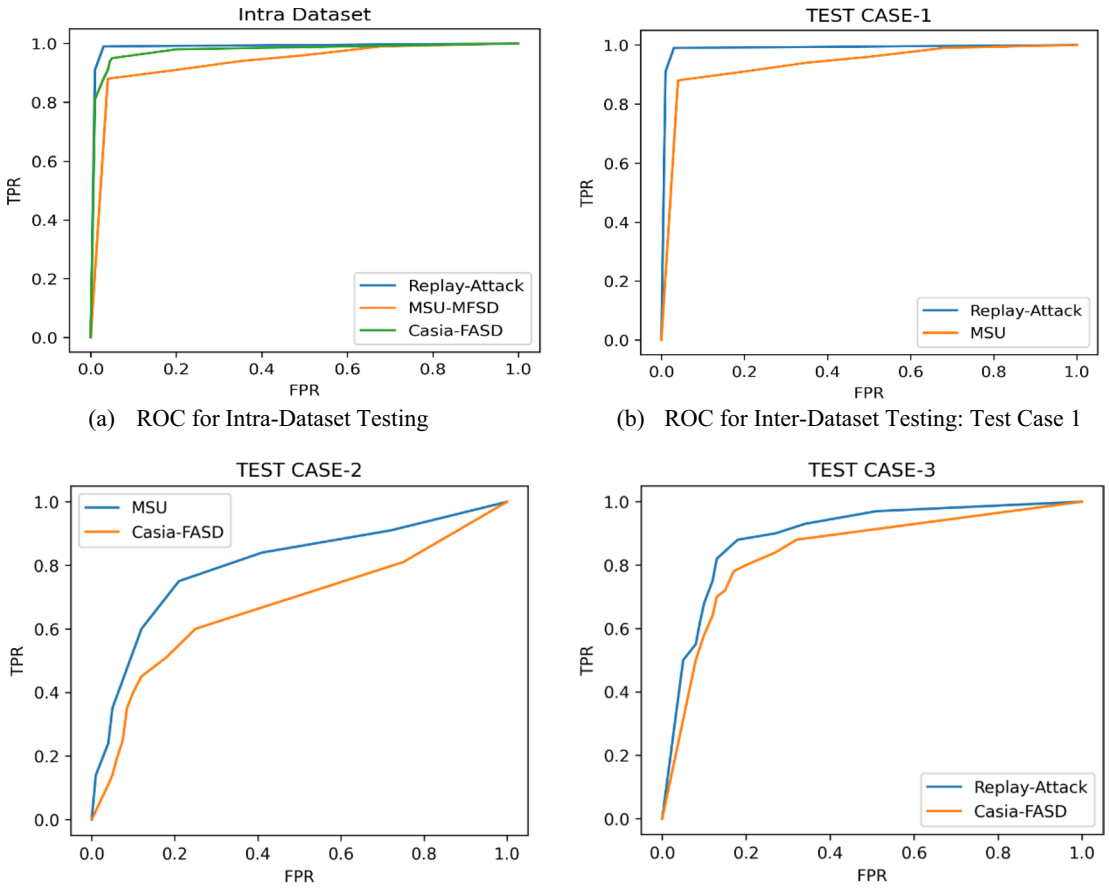
4.4.2 Inter-dataset testing

For practical applications of face anti-spoofing, ability of any anti-spoofing framework to adapt from one dataset to another dataset is important. Therefore, to test the generalizability of our framework, we extend the experiment for cross-dataset testing. Many recent studies [16, 26, 60] describe cross-dataset testing with the MSU-MFSD, CASIA FASD, and Replay-Attack datasets. The tests were conducted thrice for each combination to ascertain consistency, the results reported are the mean of the 3 observations (see Table 2).

Test Case1 *(**Train** -CASIA FASD and **Test** -Replay-Attack and MSU-MFSD):* For Model trained on CASIA FASD, the proposed work achieves an HTER of 20.05% on Replay and 9.5% on MSU-MFSD. Though, the work [38] has reported 9.6% HTER for CASIA and Chen et al. [16] has reported 2.1% HTER for MSU. The performance of these works [16, 38] is not consistent for different dataset combination.

Test Case2 *(**Train**- Replay-Attack and **Test**- CASIA FASD and MSU-MFSD):* Model trained on Replay**,** we observe almost higher HTER for both test cases, as 28% for CASIA and 21.8% for MSU except Color-SURF [26] and DRL-FAS [63] respectively. The proposed model is thus able to generalize better for Test case 2 than test case 1. This can be attributed to the large size and variety of attacks present in CASIA as compared to Replay. Despite the drop in performance as compared to previous scenario, our model still outperforms most other reported methods.

Test Case 3 *(**Train**-MSU-MFSD and **Test**- CASIA FASD and Replay-Attack):* We observed similar results for the proposed model trained on MSU-MFSD and tested on other two datasets. Though the method achieves an HTER of 23.02% and 24.3% for CASIA and Replay respectively.



(a) ROC for Intra-Dataset Testing

(b) ROC for Inter-Dataset Testing: Test Case 1

(c) ROC for Inter-Dataset Testing: Test Case 2

(d) ROC for Inter-Dataset Testing: Test Case 3

Fig. 6 ROC for **a** intra-dataset testing case; **b** Inter-dataset testing (***Test case-1)*****: Train-**CASIA FASD and **Test-** Replay-Attack and MSU-MFSD, **c** Inter-dataset testing (***Test Case-2):*** **Train**- Replay-Attack and **Test**- CASIA FASD and MSU-MFSD, **d** Inter-dataset testing (***Test case-3)*****: Train**-MSU-MFSD and **Test**- CASIA FASD and Replay-Attack

The same model achieved nearly perfect separation between attack and bonafide samples in the intra-testing scenario for the dataset MSU. These results, while subpar as compared to performance in Test case 1 and Test case 2, are superior to other state-of-arts.

Most works report great performance with less than 20% HTER/EER for specific combination for inter-dataset, while the error rates for other combinations are unusually high (> 30%). For our method, we report uniform HTER and EER (< 30%) as for all combinations of inter-dataset testing, while showing exceptional performance for intra-dataset testing (< 4% HTER for all), implying the proposed method is not biased toward any of the datasets. The generalization performance is thus fairly consistent across a wide variety of subjects, lighting, picture qualities and environmental variations. This ensures that further enhancements of our work will be able to deliver similar performance improvements for all datasets. To evaluate the performance more comprehensively, the results are presented in terms of ROC curves for both intra-dataset and cross-dataset testing. Figure 6 depicts these ROC curves.

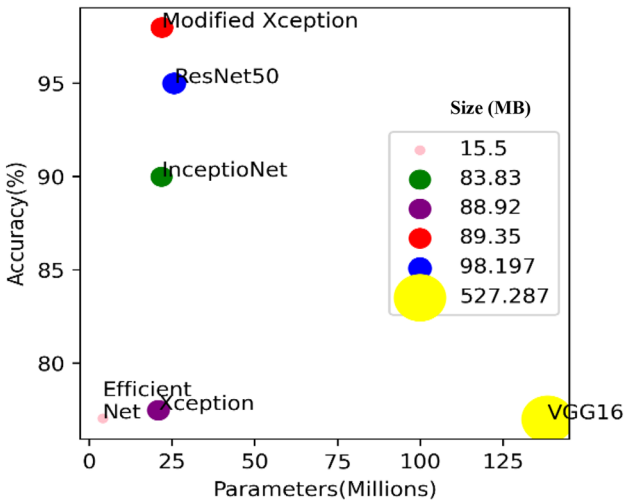


Fig. 7 Comparison of different generic architectures for search of baseline

4.5 Ablation analysis

We evaluate the effectiveness of the proposed approach by performing series of ablation experiments to choose appropriate baseline model, analyze variation in baseline model, and exhibit the effect of different color space for efficient multi-ELBP feature representation as discussed below:

4.5.1 Analysis for selection of baseline

Lack of exploration of the Xception model in the field of FAS and its superior performance in other applications motivated us to use it as our baseline for FAS task. Also, an analysis to ascertain its performance relative to other popular baseline models was conducted. Individually all models were trained on Face anti-spoofing dataset. The performance of various models i.e., VGG16 [68], ResNet50 [18], Xception [69], Efficient Net [19], and Inception [70] is analyzed in terms of trainable parameters, model size and accuracy as shown Fig. 7. It was observed that ResNet50 [18] performs comparatively better over other models at the cost of large trainable parameters. Whereas, it is observed the proposed.

modifications in Xception model results in significant rise in the performance of the classification model with a nominal change in no. of trainable parameters with comparable model size. After modification, the model performs better than ResNet50 with lesser number of parameters and size.

4.6 Analysis of variation in baseline Xception [69] model

The deep features are extracted using efficient Xception [69] model as baseline, using transfer learning. In modified Xception model, pretrained Xception [69] model is supplemented using squeeze and excitation blocks to further enhance the feature extraction at multiple levels. We introduce SE module after each block (i.e., Entry block, middle block and exit block) in the original Xception model [69], as shown in Fig. 5a. To validate the performance, we conduct various experiments to demonstrate the rationality of the modified

Table 3 Ablation Study: Intra-dataset evaluation on modified Xception architecture for CASIA FASD, Reply-Attack and MSU-MFSD (%)

Variation	CASIA FASD		Replay-Attack		MSU-MFSD	
	EER	HTER	EER	HTER	EER	HTER
Baseline Xception [69]	50.1	50	36.73	29.03	33.11	25.29
SE at Entry level	14	10.5	3.64	1.89	10.68	9.7
SE at Middle Level	11.3	9.1	8.28	4.68	8.08	4.5
Modified Xception	**7.45**	**5.08**	**2.45**	**1.5**	**0.92**	**0.46**

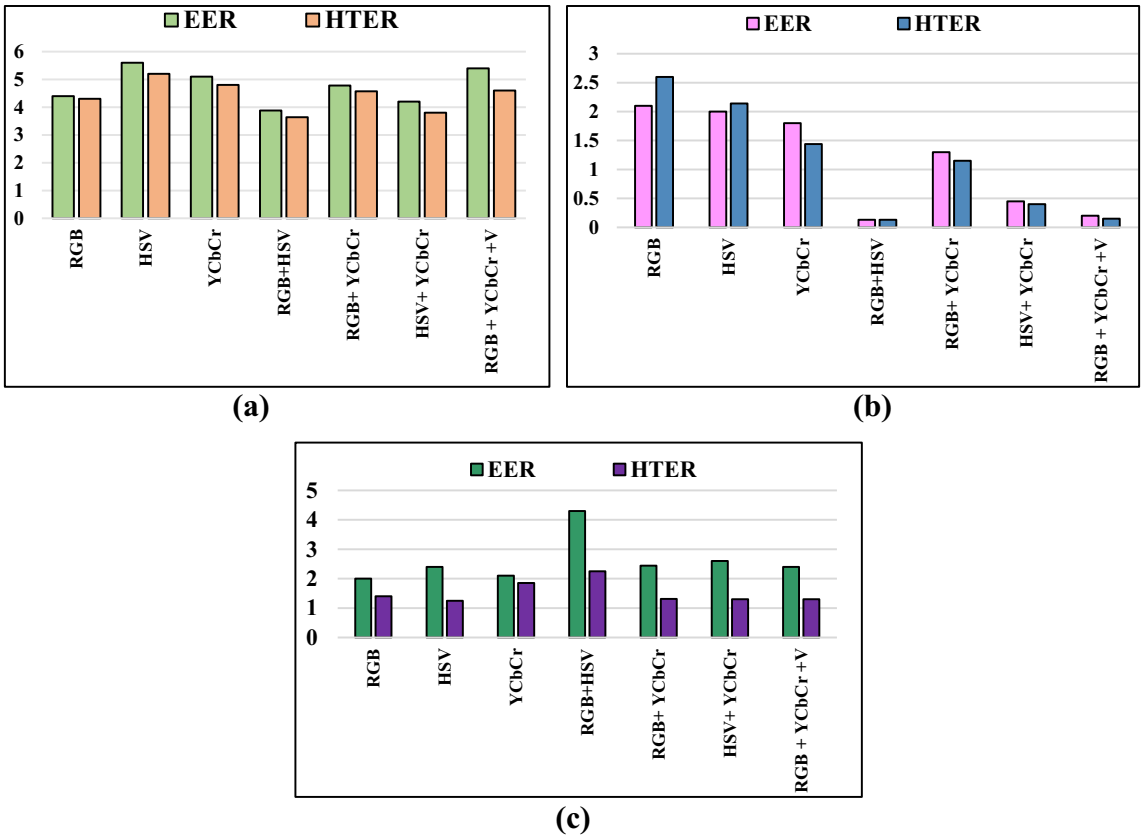


Fig. 8 HTER and EER reported for **a** CASIA FASD, **b** MSU-MFSD, **c** Replay-Attack using Intra-dataset testing

Xception model. In Table 3, the observed results for baseline Xception model [69] and modified Xception model are reported. Each variation was trained for 10 epochs for all three datasets. It can be seen from Table 3 that deep features extracted from modified Xception model result in significant drop of EER and HTER values for all three datasets. It exhibits the superiority of the modified Xception model over baseline model. Therefore, modified Xception model is employed for richer deep feature representation.

4.7 Analysis of the effect of different color space

RGB color space represents image in terms of combination of primary colors. RGB is not able to represent image in the way humans perceive image. Color spaces such as HSV and YCbCr separate color information from intensity/lamination. Modified Xception deep feature branch accepts only tri-channel images whereas ELBP can be applied on a per channel basis and for different levels of sub region. The ELBP branch of architecture has modules which convert,

Table 4 Cross-dataset evaluation of Level-0 ELBP texture feature extraction for different color space in proposed method on CASIA FASD, MSU-MFSD and Reply-Attack for ablation analysis (HTER %)

Train	CASIA FASD		MSU-MFSD		Replay-Attack	
Test	MSU MFSD	Replay-Attack	CASIA FASD	Replay-Attack	CASIA FASD	MSU MFSD
RGB	27.88	30.35	30.55	29.87	23.53	22.75
HSV	27.69	28.77	31.68	29.51	24.39	22.14
YCbCr	26.08	31.20	30.04	26.70	17.76	16.29
RGB + HSV	27.78	31.08	30.18	29.99	24.75	21.78
RGB + YCbCr	27.49	28.86	30.37	29.20	23.92	22.31
HSV + YCbCr	22.35	26.1	21.5	25.6	24.54	22.17
RGB + YCbCr + HSV	**18.61**	**20.41**	**18.21**	**19.89**	**22.69**	**18.79**

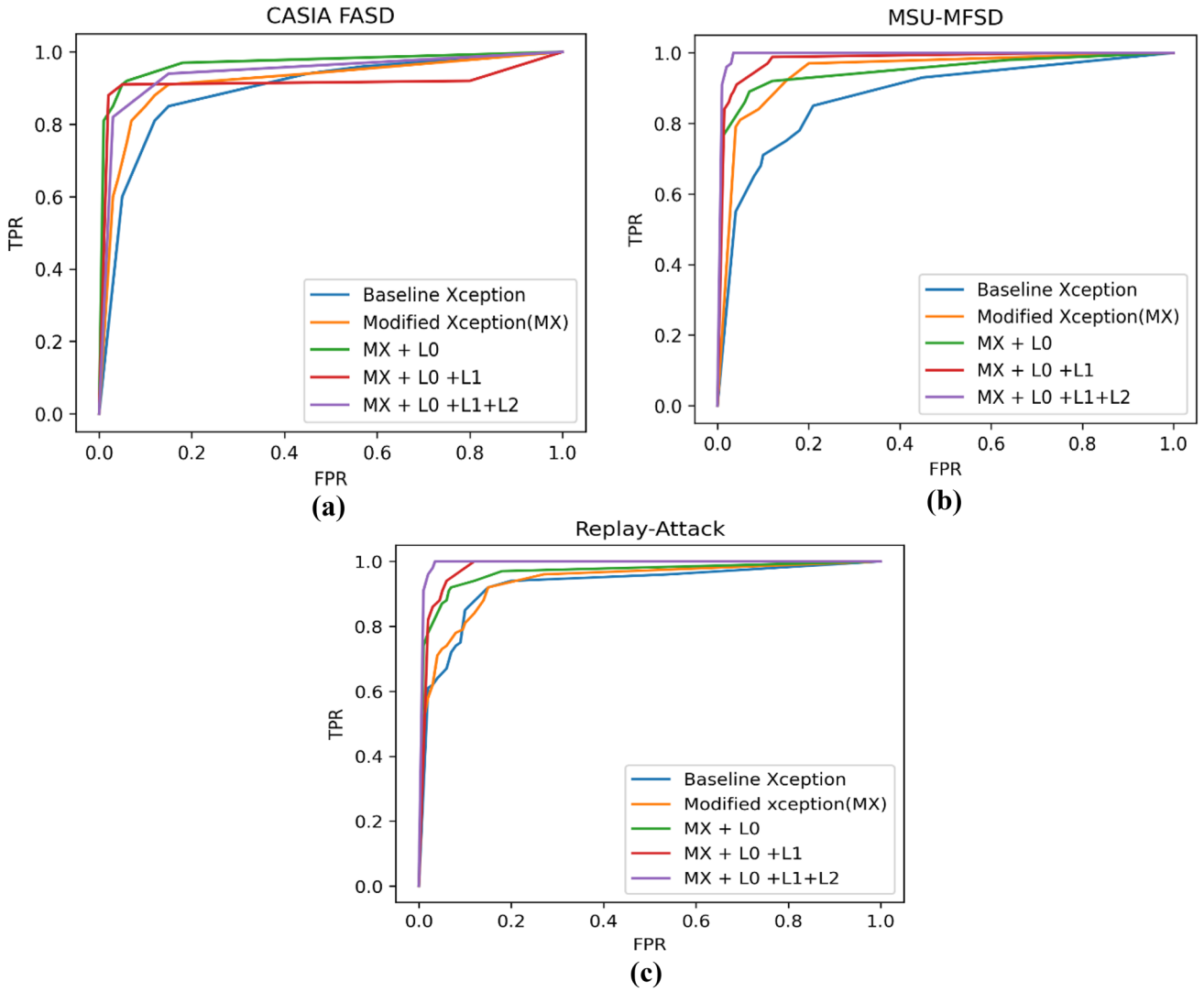


Fig. 9 Proposed model performance for different model designs on **a** CASIA FASD, **b** MSU MFSD, **c** Replay Attack using Intra-dataset testing

separate and perform ELBP on channel from each of color spaces. We performed extensive analysis for each combination of color spaces under both intra and inter-dataset testing. Figure 8a–c report *HTER* and *EER* for different color spaces on all three datasets using intra-dataset testing. We observe that in intra-dataset testing from Fig. 8a–c CASIA and MSU-MFSD perform best for $RGB + HSV$ color space with an EER of 3.88% and 0.13% respectively.

Whereas, Replay Attack dataset performs best for $RGB + HSV + YCbCr$ with 1.3% *HTER*. When the same experiment is performed in inter-dataset testing, the dataset challenges increase. In Inter-dataset testing, the test samples appear from different types of datasets acquired in different

Table 5 Selection of nodes in dense layers (D_i)

D_1	D_2	D_3	D_4	D_5	Accuracy			HTER		
					CASIA	RA	MSU	CASIA	RA	MSU
256	256	–	64	2	96.354	98.21	98.24	3.7	0.02	0.84
256	512	64	32	2	94.44	99	99.54	4.5	0	0.78
128	256	32	64	2	97.6	99.54	98.35	3.55	0.02	1.2
128	512	–	32	2	98.1	98.54	99.84	3.1	0.01	0.7
128	**512**	**32**	**64**	**2**	**98.32**	**99.98**	**99.76**	**3.2**	**0**	**0.06**

lighting conditions with different types of spoofing attacks. As reported in Table 4, high values of HTER for both RGB and HSV color spaces confirms that RGB and HSV color spaces are not sufficient to represent the unknown attack types. Out of all combinations, $RGB + HSV + YCbCr$ color space performs best for inter-dataset testing. These results correspond to L_0 *ELBP* for each color space; thus, we arrive at the conclusion that the combination $RGB + YCbCr + HSV$ color space should be selected for multi-ELBP feature representation. To support the observation the improvement in the proposed framework by inclusion of multi-ELBP features at different levels, is also shown in Fig. 9 in terms of ROC curves. It is clearly visible that fusion of ELBP features for $(L_0 + L_1 + L_2)$ levels with modified Xception based deep features exhibits superior performance.

4.8 Ablation study for number of nodes in Dense Layer

To perform binary classification, using feature vector extracted from modified Xception model, $F_{modXcept}$ $[1 \times 3504]$ (see Fig. 5a) and the multi-level ELBP feature vector, $F_{texture}$ $[1 \times 2478]$ (see Fig. 4), dense layers are introduced to gradually reduce the feature vector dimensions to two neurons. The size of each dense layer has been chosen after a rigorous experimental analysis as detailed in the Table 5. The best results were obtained for the values of $\{\boldsymbol{D_i}\}$ as $[128, 512, 32, 64, 2]$ for all three datasets.

Model running time: In testing phase, with only Modified Xception stream, the model takes approximately 0.71 s per batch of 32 images. Whereas, considering both the streams together, 0.82 s average time per batch is observed for classification. This overhead of 0.11 s is due to additional online face cropping and multi-level ELBP computations. Thus, the throughput for complete framework is ~ 40 frames per second.

5 Conclusion and future work

In this study, we proposed a novel face anti-spoofing framework using multi-level ELBP features and modified Xception based deep features. The experiment results shows that our work exhibits superior performance on intra-dataset testing and has good generalization capability compared to other state-of-arts. The method achieved near perfect spoofing detection with HTER of 0% in Replay and MSU. The results for generalization testing were consistent with less than 30% HTER for even the most complex scenarios. Thus, under certain conditions the provided technique can actually be well promoted for face spoofing detection. The modification applied on the Xception network improved the performance significantly with an 80% accuracy increase across all scenarios. The ablation study confirms the initial hypothesis that rich discriminative features are present in luminance-based representations of the RGB color space. These results and the ablation study provide direction for future performance improvements. This work thus confirms the advantage of fusion base methods for discriminative feature extraction. Our future focus would be toward designing a methodology with an excellent generalization ability in cross-datasets for face spoofing detection.

Author contributions The authors confirm contribution to the paper as follows:Study conception and design, Interpretation of results, Review, Supervision: Dr. Chhavi DhimanData processing, analysis and interpretation of results, draft manuscript preparation: Aashania Antil All authors reviewed the results and approved the final version of the manuscript.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

References

1. Rui, Z., Yan, Z.: A survey on biometric authentication: toward secure and privacy-preserving identification. IEEE Access **7**, 5994–6009 (2018)
2. Ramachandra, R., Busch, C.: Presentation attack detection methods for face recognition systems: a comprehensive survey. ACM Comput. Surv. (CSUR) **50**(1), 1–37 (2017)
3. Liu, A., Zhao, C., Yu, Z., Wan, J., Su, A., Liu, X., Tan, Z., Escalera, S., Xing, J., Liang, Y., Guo, G., Lei, Z., Li, S. Z., and Zhang, D.: Contrastive Context-Aware Learning for 3D High-Fidelity Mask Face Presentation Attack Detection. arXiv:2104.06148 [cs.CV] (2021)
4. Singh, M. S., Singh, R., Vatsa, M., Ratha, N. and Chellappa, R.: Recognizing Disguised Faces in the Wild. arXiv:1811.08837v1 [cs.CV] (2018)
5. Yu, Z., Qin, Y., Li, X., Zhao, C., Lei, Z. and Zhao, G.: Deep Learning for Face Anti-Spoofing: A Survey, arXiv:2106.14948 [cs.CV] (2021)
6. Zhang, Z., Yan, J., Liu, S., Lei, Z., Yi, D. and Li, S. Z.: A face antispoofing database with diverse attacks. In 5th IAPR International Conference on Biometrics (ICB), New Delhi, India (2012)
7. Chingovska, I., Anjos, A. and Marcel, S.: On the effectiveness of local binary patterns in face anti-spoofing. In Proceedings of the International Conference of Biometrics Special Interest Group (BIOSIG), Darmstadt, Germany (2012)
8. Tu, X. and Fang, Y.: Ultra-deep neural network for face anti-spoofing. In: International Conference on Neural Information Processing. Guangzhou, China (2017)
9. Lucena, O., Junior, A., Moia, V., Souza, R., Valle, E. and Lotufo, R.: Transfer learning using convolutional neural networks for face anti-spoofing. In International Conference Image Analysis and Recognition, Montreal, Canada (2017)

10. Turk, M. A., Pentland, A. P.: Face recognition using eigenfaces. In: IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), Maui, HI, USA (1991)
11. Belhumeur, P.N., Hespanha, J.P., Kriegman, D.J.: Eigenfaces vs. Fisherfaces: recognition using class specific linear projection. IEEE Trans. Pattern Anal. Mach. Intell. **19**(7), 711–720 (1997)
12. Vu, N.-S. and Caplier, A.: Face recognition with patterns of oriented edge magnitudes. In: Computer Vision – ECCV, Heraklion, Crete (2010)
13. Zhang, B., Shan, S., Chen, X., Gao, W.: Histogram of gabor phase patterns (HGPP): a novel object representation approach for face recognition. IEEE Trans. Image Process. **16**(1), 57–68 (2007)
14. Ahonen, T., Hadid, A., Pietikainen, M.: Face description with local binary patterns: application to face recognition. IEEE Trans. Pattern Anal. Mach. Intell. **28**(12), 2037–2041 (2006)
15. Nguyen, H. V., Bai, L. and Shen, L.: Local Gabor binary pattern whitened PCA: a novel approach for face recognition from single image per person. In: International Conference on Biometrics (ICB), Alghero, Italy (2009)
16. Chen, F.-M., Wen, C., Xie, K., Wen, F.Q., Sheng, G.Q., Tang, X.-G.: Face liveness detection: fusing colour texture feature and deep feature. IET Biometrics **8**(6), 369–377 (2019)
17. Chollet, F.: Xception: deep learning with depthwise separable convolutions. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA (2017)
18. He, K., Zhang, X., Ren, S. and Sun, J.: Deep residual learning for image recognition. arXiv:1512.03385 [cs.CV] (2015)
19. Tan, M. and Le, Q. V.: EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. arXiv:1905.11946v5 [cs.LG] (2020)
20. Kong, C., Chen, B., Li, H., Wang, S., Rocha, A., Kwong, S.: Detect and locate: exposing face manipulation by semantic- and noise-level telltales. IEEE Trans. Inf. Forensics Secur. **17**(28 April), 1741–1756 (2022)
21. Qian, Y., Yin, G., Sheng, L., Chen, Z., Shao, J.: Thinking in frequency: face forgery detection by mining frequency-aware clues. In: European Conference on Computer Vision (ECCV), Glasgow, UK (2020)
22. Kong, C., Chen, B., Yang, W., Li, H., Chen, P., Wang, S.: Appearance Matters, So Does Audio: Revealing the Hidden Face via Cross-Modality Transfer. IEEE Trans. Circuit Syst. Video Technol. **32**(1 Jan), 423–436 (2022)
23. Hu, J., Shen, L., Sun, G.: Squeeze-and-excitation networks. In: IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA (2018)
24. Gragnaniello, D., Poggi, G., Sansone, C., Verdoliva, L.: An investigation of local descriptors for biometric spoofing detection. IEEE Trans. Inf. Forensics Secur. **10**(4), 849–863 (2015)
25. Boulkenafet, Z., Komulainen, J., Hadid, A.: Face spoofing detection using color texture analysis. IEEE Trans. Inf. Forensics Secur. **11**(8), 1818–1830 (2016)
26. Boulkenafet, Z., Komulainen, J., Hadid, A.: Face antispoofifing using speeded-up robust features and fisher vector encoding. IEEE Signal Process. Lett. **24**(2), 141–145 (2017)
27. Peng, F., Qin, L. and Long, M.: CCoLBP: Chromatic co-occurrence of local binary pattern for face presentation attack detection," in *27th International Conference on Computer Communication and Networks (ICCCN)*, Hangzhou, China , 2018.
28. Li, L., Feng, X.: Face anti-spoofing via deep local binary pattern. In Deep Learning in Object Detection and Recognition, pp. 91–111. Springer, Singapore (2019)
29. Li, L., Feng, X., Xia, Z., Jiang, X., Hadid, A.: Face spoofing detection with local binary pattern network. J. Vis. Commun. Image Represent. **54** (July), 182–192 (2018)
30. Zhao, X., Lin, Y., Heikkilä, J.: Dynamic texture recognition using volume local binary count patterns with an application to 2D face spoofing detection. IEEE Trans. Multimedia **20**(3), 552–566 (2018)
31. Shu, X., Pan, L., Wu, X.-J.: Multi-scale contour flexibility shape signature for Fourier descriptor. J. Vis. Commun. Image Represent. **26**, 161–167 (2015)
32. Guo, Z., Zhang, L., Zhang, D.: Rotation invariant texture classification using LBP variance (LBPV) with global matching. Pattern Recog. **43**(3), 706–719 (2010)
33. Tan, X., Triggs, B.: Enhanced local texture feature sets for face recognition under difficult lighting conditions. IEEE Trans. Image Process. **19**(6), 1635–1650 (2010)
34. Chen, H., Hu, G., Lei, Z., Chen, Y., Robertson, N.M., Li, S.Z.: Attention-based two-stream convolutional networks for Face Spoofing Detection. IEEE Trans. Inf. Forensics Secur. **15**, 578–593 (2019)
35. Almeida, W. R., Andaló, F. A., Padilha, R., Bertocco, G., Dias, W., Torres, R. d. S., Wainer, J. and Rocha, A.: Detecting face presentation attacks in mobile devices with a patch-based CNN and a sensor-aware loss function. PLos One 15(9) (2020)
36. Chen, B., Yang, W., Li, H., Wang, S., Kwong, S.: Camera invariant feature learning for generalized face anti-spoofing. IEEE Trans. Inf. Forensics Secur. **16**, 2477–2492 (2021)
37. Xu, X., Xiong, Y. and Xia, W.: On Improving Temporal Consistency for Online Face Liveness Detection. arXiv:2006.06756 [cs.CV] (2020)
38. Liu, Y., Jourabloo, A. and Liu, X.: Learning deep models for face anti-spoofing: binary or auxiliary supervision. In IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA (2018)
39. Hossain, M. S., Rupty, L., Roy, K. and Hasan, M.: A-DeepPixBis: attentional angular margin for face anti-spoofing. In Digital Image Computing: Techniques and Applications (DICTA), Melbourne, Australia (2020)
40. Yu, Z., Qin, Y., Xu, X., Zhao, C., Wang, Z., Lei, Z. and Zhao, G.: Auto-Fas: searching lightweight networks for face anti-spoofing. In IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Barcelona, Spain (2020)
41. Kong, C., Zheng, K., Wang, S., Rocha, A., Li, H.: Beyond the pixel world: a novel acoustic-based face anti-spoofing system for smartphones. IEEE Trans. Inf. Forensics Secur. **17**, 3238–3253 (2022)
42. Wang, Z., Zhao, C., Qin, Y., Zhou, Q., Qi, G., Wan, J. and Lei, Z.: Exploiting temporal and depth information for multi-frame face anti-spoofing. arXiv:1811.05118 [cs.CV] (2019)
43. Li, H., He, P., Wang, S., Rocha, A., Jiang, X., Kot, A.C.: Learning Generalized Deep Feature Representation for Face Anti-Spoofing,". IEEE Trans. Inf. Forensics Secur. **13**(10), 2639–2652 (2018)
44. Yu, Z., Zhao, C., Wang, Z., Qin, Y., Su, Z., Li, X., Zhou, F., Zhao, G.: Searching central difference convolutional networks for face anti-spoofing. In: International Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, Washington (2020)
45. Yu, Z., Wan, J., Qin, Y., Li, X., Li, S.Z., Zhao, G.: Nas-fas: static-dynamic central difference network search for face anti spoofing. IEEE Trans. Pattern Anal. Mach. Intell. **43**(9), 3005–3023 (2021)
46. Yu, Z., Li, X., Niu, X., Shi, J., Zhao, G.: Face anti-spoofing with human material perception. In European Conference on Computer Vision (ECCV), Glasgow, United Kingdom (2020)
47. Qin, Y., Yu, Z., Yan, L., Wang, Z., Zhao, C., Lei, Z.: Meta-teacher for face anti-spoofing. IEEE Trans. Pattern Anal. Mach. Intell. (2021)
48. V. J. and S.: Detecting pedestrians using patterns of motion and appearance. In Ninth IEEE International Conference on Computer Vision, Nice, France (2003)
49. Zhang, K., Zhang, Z., Li, Z., Qiao, Y.: Joint face detection and alignment using multi-task cascaded convolutional networks. In:

Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA (2016)
50. Zhang, L.-B., Peng, F., Qin, L., Long, M.: Face spoofing detection based on color texture markov feature and support vector machine recursive feature elimination. J. Visual Commun. Image Represent. **51**, 56–69 (2018)
51. Zhao, W.-Y., Chellappa, R., Phillips, P.J., Rosenfeld, A.: Face recognition: a literature survey. ACM Comput. Surv. **35**(4), 399–458 (2003)
52. Nguyen, H.-T. and Caplier, A.: Elliptical local binary patterns for face recognition. In Asian Conference on Computer Vision. Daejeon, Korea (2012)
53. Nair, V. and Hinton, G. E.: Rectified linear units improve restricted boltzmann machines. In ICML, Haifa, Israel (2010)
54. Wen, D., Han, H., Jain, A.K.: Face spoof detection with image distortion analysis. IEEE Trans. Inf. Forensics Secur. **10**(4), 746–761 (2015)
55. Boulkenafet, Z., Komulainen, J. and Hadid, A.: Face anti-spoofing based on color texture analysis. In IEEE International Conference on Image Processing (ICIP), Quebec City, QC, Canada (2015)
56. Tirunagari, S., Poh, N., Windridge, D., Iorliam, A., Suki, N., Ho, A.T.S.: Detection of face spoofing using visual dynamics. IEEE Trans. Inf. Forensics Secur. **10**(4), 762–777 (2015)
57. Asim, M., Ming, Z. and Javed, M. Y.: CNN based spatio-temporal feature extraction for face anti-spoofing. In: 2nd International Conference on Image, Vision and Computing (ICIVC), Chengdu, China (2017)
58. Li, L., Feng, X., Boulkenafet, Z., Xia, Z., Li, M. and Hadid, A.: An original face anti-spoofing approach using partial convolutional neural network. In Sixth International Conference on Image Processing Theory, Tools and Applications (IPTA), Oulu, Finland (2016)
59. Li, L., Xia, Z., Li, L., Jiang, X., Feng, X. and Roli, F.: Face anti-spoofing via hybrid convolutional neural network. In 2017 International Conference on the Frontiers and Advances in Data Science (FADS), Xi'an, China (2017)
60. Boulkenafet, Z., Komulainen, J., Hadid, A.: On the generalization of color texture-based face anti-spoofing. Image Vis. Comput. **77**(September), 1–9 (2018)
61. Souza, G. B. d., Papa, J. P. and Marana, A. N.: On the learning of deep local features for robust face spoofing detection. In 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), Parana, Brazil (2018)
62. Wang, G., Han, H., Shan, S. and Chen, X.: Improving cross-database face presentation attack detection via adversarial domain adaptation. In International Conference on Biometrics (ICB), Crete, Greece (2019)
63. Cai, R., Li, H., Wang, S., Chen, C., Kot, A.C.: Drl-fas: a novel framework based on deep reinforcement learning for face anti-spoofing. IEEE Trans Inf. Forensics Secur. **19**, 937–951 (2020)
64. Chen, B., Yang, W. and Wang, S.: Face anti-spoofing by fusing high and low frequency features for advanced generalization capability. In IEEE Conference on Multimedia Information Processing and Retrieval (MIPR), Shenzhen, China (2020)
65. Tu, X., Ma, Z., Zhao, J., Du, G., Xie, M., Feng, J.: Learning generalizable and identity-discriminative representations for face anti-spoofing. ACM Trans. Intell. Syst. Technol. **11**(5), 1–19 (2019)
66. Sun, W., Song, Y., Chen, C., Huang, J., Kot, A.C.: Face spoofing detection based on local ternary label supervision in fully convolutional networks. IEEE Trans. Inf. Forensics Secur. **15**, 3181–3196 (2020)
67. Wang, G., Han, H., Shan, S., Chen, X.: Unsupervised adversarial domain adaptation for cross-domain face presentation attack detection. IEEE Trans. Inf. Forensics Secur. **16**, 56–69 (2020)
68. Simonyan, K. and Zisserman, A. Very deep convolutional newtorks for large-scale image recognition. In ICLR, San Diego, CA (2015)
69. Chollet, F.: Xception: deep learning with depthwise separable convolutions. arXiv:1610.02357v3 [cs.CV]
70. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V. and Rabinovich, A.: Going deeper with convolutions. arXiv:1409.4842v1 [cs.CV], (2014)

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.

Environmental Science and Pollution Research
https://doi.org/10.1007/s11356-023-25561-0

RESEARCH ARTICLE

Analyzing the barriers for aquaponics adoption using integrated BWM and fuzzy DEMATEL approach in Indian context

Girish Kumar[1] · Ram C. Bhujel[2] · AniketAggarwal[1] · Divyansh Gupta[1] · Ashish Yadav[1] · Mohammad Asjad[3]

Received: 19 August 2022 / Accepted: 22 January 2023
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2023

Abstract

Aquaponic system in greenhouses which can recycle and reuse the water and nutrients is gaining importance across the world to counter the uncertainties due to weather fluctuations. However, there is a slow pace of growth in aquaculture practices around the globe in general and India in particular. There are many barriers to adopt the aquaponic culture. In this study an analysis of the barriers for aquaponics culture in Indian context during the COVID-19 period is presented. Literature review and interactions with various stakeholders help to find out the list of potential factors while gauging the success of their prospective aquaponics project. The "best-worst" methodology (BWM) is employed for ranking of barriers, whereas categorizing of barriers is carried out with the help of fuzzy DEMATEL. Furthermore, the results of this research work are of great value to corporations or start-up companies looking to invest in this technology as well as to farmers who wish to adopt this farming technique.

Keyword India · Sustainable farming Aquaponics · Barriers · Fuzzy DEMATEL

Introduction

One of the great lessons COVID-19 has taught us is that food security should be the most important priority at all levels: at family, local, state, and country levels. No one should depend entirely on others when crisis of that magnitude occurs. COVID-19 pandemic has taken as wake-up call to rethink and redesign our food systems (Bhavani and Gopinath 2020; Singh et al. 2021). It has been made clear that food production systems should be robust so that food is produced adequately and people have access to it during the time of crisis. During the last few decades when globalization occurred at a rapid pace, many countries were rushing on certain areas which gave better revenues through trade. They were shifting to other businesses from the food production sector or agriculture considering it as one of the least profitable sectors. Even the agriculture became more export-based. It brought big changes in the world including rural areas. Many people left the villages stopping the farming to migrate to cities where they could find jobs and enjoy better life. Some countries put ban on selling agriculture products (Jain et al. 2020). More than that COVID-19 resulted in millions of job losses. Most of those who lost their jobs have gone back to do the farming realizing that agriculture is the ultimate destination when the situation gets worse.

Responsible Editor: Philippe Garrigues

✉ Mohammad Asjad
asjad_alig@rediffmail.com

Girish Kumar
girish.kumar154@gmail.com

Ram C. Bhujel
bhujel@ait.asia

AniketAggarwal
aniketagg1@gmail.com

Divyansh Gupta
divyansh0113@gmail.com

Ashish Yadav
ashishgeorgian3657@gmail.com

[1] Department of Mechanical Engineering, Delhi Technological University, Delhi, India

[2] School of Environment, Resources and Development, Asian Institute of Technology, Bangkok, Thailand

[3] Department of Mechanical Engineering, Faculty of Engineering and Technology, Jamia Millia Islamia, New Delhi, India

Published online: 07 February 2023

India has always been a country where its majority (58%) of population depend on agriculture for their livelihoods (IBEF 2020). The growth rate and gross value added (GVA) by agriculture and allied sectors had improved from − 0.2% in 2015 to + 6.3% in 2017 with some improvement (+2.9%) in 2019 (Government of India 2019). This variation is due to the dependence on the monsoon, inefficient irrigation, injudicious and uncontrolled use of soil nutrients resulting in loss of fertility of soil, uneven access to modern technology in different parts of the country, and various other factors (Kala et al. 2018).

Despite all the rich farming culture and produce, Indian agriculture still has scope for improvement in terms of sustainable production. There is a need for promoting modern technologies and reforming agricultural research and extension as there was underfunding of infrastructure and operations, and limited access to state-of-the-art technologies in the past (World Bank 2012). The adaptation of a modern agriculture technique such as aquaponics can help in achieving the sustainable development goal.

Aquaponics in the simplest form is described as the fusion of two leading modern farming techniques, namely *Aquaculture* and *Hydroponics*. Aquaculture is the cultivation of fish and other organisms in a controlled environment while hydroponics is a soil-less farming method where crops are grown on water itself (Somerville et al. 2014). In aquaponics, the undigested and uneaten food of fishes mostly accumulated as ammoniacal nitrogen provide necessary nitrogen to plants for their growth after bacteria convert it to nitrates, and the plants perform their role by absorbing nitrates from the water, thus making it safe and clean enough for fishes to grow in. The water is re-circulated back to the fish tanks and the cycle repeats again and again. Therefore, aquaponics is one of the best examples of micro unit of a natural world wherein relationship among human, animals, plants, and microbes persists; importance of it has been recently realized and highlighted due to COVID-19 (Altieri and Nicholls 2020).

The soil-less feature of aquaponics not only reduces the dependence on availability of rich fertile soil for cultivation, but eliminates the limitation of agriculture being performed on land altogether. Thus, agricultural practices no longer need to be restricted on soil as now they can be performed easily on rooftops, basements, etc. This practice reduces the need for fertilizers or manure as plants obtain their nutrition directly from the fish excretes. The approximate decrease of about 90% of water requirement not only removes the dependence on monsoon for agriculture but opens up new avenues for performing agriculture in areas prone to droughts or areas with depleting groundwater levels (Simanovski and Pirkebner 2018).

To mitigate the adverse effects of nature, aquaponics has been rising in many parts of India like in Cherai, a village in Kochi, Kerala (Karthika 2018), and outside India like Wellington, New Zealand (WWF 2013). In Kerala, many farmers have paired the technology with rooftop solar panels to ensure continuous power supply and to utilize the full potential of the integrated system. Even though the Indian project started just 4 years back, many farmers have been successful in growing vegetables in hundred bags by using 14,000-L fish tanks, which contain more than 1500 fish. This technique has benefited farmers and fishermen alike. The practice of aquaponics has various incentives including organic-like produce, soil-less nature, zero fertilizer use, higher control, being extremely water-efficient and making farming possible on non-arable small piece of land. Because of its ability to be installed in different landscapes especially in backyard gardens, numerous aquaponics setups have also been installed on rooftops in the West Bank and Gaza Strip, by the Food and Agriculture Organization of the United Nations (Somerville et al. 2014). These setups are in effect to tackle the chronic food and nutrition security issues seen across the region. It became so attractive to individual families during lock-downs due to COVID-19.

Aquaponics technology finds its origins during the times of the Aztec Indians, but as a modern technology, it is still in the research and development phase. There have been many studies and researches about how to improve the crop yield, how to incorporate new methodologies in aquaponics, and other technological developments (Monsees et al. 2017; Yang and Kim 2019). One such study advocates the introduction of a vital index like nitrogen utilization efficiency (NUE), for the assessment of aquaponics and the improvement of the system through micro- and macro-nutrient addition (Ru et al. 2017). There also have been some comparative studies between aquaponics and conventional farming, stating that increased productivity and water use efficiency are the key advantages of modern farming technologies like aquaponics (Alshrouf 2017). There are studies in different parts of the world which compare the final produce of the different agricultural technologies to dismiss the safety concerns about fish and plant cycle integration and conclude that aquaponics can be considered a strong alternative to conventional farming (El-Essawy et al. 2019; Rosgren and Grahler 2022). A vision document prepared for Washington DC and Netherland emphasized to capture the full potential of a symbiotic that effectively integrates the value of nature into the urban scape and its social, built, and geographic characteristics (Stuiver and O'hara 2022).

Beyond the hype of the technology, there are several challenges associated with the implementation of aquaponics such as lack of knowledge and expertise, lack of capital investment, poor pest and disease management, and others (Love et al. 2014; Turnsek et al. 2020; Yep and Zheng 2019). Some factors are more important than others in different climatic, geographical, and socio-economic contexts.

Technological, economic, and social considerations are crucial in setting up, wide-scale adoption, and in-turn success of the aquaponics technology. Many studies have focused on the diversity of these fields that need to be addressed including various technological, socio-economic, and system design trends (Goddek et al. 2015; Junge etal. 2017; Turnsek et al. 2020)

However, there is a lack of qualitative research on the barriers to the adoption of aquaponics culture across the India and especially during the COVID-19, as the people were not much aware about its exhaustive application. Thus, there was the potential to identify the barrier and their remedial action for efficient and effective utilization of aquaponics environment. From the literature review, it can be concluded that the previous researches were mainly focused on the technical aspects of the aquaponics setup, less on the economic front, and even lesser considerations of the social aspects. Moreover, those studies were not exhaustive in listing the challenges and they did not employ any scientific decision-making tool to rank and categorize them. The present study focused on the barriers that are significant and that should be considered while setting up an aquaponics unit. This research aimed at identifying numerous challenges under the domain of economic, technological, social, educational, etc. that are crucial in determining the success of aquaponics in Indian context. More importantly, all these factors have been ranked according to their importance or the priority of their consideration, using a best-worst method (BWM). The factors are also categorized into cause-and-effect groups using fuzzy DEMATEL approach. Addressing all the mentioned challenges, the incorporation of aquaponics into the current farming scenario would see a phenomenal rise in both efficiency and productivity of the agricultural sector. Thus, this setup provides a sustainable modern farming technique that shall boost the agricultural sector to its highest potential.

This research paper presents the current research scenario of aquaponics, its research gaps and challenges identified through literature review and expert interviews. It also describes fuzzy DEMATEL tool which uses best-worst method of ranking and helps compare among the categories. Finally, it concludes and recommends important areas for future research.

Barriers in adoption of aquaponics

The following list of barriers have been created by extensive literature review and expert interviews.

Lack of knowledge and expertise

Aquaponics requires a symbiotic environment with appropriate levels of pH, temperature, oxygen levels, etc. in accordance with the life forms of animals and plants. This requires high-level expertise not just in the field of farming and fish culture but also in the fields of basic sciences and biological systems. This proves to be a significant barrier in India due to lack of knowledge and awareness among various stakeholders, especially farmers and extension workers. Studies conducted even in developed counties such as Canada and UK showed that there is lack of knowledge and expertise and it serves as a significant barrier to implementation of aquaponics (Matthews 2017; Cammies 2021). Although highly qualified people are involved in aquaponics in European countries as it is evident from a survey conducted in Europe on current aquaponics systems showed that 91.7% of the people involved in aquaponics hold at least post-graduate degree (Villarroel et al. 2016).

Absence of stable environment

There are different climatic conditions which are specific to the species grown, their age, size, technology, etc. that determine the success of any aquaponic system. Whatever these conditions may be, they vary to a great extent by season throughout the year and these fluctuations might have a considerable impact on the health and growth of animals and plants used in aquaponics system (Goddek et al. 2015). Fluctuations in temperature specifically have a tremendous impact on the fish, plant, and bacteria cycles and the nitrification process as well (Zhu and Chen 2002). In a study conducted in Brazil, it was observed that absence of stable environment is one of the barriers to the success of aquaponics (Brewer et al. 2021).

High capital requirements

The initial investments and operating expenses of initiating and running an aquaponic unit are also barriers to the success of the technology (Matthews 2017). Despite considerable research in the area, it is still uncertain whether the aquaponics is economically profitable (Greenfeld et al. 2019). Though studies are being conducted to explore more cost-effective ways to implement aquaponics, they are still a long way to go to become a good option for commercial ventures (Sunny et al. 2019).

Nutrient limitation in fish excreta

In a closed loop aquaponics system, plants derive almost all of their nutrients from the fish excreta. However, certain essential nutrients like Fe, Ca, Mg, and K which are required for the plant growth (Njinga et al. 2013) are limited in the fish excreta and fish meals since fishes have minimal usage of these metals (Savidov et al. 2007). Synthetic fertilizers may be used to compensate these deficiencies but aquaponic

systems rely little on them (Yep and Zheng 2019). Surveys show that some aquaponics practitioners also have problems in understanding and managing these nutrient deficiencies (Matthews 2017).

Maintaining pH

The life forms in the aquaponics unit, i.e., fishes, plants, and microbes, have different optimal pH ranges (Somerville et al. 2014). Given the difference in pH ranges, there is no single value of pH that would ensure optimal growth of all the involved species. There is a constant influx of acid (H+) from the nitrification of ammonia and of hydroxide (OH−) or bicarbonate from most plants, both of which keep on trying to shift the pH from the decided value (Yep and Zheng 2019). The working pH is based on trade-off between the growth of one organism at the cost of the other organism and thus has to be carefully decided and maintained (Yep and Zheng 2019).

Poor pest and disease management

The pest and disease management front is another aspect that poses as a barrier for the aquaponic systems (Vermeulen and Kamstra 2013). There is a need for innovative pest and disease management solutions which do not disturb the balance of the cycle. The pest prevention solutions have to be in accordance with the fishes so as to not harm them and the antibiotics have to be suitable for the plants for similar reasons (Goddek et al. 2015).

Interrupted power supply

Aquaponics setup requires continuous electricity for aeration, water pumps, and possibly temperature regulation systems (Matthews 2017). A survey of aquaponic practitioners in the USA and internationally shows that about 95% of the respondents relied on power from the main supply grid but because of its unreliability, about 57% of the respondents used alternate sources of renewable energy, e.g., solar cells (Love et al. 2014). This adds to which is already high capital investment of the aquaponics setup and thus restricts its adoption and expansion.

Limited plant and fish combinations

In an aquaponic setup with multiple living organisms, it is difficult to set one value for each water quality parameter that would suit all the organisms living within the system (Estim et al. 2020). Thus, these parameters have to be maintained within the tolerance levels of each organism (Estim et al. 2020). This leads to sub-optimal conditions for individual organisms but aims to optimize the overall harvest. Even today, finding appropriate fish and plant combinations that would have an optimal yield while trading off and growing at these sub-optimal conditions is a key challenge (König et al. 2018). This problem is enhanced further when market restraints are applied as the pool of crops that could be grown in the system is further narrowed.

Limited options of crops to be produced

Aquaponics as a technology is capable of producing numerous types of crops commercially. But practically it is restricted to producing only few high- and middle-value crops like tomatoes and lettuces. This is so, because the returns in producing low-value crops like potatoes fall short to cover the cost incurred to establish the aquaponic setup per unit space. Since the low-value crops form a significant part of the diet, the inability to produce them profitably accompanied by availability of cheap counterparts, limits the market of aquaponic products thereby the possibility of adoption and growth (Mukherjee 2019; Turnšek et al. 2019). The summary of barriers identified based on literature review is presented in Table 1.

Studies pertaining to the applicability of aquaponics and the barriers that are associated with it are not exhaustive. They fall short to account for various technical and sociological parameters which are significant in determining the success of aquaponics. Moreover, no current research has conducted scientific studies that employs decision-making tools to categorize the challenges in implementing the aquaponics in Indian context.

Methodology

The challenges in implementation of aquaponics in Indian context are identified on the basis of extensive literature review and discussion with the domain experts. In further analysis, best-worst method (BWM) developed by Razaei (2015) is used for prioritizations of the barriers and fuzzy DEMATEL is employed for further categorization of barriers into cause and effect groups.

There are many Multi Criteria Decision Making (MCDM) methods available in literature for ranking such as Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) (Esfandiari and Rizvandi 2014), Analytic Hierarchy Process (AHP) (Cheng et al. 1999), and Grey Weighted Sum Model (GWSM) (Esangbedo and Che 2016). However, BWM is chosen for this study as it requires less comparison data and provides a more consistent result (Rezaei 2015), which not only reduces the number of pairwise comparisons and inconsistency in such a task (Labella et al. 2021) but it also performs better than other multi-attribute decision-making models (Bai 2018.)

Table 1 List of barriers in adoption of aquaponics based on the literature survey

Code	Name of barrier	Description	Country	Source
B1	Lack of knowledge and expertise	Educated and expert labor is required to run the system and maintain a balance of all water quality parameters.	Canada	(Matthews 2017)
B2	Absence of stable environment	Fluctuating environmental conditions will impact the system negatively. Stable environment is preferred.	United States	(Zhu and Chen 2002)
B3	High capital requirements	Huge initial investments and ongoing expenditures make the projects risky.	Canada	(Matthews 2017)
B4	Nutrient limitation in fish excreta	Plants cannot always derive all the nutrients from fish excreta; external help might be required.	Berlin, Spain, Switzerland	(König et al. 2018)
B5	Maintaining pH	Maintaining one pH value which is favorable for all organisms in the system is challenging	Canada	(Yep and Zheng 2019)
B6	Poor pest and disease management	Conventional pest and disease management techniques do not work in the aquaponics setup. New techniques need to be developed and implemented.	Germany, Belgium, Iceland	(Goddek et al. 2015)
B7	Interrupted power supply	Continuous electricity supply at all times is difficult to ensure. Fail-safes and alternate sources add to the already high capital expenditures.	Canada	(Matthews 2017)
B8	Limited plant and fish combinations	Finding a combination of plants and fishes that have a common range for all water quality parameters for their optimal growth is difficult to find.	Berlin, Spain, Switzerland	(König et al. 2018)
B9	Limited options of crops to be produced	Not all market crops can be grown economically through aquaponics.	India	Self-developed

Commonly used methods for categorization of factors are Analytic Network Process (ANP) (Lee et al. 2013), Interpretive Structural Modelling (ISM) (Al-Muftah et al. 2018; Nagpal et al. 2017; Pitchaimuthu et al. 2019), and Interpolative Boolean Algebra (IBA) (Mandic and Delibasic 2014). The DEMATEL method not only converts the interdependency relationships into a cause and effect group via matrixes but also finds the critical factors of a complex structure system with the help of an impact relation diagram (Si et al. 2018). However, fuzzy DEMATEL is selected for this work as it ranks factors and also finds out the critical evaluation criteria as well as the mutual influence of various factors on each other (Si et al. 2018).

The procedures for BWM and fuzzy DEMATEL are detailed in the following:

Best-worst method (BWM)

Best-worst method determines weights of factors in reduced number of comparisons. The factor with the most vital role is considered most important which becomes evident on the basis of weights determined by BWM method.

Steps involved in BWM:

Step 1: Criteria determination

In the present study there are 9 barriers, denoted by B_1, B_2, B_3……. B_9. Select one best criteria (most desirable/most important) and one worst criteria (least desirable/least important) based on the opinion of industry experts.

Step 2: Comparing best and worst criteria

Assign preference values (preference is indicated by a number 1 to 9, where 1 denotes equal importance) after comparing the most important barrier with all barriers denoted by p_{m1}, p_{m2}, p_{m3}, ….p_{m9}. Similarly compare the least important barrier with all barriers and assign a quantitative preference value to each barrier denoted as p_{1l}, p_{2l}, p_{3l}, …. p_9.

Step 3: Optimal weight calculation

Let weights of barriers B_1, B_2, B_3……. B_9 be denoted by α_1, α_2, α_3…… α_9. After assigning preference values, obtain optimal weights by solving the linear programming model from Eqs. (1)–(4) (Rezaei 2015)

Objective: Minimize β

Subject to:

$$\left|\frac{\alpha_m}{\alpha_j} - p_{mj}\right| \le \beta \qquad j = 1, 2, 3 \ldots, 9 \tag{1}$$

α_m : denotes the weight of the most important factor as selected in step 1.

$$\left|\frac{\alpha_j}{\alpha_l} - p_{jl}\right| \leq \beta \qquad j = 1, 2, 3. \ldots ., 9 \quad (2)$$

α_l : denotes the weight of least important factor as selected in step 2.

$$\sum_{j=1}^{9} \alpha_j = 1 \qquad j = 1, 2, 3 \ldots .9 \quad (3)$$

(Equation 3 is the expression denoting the sum of weights is equal to unity.)

$$\alpha_j \geq 0 \qquad j = 1, 2, 3 \ldots .9 \quad (4)$$

Step 4: Solving equations

On solving the linear equations, we get the weights for all 9 barriers on the basis of which they are ranked.

The fuzzy DEMATEL method

The Decision-Making Trial and Evaluation Laboratory (DEMATEL) method was first introduced by the Geneva Research Centre of the Battelle Memorial Institute to visualize the structure of complicated causal relationships through matrices or digraphs (Gabus and Fontela 1973). Fuzzy DEMATEL method is employed in order to visualize the problem within a fuzzy environment. In the present study we use the fuzzy DEMATEL method to categorize the barriers in implementation of aquaponics into two categories, the cause group and the effect group. It helps in determining the relative importance of the factors and thereby ranking/prioritizing them.

DEMATEL method is used to understand the relationship between various factors and analyze how they influence each other. This method of ranking had been used previously in many domains such as improvising emergency systems (Han and Deng 2018), barriers to coastal shipping development (Venkatesh et al. 2017), remanufacturing industry (Bhatia and Srivastava 2018), supplier selection (Chang et al. 2011), and safety management system for airlines (Liou et al. 2008). In this method directed graphs are employed, which help in separating the factors into two groups, the cause group and the effect group. DEMATEL method also aids the making of causal diagram which helps in visualizing the groups and their influence on other factors.

Fuzzy logic was proposed by Zadeh (1965). He introduced the concept of fuzzy set theory and the concept of membership function (Zadeh 1965). A membership function defines the degree of truth in the logic. The membership function in fuzzy logic plays a vital role in selecting the best alternative among the feasible one, when applied to any research problem. However, in literature, a number of membership functions are reported with their application, advantages, and limitations, among which the response of triangular membership function out performed to other membership functions (Zhao and Bose 2002). Thus, in this study, a triangular membership function is selected for further analysis (Fig. 1). In order to obtain crisp values, defuzzification is done using the center of area (COA) method also known as the centroid method.

The procedure of fuzzy DEMATEL method uses reviews of domain experts to form the initial matrix (Lin and Wu 2008). All the experts were presented with 9 factors/challenges in implementation of aquaponics in India. The challenges are denoted as B_1, B_2, B_3……. B_9. Each expert graded the factors based on relative importance between 0 and 4 as denoted in Table 31 (Appendix 3).

Step 1: Direct relation matrix (D_m)

A direct relation matrix (D_m) is formed on the basis of influence score given by domain experts.

Step 2: Transform in triangular fuzzy numbers

The influence scores in the direct relation matrix (D_m) are replaced by the corresponding triangular fuzzy numbers as given in Table 2.

Step 3: Defuzzification of matrix (Z_m)

In order to convert the triangular fuzzy numbers to crisp values, the given matrix is now defuzzified by centroid method using Eq. (5) (Liou et al. 2008; Si et al. 2018).

$$\left(CV_{ij}\right) = \frac{S_{ij} + M_{ij} + L_{ij}}{3} \mathrm{i} = 1, 2, 3 \ldots ..9; \mathrm{j} = 1, 2, 3 \ldots ..9 \quad (5)$$

CV_{ij} : denotes the crisp value for the particular cell i, j.

S_{ij} : denotes the smallest likely value of the particular cell i, j.

M_{ij} : denotes the most probable value of the particular call i, j.

L_{ij} : denotes the largest possible value of the particular cell i, j.

Step 4: Form a single matrix (S_m)

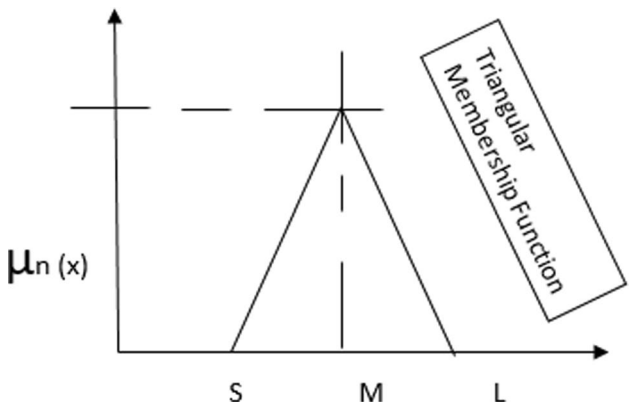


Fig. 1 Triangular membership function. $\mu_{n(x)}$, membership function; *S*, smallest likely value; *M*, most probable value; *L*, largest possible value

Table 2 Comparison of most important criterion with all criteria

	B_1	B_2	B_3	B_4	B_5	B_6	B_7	B_8	B_9
B_1	1	3	2	5	6	8	7	4	9

Until now individual matrices of experts were formulated separately, but now one single matrix (S_m) is obtained by averaging corresponding cells of all the matrices using Eqs. (6)–(8) (Lin and Wu 2008).

$$S_m = \frac{Z_{mij}^{<1>} + Z_{mij}^{<2>} + Z_{mij}^{<3>} + \cdots + Z_{mij}^{<N>}}{N_e} \tag{6}$$

$Z_{mij}^{<N>}$: denotes the value crisp value in the cell i, j of the defuzzified matrix of N^{th} expert.

$$S_m = \left[NV_{11}NV_{12}\cdots NV_{19} \vdots \ddots \vdots NV_{91}NV_{92}\cdots NV_{99}\right] \tag{7}$$

$$NV_{ij} = \frac{\sum_{g=1}^{N_e} OV_{ij}^{<g>}}{N_e} \quad \mathrm{i} = 1,2,3\ldots.9;\ \mathrm{j} = 1,2,3\ldots..9 \tag{8}$$

NV_{ij} : denotes the new crisp value in the cell i, j in the single matrix.

$OV_{ij}^{<N>}$: denotes the crisp value in the cell i, j corresponding to N^{th} matrix.

N_e : denotes the number of experts.

Step 5: Normalized direct relation matrix (G_m)

The single direct relation matrix (S_m) is normalized/generalized using Eqs. (9) and (10) (Lin and Wu 2008).

$$G_m = X * S_m \tag{9}$$

$$X = \frac{1}{\sum_{j=1}^{9} NV_{ij}} \mathrm{i} = 1,2,3\ldots.9;\mathrm{j} = 1,2,3\ldots..9 \tag{10}$$

Step 6: Total relation matrix (Y_m)

Normalized direct relation matrix (G_m) is formulated using Eq. (11) in order to obtain the total relation matrix (Y_m) (Lin and Wu 2008).

$$Y_m = G_m * \left(I_m - G_m\right)^{-1} \tag{11}$$

I_m : denotes the Identity matrix.

Step 7: Obtain sum of rows and columns:

Sum of rows (U_i) is calculated using Eq. (12), and similarly, sum of columns (V_j) is calculated using Eq. (13) (Wang and Chen 2012).

$$U_i = \sum_{j=1}^{9} Y_{mij} \mathrm{i} = 1,2,3\ldots.9;\ \mathrm{j} = 1,2,3\ldots..9 \tag{12}$$

$$V_j = \sum_{i=1}^{9} Y_{mij}\ \mathrm{i} = 1,2,3\ldots.9;\ \mathrm{j} = 1,2,3\ldots..9 \tag{13}$$

Y_{mij} : denotes the value in the cell i, j of the Total relation matrix (Y_m).

Step 8: Causal diagram

(U_i+V_j) and (U_i−V_j) are calculated for equal values of i and j followed by a graphical depiction with (U_i+V_j) on the X-axis and (U_i−V_j) on the Y-axis, known as the causal diagram (Lin and Wu 2008; Liou et al. 2008).

A case illustration

In this research work, a case of Indian context is considered and the data were collected during the COVID-19. This section is categorized into three sub-section, viz, "Data collection", "Ranking by BWM", and "Categorization by fuzzy DEMATEL" respectively.

Data collection

The questionnaires for ranking and cause-effect categorization of the barriers for aquaponics adoption using integrated BWM and fuzzy DEMATEL approach (Appendix 1 and Appendix 3) were finalized after a review by six experts (Appendix 1). Out of the six experts selected for this study, two are from academia, two from the industry, and two from not-for-profit organization. Their inputs helped to figure out one important barrier for this study, which resulted in a final list of nine barriers. The experts also helped in validation of the literature review findings and also facilitated framing the questionnaire more specific to the Indian context. They commented on the intelligibility, subject, and illustration of the survey questionnaire. Suggestions were incorporated and improvements in the questionnaire were completed before distributing for data collection. The final questionnaires consist of all the nine significant challenges identified through extensive literature review and inputs from experts.

The data for questionnaires (Appendix 2: Table 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28 and 29) was collected from six experts (Appendix 1) and used for BWM method. Pairwise comparison data was obtained on a scale of 1 to 9 for comparison of best to others and worst to others.

The data for questionnaire (Appendix 3: Table 30, 31 and 32) was obtained though survey method. Google form was chosen as a medium to collect data from professionals who are linked directly or indirectly to aquaponics industry. Initially, 200 emails were sent to the professionals who are working in operations, strategy, marketing, and environment domain. Out of the 200 emails sent, approximately 100 responses were received. Half

of the emails received were either incomplete or not appropriate for the purpose of analysis. Overall, 50 responses received were carefully analyzed and checked for anomalies/errors, and minor corrections were made after checking with the respondent. The professionals were asked to rate these challenges on a 5-point Likert scale (e.g., 1 = no influence and 5 = extremely high influence) illustrating the influence of each challenge on setting up of an aquaponics setup in India. The average of the valid responses was used to generate the pairwise comparison matrix of the selected barriers. After processing all the responses, the fuzzy DEMATEL approach was used for cause-effect categorization of barriers.

Ranking by BWM

In the present study, nine barriers, as identified from the literature and through expert feedback, are being utilized in the implementation of aquaponics in Indian context. After receiving the expert's feedback, the proposed best-worst method was applied for calculating the weights of respective barriers. Table 2 and Table 3 represent the input from all 6 experts for barriers B1 and B9 respectively. In the similar manner, the comparison was made in respect of other barriers. The values represent the average of all the responses rounded off to nearest integer. This data is further utilized for constructing the matrix in Table 4.

Now a linear programming model is formed using Eqs. (1) to (4). The equations obtained (1a) to (4a) are then solved in order to obtain the weights of respective barriers.

$$\left|\frac{\alpha_1}{\alpha_2} - 3\right| \le \beta \tag{1a}$$

$$\left|\frac{\alpha_1}{\alpha_3} - 2\right| \le \beta \tag{1b}$$

$$\left|\frac{\alpha_1}{\alpha_4} - 5\right| \le \beta \tag{1c}$$

$$\left|\frac{\alpha_1}{\alpha_5} - 6\right| \le \beta \tag{1d}$$

$$\left|\frac{\alpha_1}{\alpha_6} - 8\right| \le \beta \tag{1e}$$

$$\left|\frac{\alpha_1}{\alpha_7} - 7\right| \le \beta \tag{1f}$$

$$\left|\frac{\alpha_1}{\alpha_8} - 4\right| \le \beta \tag{1g}$$

$$\left|\frac{\alpha_1}{\alpha_9} - 9\right| \le \beta \tag{1h}$$

$$\left|\frac{\alpha_2}{\alpha_9} - 7\right| \le \beta \tag{2a}$$

$$\left|\frac{\alpha_3}{\alpha_9} - 8\right| \le \beta \tag{2b}$$

$$\left|\frac{\alpha_4}{\alpha_9} - 5\right| \le \beta \tag{2c}$$

$$\left|\frac{\alpha_5}{\alpha_9} - 3\right| \le \beta \tag{2d}$$

$$\left|\frac{\alpha_6}{\alpha_9} - 2\right| \le \beta \tag{2e}$$

$$\left|\frac{\alpha_7}{\alpha_9} - 3\right| \le \beta \tag{2f}$$

$$\left|\frac{\alpha_8}{\alpha_9} - 6\right| \le \beta \tag{2g}$$

$$\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 + \alpha_6 + \alpha_7 + \alpha_8 + \alpha_9 = 1 \tag{3a}$$

Table 3 Comparison of least important criterion with all criteria

	B_9
B_1	9
B_2	7
B_3	8
B_4	5
B_5	3
B_6	2
B_7	3
B_8	6
B_9	1

Table 4 Weights and ranking of criteria

Criteria	B_1 α_1	B_2 α_2	B_3 α_3	B_4 α_4	B_5 α_5	B_6 α_6	B_7 α_7	B_8 α_8	B_9 α_9
Weights	0.3146	0.1277	0.1915	0.0766	0.0638	0.0479	0.0547	0.0957	0.0273
Ranking	1	3	2	5	6	8	7	4	9

$$\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6, \alpha_7, \alpha_8, \alpha_9 \geq 0 \quad (4a)$$

On solving the above equations, the weights are obtained and are further used to rank the barriers as shown in Table 4.

Table 4 depicts that the Lack of knowledge and expertise (B_1) has the highest weight (α_1) 0.3146 while Limited option of crops to be produced (B_9) has the lowest weight (α_9) 0.0273. Since $\alpha_1 > \alpha_3 > \alpha_2 > \alpha_8 > \alpha_4 > \alpha_5 > \alpha_7 > \alpha_6 > \alpha_9$, therefore $B_1 > B_3 > B_2 > B_8 > B_4 > B_5 > B_7 > B_6 > B_9$ are the ranks of importance.

Categorization by fuzzy DEMATEL

After ranking the barriers, fuzzy DEMATEL is applied for categorizing the barriers into cause-and-effect group. The procedural steps are applied as explained in the "Methodology" section, initially a questionnaire was prepared with the help of experts (Appendix 1), and survey method was employed for the data collection to prepare the direct-relation matrix containing the average influence scores given by 50 survey respondents. The values represent the average of all the responses rounded off to nearest integer. This data is further utilized for constructing the matrix in Table 5.

Now the influence scores in Table 5 are replaced by respective triangular fuzzy numbers using Table 31 (Appendix 3) as shown in Table 6.

The next step is to defuzzify the values in Table 6 and convert them into crisp values. This defuzzification is done by applying the centroid method using Eq. (5). The crisp values obtained after defuzzification are shown in Table 7.

Now a single matrix is formed using all the processed matrices of the domain-experts (i.e., Six, as mentioned in Appendix 1); this formation of single matrix is done using Eqs. (6), (7), and (8). This single matrix is shown in Table 8.

The next step is to normalize/generalize the single matrix in Table 8, using Eqs. (9) and (10). This results in normalized direct relation matrix as shown in Table 9.

Normalized total relation matrix is formulated using Eq. (11) and total relation matrix is obtained as shown in Table 10.

From the total relation matrix, sum of rows (U_i) and sum of columns (Vj) are calculated using Eqs. (12) and (13). The next step is to form the causal diagram by plotting ($U_i + V_j$) on the X-axis and ($U_i - V_j$) on the Y-axis as shown in Fig. 2. The degree of central role is shown in Table 11.

The barriers were arranged and ranked on the basis of their respective weights as shown in Table 4. Lack of knowledge and expertise (B_1) has the highest weight (α_1) 0.3146 while Limited option of crops to be produced (B_9) has the lowest weight (α_9) 0.0273. Since $\alpha_1 > \alpha_3 > \alpha_2 > \alpha_8 > \alpha_4 > \alpha_5 > \alpha_7 > \alpha_6 > \alpha_9$, therefore $B_1 > B_3 > B_2 > B_8 > B_4 > B_5 > B_7 > B_6 > B_9$ are the ranks of importance.

These results are consistent with the existing literature wherein lack of knowledge and expertise and high capital requirements are two of the most prominent challenges in setting up an aquaponics unit (Brewer 2019; Greenfeld et al. 2020). Farmers with lack of knowledge and lack of information about aquaponics accounted for a huge proportion of farmers, out of which about half had financial concerns about initial capital requirements as well (Brewer 2019). Absence of stable environment, limited fish and plant combinations restricting the expansion of production, and other technical challenges constitute the next prominent set of challenges in setting up an aquaponics unit (Brewer 2019; Goddek et al. 2015).

According to Table 11, the value of ($U_i + V_j$) denotes the importance of the barrier. Lack of knowledge and expertise (B_1) has the highest ($U_i + V_j$) score of 7.0817, followed by $B_3 > B_2 > B_8 > B_4 > B_5 > B_7 > B_6 > B_9$. Using the fuzzy DEMATEL method, the barriers are categorized into cause-and-effect groups using the ($U_i - V_j$) score. On the basis of ($U_i - V_j$) score, the evaluation barriers, namely, High capital requirements (B_3), Maintaining pH (B_5), Interrupted power supply (B_7), Limited plant and fish combinations (B_8), and Limited options of crops to be produced (B_9), fall under the category of cause group, whereas Lack of knowledge and expertise (B_1), Absence of stable environment (B_2), Nutrient limitation in fish excreta (B_4), and Poor pest and disease management (B_6) fall under the category of effect group.

The cause group factors have an impact on many other factors. Interrupted power supply (B_7) has the highest ($U_i - V_j$) value of 2.5395; hence, barrier B_7 has more impact on the

Table 5 Direct-relation matrix (Dm)

	B_1	B_2	B_3	B_4	B_5	B_6	B_7	B_8	B_9
B_1	0	4	4	2	1	0	0	1	1
B_2	4	0	2	1	0	2	0	0	0
B_3	4	3	0	2	1	3	0	4	3
B_4	3	3	1	0	1	1	0	2	1
B_5	3	3	2	3	0	3	0	1	1
B_6	1	1	1	1	0	0	2	2	0
B_7	4	4	3	3	3	2	0	2	0
B_8	4	3	3	2	2	2	0	0	0
B_9	3	4	2	0	0	2	0	0	0

Table 6 Triangular fuzzy numbers

	B_1	B_2	B_3	B_4	B_5	B_6	B_7	B_8	B_9
B_1	0	(0.6, 0.75, 1)	(0.6, 0.75, 1)	(0.25, 0.5, 0.6)	(0.1, 0.25, 0.5)	(0, 0.1, 0.25)	(0, 0.1, 0.25)	(0.1, 0.25, 0.5)	(0.1, 0.25, 0.5)
B_2	(0.6, 0.75, 1)	0	(0.25, 0.5, 0.6)	(0.1, 0.25, 0.5)	(0, 0.1, 0.25)	(0.25, 0.5, 0.6)	(0, 0.1, 0.25)	(0, 0.1, 0.25)	(0, 0.1, 0.25)
B_3	(0.6, 0.75, 1)	(0.5, 0.6, 0.75)	0	(0.25, 0.5, 0.6)	(0.1, 0.25, 0.5)	(0.5, 0.6, 0.75)	(0, 0.1, 0.25)	(0.6, 0.75, 1)	(0.5, 0.6, 0.75)
B_4	(0.5, 0.6, 0.75)	(0.5, 0.6, 0.75)	(0.1, 0.25, 0.5)	0	(0.1, 0.25, 0.5)	(0.1, 0.25, 0.5)	(0, 0.1, 0.25)	(0.25, 0.5, 0.6)	(0.1, 0.25, 0.5)
B_5	(0.5, 0.6, 0.75)	(0.5, 0.6, 0.75)	(0.25, 0.5, 0.6)	(0.5, 0.6, 0.75)	0	(0.5, 0.6, 0.75)	(0, 0.1, 0.25)	(0.1, 0.25, 0.5)	(0.1, 0.25, 0.5)
B_6	(0.1, 0.25, 0.5)	(0.1, 0.25, 0.5)	(0.1, 0.25, 0.5)	(0.1, 0.25, 0.5)	(0, 0.1, 0.25)	0	(0.25, 0.5, 0.6)	(0.25, 0.5, 0.6)	(0, 0.1, 0.25)
B_7	(0.6, 0.75, 1)	(0.6, 0.75, 1)	(0.5, 0.6, 0.75)	(0.5, 0.6, 0.75)	(0.5, 0.6, 0.75)	(0.25, 0.5, 0.6)	0	(0.25, 0.5, 0.6)	(0, 0.1, 0.25)
B_8	(0.6, 0.75, 1)	(0.5, 0.6, 0.75)	(0.5, 0.6, 0.75)	(0.25, 0.5, 0.6)	(0.25, 0.5, 0.6)	(0.25, 0.5, 0.6)	(0, 0.1, 0.25)	0	(0, 0.1, 0.25)
B_9	(0.5, 0.6, 0.75)	(0.6, 0.75, 1)	(0.25, 0.5, 0.6)	(0, 0.1, 0.25)	(0, 0.1, 0.25)	(0.25, 0.5, 0.6)	(0, 0.1, 0.25)	(0, 0.1, 0.25)	0

Table 7 Crisp values matrix (Zm)

	B_1	B_2	B_3	B_4	B_5	B_6	B_7	B_8	B_9
B_1	0	0.7833	0.7833	0.45	0.2833	0.1167	0.1167	0.2833	0.2833
B_2	0.7833	0	0.45	0.2833	0.1167	0.45	0.1167	0.1167	0.1167
B_3	0.7833	0.6167	0	0.45	0.2833	0.6167	0.1167	0.7833	0.6167
B_4	0.6167	0.6167	0.2833	0	0.2833	0.2833	0.1167	0.45	0.2833
B_5	0.6167	0.6167	0.45	0.6167	0	0.6167	0.1167	0.2833	0.2833
B_6	0.2833	0.2833	0.2833	0.2833	0.1167	0	0.45	0.45	0.1167
B_7	0.7833	0.7833	0.6167	0.6167	0.6167	0.45	0	0.45	0.1167
B_8	0.7833	0.6167	0.6167	0.45	0.45	0.45	0.1167	0	0.1167
B_9	0.6167	0.7833	0.45	0.1167	0.1167	0.45	0.1167	0.1167	0

system and other challenges as interrupted power supply would directly affect the capital requirements, ability to sustain limited plant and fish combinations, etc.; however, the low $(U_i + V_j)$ score of 4.9165 for B_7 can be justified by the relative low occurrence of the challenge in aquaponics plants compared to other challenges (El-Sayed 2020). The last barrier in the list of cause group is High capital requirements (B3) with a $(U_i - V_j)$ score of 0.0115 and a high $(U_i + V_j)$ score of 6.7355 giving it a higher ranking in importance, which is consistent with the results from the best-worst method. The barriers categorized as the effect group are influenced by the other barriers. Absence of stable environment (B_2) has the lowest $(U_i - V_j)$ score of −2.1015 implying that barrier B_2 is influenced more compared to other factors. Lack of capital, lack of knowledge, and lack of appropriate plant and fish combination would all negatively affect the ability to provide stable environment for the aquaponics unit.

Conclusions

The present research identified nine criteria or the challenges in the implementation of aquaponics shortlisted considering the social, economic, and technical variables by aggregating the expert's inputs. When they were categorized using

Table 8 Single matrix (Sm)

	B_1	B_2	B_3	B_4	B_5	B_6	B_7	B_8	B_9
B_1	0	0.7833	0.7278	0.45	0.2833	0.1167	0.1167	0.2278	0.2278
B_2	0.7278	0	0.3944	0.2833	0.1722	0.5056	0.1167	0.1722	0.1167
B_3	0.7833	0.5611	0	0.3944	0.1722	0.6167	0.1167	0.7278	0.5611
B_4	0.6722	0.6167	0.3389	0	0.3389	0.2833	0.1167	0.3944	0.3389
B_5	0.6722	0.6167	0.3944	0.5056	0	0.5611	0.1722	0.2278	0.2833
B_6	0.2833	0.2833	0.2278	0.2833	0.1167	0	0.3944	0.3944	0.1167
B_7	0.6722	0.7278	0.6167	0.6167	0.6722	0.3944	0	0.3944	0.1167
B_8	0.7278	0.7278	0.6167	0.45	0.5611	0.3944	0.1167	0	0.1167
B_9	0.6167	0.7278	0.45	0.1167	0.1722	0.45	0.1167	0.1167	0

Table 9 Normalized direct relation matrix (Gm)

	B_1	B_2	B_3	B_4	B_5	B_6	B_7	B_8	B_9
B_1	0	0.1860	0.1728	0.1069	0.0673	0.0277	0.0277	0.0541	0.0541
B_2	0.1728	0	0.0937	0.0673	0.0409	0.1201	0.0277	0.0409	0.0277
B_3	0.1860	0.1332	0	0.0937	0.0409	0.1464	0.0277	0.1728	0.1332
B_4	0.1596	0.1464	0.0805	0	0.0805	0.0673	0.0277	0.0937	0.0805
B_5	0.1596	0.1464	0.0937	0.1201	0	0.1332	0.0409	0.0541	0.0673
B_6	0.0673	0.0673	0.0541	0.0673	0.0277	0	0.0937	0.0937	0.0277
B_7	0.1596	0.1728	0.1464	0.1464	0.1596	0.0937	0	0.0937	0.0277
B_8	0.1728	0.1728	0.1464	0.1069	0.1332	0.0937	0.0277	0	0.0277
B_9	0.1464	0.1728	0.1069	0.0277	0.0409	0.1069	0.0277	0.0277	0

Table 10 Total relation matrix (Ym)

	B_1	B_2	B_3	B_4	B_5	B_6	B_7	B_8	B_9
B_1	0.3587	0.5024	0.4173	0.3120	0.2217	0.2654	0.1182	0.2488	0.2006
B_2	0.4406	0.2842	0.3100	0.2450	0.1733	0.2965	0.1078	0.2042	0.1490
B_3	0.6065	0.5545	0.3432	0.3582	0.2475	0.4201	0.1481	0.3961	0.2993
B_4	0.5033	0.4830	0.3521	0.2225	0.2408	0.2993	0.1228	0.2810	0.2216
B_5	0.5302	0.5084	0.3823	0.3495	0.1786	0.3748	0.1459	0.2670	0.2229
B_6	0.3293	0.3219	0.2555	0.2305	0.1586	0.1728	0.1580	0.2336	0.1333
B_7	0.6382	0.6314	0.5057	0.4414	0.3746	0.4118	0.1318	0.3604	0.2329
B_8	0.5851	0.5696	0.4583	0.3659	0.3175	0.3703	0.1426	0.2394	0.2079
B_9	0.4447	0.4566	0.3376	0.2215	0.1799	0.3056	0.1133	0.2021	0.1288

the best-worst method for ranking and the fuzzy DEMATEL method, lack of knowledge and expertise (B1) received the highest weight 0.3146 (α 1) followed by high capital investment (0.191, i.e., α 3) and stable environment (0.127. i.e., α 2) while limited option of crops to be produced (B9) has the lowest weight (0.0273, i.e., α 9). If these factors, especially the first three, are considered while planning and implementation of aquaponics, it has a great potential. The present research highlights the nine major challenges that any new entrant in the field has to address according to their importance in order to set up the aquaponics unit efficiently. Managers of allied businesses can use the proposed rankings and categorizations to evaluate the correct plan of action before incurring unnecessary costs to establish the unit. Thus, the obtained results not only address the concerns of the start-ups and corporations who want to adopt aquaponics, but they also in-turn increase the total investment and number of investors, who are interested to adopt this technology, by making it easier to enter the field. The outcomes can be used as a guideline to systematically tackle and eliminate all challenges as obstacles involved in setting up an aquaponics unit. This will help in extensive reduction of capital requirements and of negative environmental impacts.

The challenges faced in the implementation of aquaponics are gaining much attention as the importance of aquaponic

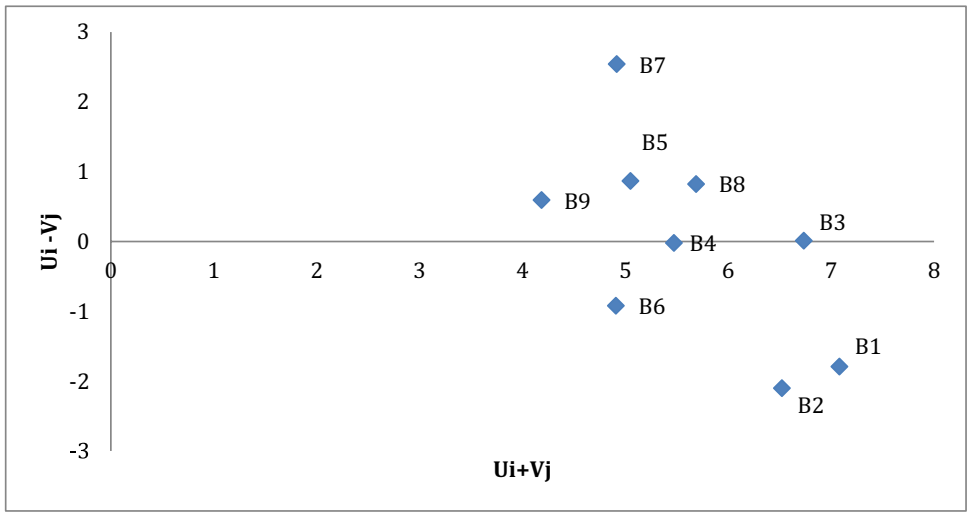


Fig. 2 Causal diagram

Table 11 The degree of central role

	B_1	B_2	B_3	B_4	B_5	B_6	B_7	B_8	B_9
U_i	2.6451	2.2107	3.3735	2.7265	2.9596	1.9936	3.7280	3.2564	2.3901
V_j	4.4366	4.3121	3.3620	2.7464	2.0925	2.9166	1.1885	2.4326	1.7963
$U_i + V_j$	7.0817	6.5228	6.7355	5.4729	5.0521	4.9102	4.9165	5.6890	4.1864
$U_i - V_j$	-1.7915	-2.1015	0.0115	-0.0199	0.8671	-0.9229	2.5395	0.8238	0.5938
Cause /effect group	Effect	Effect	Cause	Effect	Cause	Effect	Cause	Cause	Cause
Rank	8	9	5	6	2	7	1	3	4

technology is increasingly realized in the context of COVID-19 for the food security purpose and environmental point of view. Therefore, further research should also be done accommodating the sub-factors of the challenges identified in this research. Subsequently, local-weights for these sub-factors can be identified following the same fuzzy DEMATEL model. Overall, weights should be formed by including both parameters, i.e., weight of the parent challenge and weight of the sub-factor. Future research could also include the implementation frameworks to standardize the steps to be taken to tackle the ranked challenges. The research could further be focused on either reducing the time required to set up or be focused on capital required to set up the aquaponics unit.

Appendix 1. Experts profile

S. No.	Background	Designation	Experience	Location
1	Aquaponics Industry	Manager	8 years	National Capital Region
2	Aquaponics Industry	Owner	11Years	Kolkata
3	Academic	Professor (Botany)	15 Years	Bombay
4	Academic	Professor (Agriculture Science)	13 years	Chennai
5	Policy maker	Government official	12 years	Delhi
6	Policy maker	Government official	10 years	Bengaluru

Appendix 2. Questionnaire for best-worst method

Rank the most important barriers of aquaponics adoption as compared to others; assign a number from 1 to 9 to show the preference of a criterion over the others (Tables 12, 13, 14, 15, 16, 17, 18, 19, and 20). Also, rank the lease important barriers of aquaponics adoption as compared to others, assign a number from 1 to 9 to show the preference of a criterion over the others (Tables 21, 22, 23, 24, 25, 26, 27, 28, and 29).

Table 12 Rank of the most important barrier in context to Lack of knowledge and expertise

Rank the most important barriers	Lack of knowledge and expertise (B1)	Absence of Stable environment (B2)	High capital requirements (B3)	Nutrient limitation in fish excreta (B4)	Maintaining pH (B5)	Poor pest and disease management (B6)	Interrupted power supply (B7)	Limited plant and fish combinations (B8)	Limited options of crops to be produced (B9)
Lack of knowledge and expertise (B1)									

Table 13 Rank of the most important barrier in context to Absence of stable environment

Rank the most important Barriers	Lack of knowledge and expertise (B1)	Absence of Stable environment (B2)	High capital requirements (B3)	Nutrient limitation in fish excreta (B4)	Maintaining pH (B5)	Poor pest and disease management (B6)	Interrupted power supply (B7)	Limited plant and fish combinations (B8)	Limited options of crops to be produced (B9)
Absence of Stable environment (B2)									

Table 14 Rank of the most important barrier in context to High capital requirements

Rank the most important Barriers	Lack of knowledge and expertise (B1)	Absence of Stable environment (B2)	High capital requirements (B3)	Nutrient limitation in fish excreta (B4)	Maintaining pH (B5)	Poor pest and disease management (B6)	Interrupted power supply (B7)	Limited plant and fish combinations (B8)	Limited options of crops to be produced (B9)
High capital requirements (B3)									

Table 15 Rank of the most important barrier in context to Nutrient limitation in fish excreta

Rank the most important Barriers	Lack of knowledge and expertise (B1)	Absence of Stable environment (B2)	High capital requirements (B3)	Nutrient limitation in fish excreta (B4)	Maintaining pH (B5)	Poor pest and disease management (B6)	Interrupted power supply (B7)	Limited plant and fish combinations (B8)	Limited options of crops to be produced (B9)
Nutrient limitation in fish excreta (B4)									

Table 16 Rank of the most important barrier in context to Maintaining pH

Rank the most important Barriers	Lack of knowledge and expertise (B1)	Absence of Stable environment (B2)	High capital requirements (B3)	Nutrient limitation in fish excreta (B4)	Maintaining pH (B5)	Poor pest and disease management (B6)	Interrupted power supply (B7)	Limited plant and fish combinations (B8)	Limited options of crops to be produced (B9)
Maintaining pH (B5)									

Table 17 Rank of the most important barrier in context to Poor pest and disease management

Rank the most important Barriers	Lack of knowledge and expertise (B1)	Absence of Stable environment (B2)	High capital requirements (B3)	Nutrient limitation in fish excreta (B4)	Maintaining pH (B5)	Poor pest and disease management (B6)	Interrupted power supply (B7)	Limited plant and fish combinations (B8)	Limited options of crops to be produced (B9)
Poor pest and disease management (B6)									

Table 18 Rank of the most important Barrier in context to Interrupted power supply

Rank the most important Barriers	Lack of knowledge and expertise (B1)	Absence of Stable environment (B2)	High capital requirements (B3)	Nutrient limitation in fish excreta (B4)	Maintaining pH (B5)	Poor pest and disease management (B6)	Interrupted power supply (B7)	Limited plant and fish combinations (B8)	Limited options of crops to be produced (B9)
Interrupted power supply (B7)									

Table 19 Rank of the most important Barrier in context to Limited plant and fish combinations

Rank the most important Barriers	Lack of knowledge and expertise (B1)	Absence of Stable environment (B2)	High capital requirements (B3)	Nutrient limitation in fish excreta (B4)	Maintaining pH (B5)	Poor pest and disease management (B6)	Interrupted power supply (B7)	Limited plant and fish combinations (B8)	Limited options of crops to be produced (B9)
Limited plant and fish combinations (B8)									

Table 20 Rank of the most important barrier in context to Limited options of crops to be produced

Rank the most important Barriers	Lack of knowledge and expertise (B1)	Absence of Stable environment (B2)	High capital requirements (B3)	Nutrient limitation in fish excreta (B4)	Maintaining pH (B5)	Poor pest and disease management (B6)	Interrupted power supply (B7)	Limited plant and fish combinations (B8)	Limited options of crops to be produced (B9)
Limited options of crops to be produced (B9)									

Table 21 Rank of the least important barrier in context to Lack of knowledge and expertise

Rank the least important barriers	Lack of knowledge and expertise (B1)
Lack of knowledge and expertise (B1)	
Absence of Stable environment (B2)	
High capital requirements (B3)	
Nutrient limitation in fish excreta (B4)	
Maintaining pH (B5)	
Poor pest and disease management (B6)	
Interrupted power supply (B7)	
Limited plant and fish combinations (B8)	
Limited options of crops to be produced (B9)	

Table 22 Rank of the least important Barrier in context to Absence of stable environment

Rank the least important barriers	Absence of stable environment (B2)
Lack of knowledge and expertise (B1)	
Absence of Stable environment (B2)	
High capital requirements (B3)	
Nutrient limitation in fish excreta (B4)	
Maintaining pH (B5)	
Poor pest and disease management (B6)	
Interrupted power supply (B7)	
Limited plant and fish combinations (B8)	
Limited options of crops to be produced (B9)	

Table 23 Rank of the least important barrier in context to High capital requirements

Rank the least important barriers	High capital requirements (B3)
Lack of knowledge and expertise (B1)	
Absence of stable environment (B2)	
High capital requirements (B3)	
Nutrient limitation in fish excreta (B4)	
Maintaining pH (B5)	
Poor pest and disease management (B6)	
Interrupted power supply (B7)	
Limited plant and fish combinations (B8)	
Limited options of crops to be produced (B9)	

Table 24 Rank of the least important barrier in context to Nutrient limitation in fish excreta

Rank the least important barriers	Nutrient limitation in fish excreta (B4)
Lack of knowledge and expertise (B1)	
Absence of stable environment (B2)	
High capital requirements (B3)	
Nutrient limitation in fish excreta (B4)	
Maintaining pH (B5)	
Poor pest and disease management (B6)	
Interrupted power supply (B7)	
Limited plant and fish combinations (B8)	
Limited options of crops to be produced (B9)	

Table 25 Rank of the least important Barrier in context to Maintaining pH

Rank the least important barriers	Maintaining pH (B5)
Lack of knowledge and expertise (B1)	
Absence of Stable environment (B2)	
High capital requirements (B3)	
Nutrient limitation in fish excreta (B4)	
Maintaining pH (B5)	
Poor pest and disease management (B6)	
Interrupted power supply (B7)	
Limited plant and fish combinations (B8)	
Limited options of crops to be produced (B9)	

Table 26 Rank of the least important barrier in context to Poor pest and disease management

Rank the least important barriers	Poor pest and disease management (B6)
Lack of knowledge and expertise (B1)	
Absence of Stable environment (B2)	
High capital requirements (B3)	
Nutrient limitation in fish excreta (B4)	
Maintaining pH (B5)	
Poor pest and disease management (B6)	
Interrupted power supply (B7)	
Limited plant and fish combinations (B8)	
Limited options of crops to be produced (B9)	

Table 27 Rank of the least important barrier in context to Interrupted power supply

Rank the least important barriers	Interrupted power supply (B7)
Lack of knowledge and expertise (B1)	
Absence of Stable environment (B2)	
High capital requirements (B3)	
Nutrient limitation in fish excreta (B4)	
Maintaining pH (B5)	
Poor pest and disease management (B6)	
Interrupted power supply (B7)	
Limited plant and fish combinations (B8)	
Limited options of crops to be produced (B9)	

Table 28 Rank of the least important barrier in context to Limited plant and fish combinations

Rank the least important barriers	Limited plant and fish combinations (B8)
Lack of knowledge and expertise (B1)	
Absence of Stable environment (B2)	
High capital requirements (B3)	
Nutrient limitation in fish excreta (B4)	
Maintaining pH (B5)	
Poor pest and disease management (B6)	
Interrupted power supply (B7)	
Limited plant and fish combinations (B8)	
Limited options of crops to be produced (B9)	

Table 29 Rank of the least important barrier in context to Limited options of crops to be produced

Rank the least important barriers	Limited options of crops to be produced (B9)
Lack of knowledge and expertise (B1)	
Absence of Stable environment (B2)	
High capital requirements (B3)	
Nutrient limitation in fish excreta (B4)	
Maintaining pH (B5)	
Poor pest and disease management (B6)	
Interrupted power supply (B7)	
Limited plant and fish combinations (B8)	
Limited options of crops to be produced (B9)	

Appendix 3. Questionnaire for fuzzy DEMATEL method

Table 30 and 31 present the barriers and scale on which the experts are supposed to rank. Table 32 represents the questionnaire utilized for conducting the DEMATEL study. Each expert was asked to evaluate the impact of one indicator over the other indicators using an integer scale (from 0 to 4). Table 31 shows that if Indicator (i) has a weak direct influence on indicator (j), then a score of "1" is given to represent this weak influence. Conversely, if the indicator (i) has a strong direct influence on the indicator (j) then, a score of "3" is assigned and so on. A high score represents the belief of a higher influence of indicator (i) over indicator (j). The detailed scale is shown in Table 31.

Table 30 Barriers of aquaponics adoption

Symbol	Name of barrier
B1	Lack of knowledge and expertise
B2	Absence of Stable environment
B3	High capital requirements
B4	Nutrient limitation in fish excreta
B5	Maintaining pH
B6	Poor pest and disease management
B7	Interrupted power supply
B8	Limited plant and fish combinations
B9	Limited options of crops to be produced

Table 31 Linguistic terms and corresponding fuzzy numbers

Linguistic terms	Influence score	Triangular fuzzy numbers
No influence	0	0, 0.1, 0.25
Low influence	1	0.1, 0.25, 0.5
Moderate influence	2	0.25, 0.5, 0.6
High influence	3	0.5, 0.6, 0.75
Very high influence	4	0.6, 0.75, 1

Table 32 Barriers of aquaponics adoption and symbol

Indicators (j) Indicators (i)	B1	B2	B3	B4	B5	B6	B7	B8	B9
B1	0.00								
B2		0.00							
B3			0.00						
B4				0.00					
B5					0.00				
B6						0.00			
B7							0.00		
B8								0.00	
B9									0.00

Author contribution All authors contributed in every stage including conception and design of this research. Questionnaire design, data collection and analysis were performed by Dr. Girish Kumar and Dr. Ram C. Bhujel. Results from the study were checked by Dr. Ram C. Bhujel and Dr. Mohammad Asjad. The first draft of the manuscript was written by Mr. Aniket Aggarwal, Divyansh Gupta, and Ashish Yadav, modified by Dr. Girish Kumar and Dr. Ram C. Bhujel, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Declarations

Ethical approval Full compliance with ethical standards policy of journal

Consent to participate Not applicable.

Consent for publication Not applicable.

Conflict of interest The authors declare no competing interests.

References

Al-Muftah H, Weerakkody V, Rana NP, Sivarajah U, Irani Z (2018) Factors influencing e-diplomacy implementation: exploring causal relationships using interpretive structural modelling. Gov Inf Q 35(3):502–514. https://doi.org/10.1016/j.giq.2018.03.002

Alshrouf A (2017) Hydroponics, aeroponic and aquaponic as compared with conventional farming. Am Sci Res J Eng Technol Sci 27(1) https://asrjetsjournal.org/index.php/American_Scientific_Journal/article/view/2543/1028

Altieri MA, Nicholls CI (2020) Agroecology and the emergence of a post COVID-19 agriculture. Agric Human Values 37(3):525–526. https://doi.org/10.1007/s10460-020-10043-7

Bai C (2018) Best-worst multi-criteria decision-making method: some limits and improved models. University of Electronic Science and Technology of China

Bhatia MS, Srivastava RK (2018) Analysis of external barriers to remanufacturing using grey-DEMATEL approach: an Indian perspective. Resour Conserv Recycling 136(December 2017):79–87. https://doi.org/10.1016/j.resconrec.2018.03.021

Bhavani RV, Gopinath R (2020) The COVID19 pandemic crisis and the relevance of a farm-system-for-nutrition approach. Food Security 12(4):881–884. https://doi.org/10.1007/s12571-020-01071-6

Brewer A (2019) Barriers and incentives to engaging in aquaponics, p 45 https://deepblue.lib.umich.edu/bitstream/handle/2027.42/148828/Brewer_Alexandria_Thesis.pdf?sequence=1&isAllowed=y

Brewer A, Alfaro JF, Malheiros TF (2021) Evaluating the capacity of small farmers to adopt aquaponics systems: empirical evidence from Brazil. Renew Agric Food Syst 36(4):375–383

Cammies C, Mytton D, Crichton R (2021) Exploring economic and legal barriers to commercial aquaponics in the EU through the lens of the UK and policy proposals to address them. Aquac Int 29(3):1245–1263

Chang B, Chang CW, Wu CH (2011) Fuzzy DEMATEL method for developing supplier selection criteria. Expert Syst Appl 38(3):1850–1858. https://doi.org/10.1016/j.eswa.2010.07.114

Cheng CH, Yang KL, Hwang CL (1999) Evaluating attack helicopters by AHP based on linguistic variable weight. Eur J Oper Res 116(2):423–435. https://doi.org/10.1016/S0377-2217(98)00156-8

El-Essawy H, Nasr P, Sewilam H (2019) Aquaponics: a sustainable alternative to conventional agriculture in Egypt – a pilot scale investigation. Environ Sci Pollut Res 26(16):15872–15883. https://doi.org/10.1007/s11356-019-04970-0

El-Sayed A-FM (2020) Intensive culture. In: Tilapia Culture. https://doi.org/10.1016/b978-0-12-816509-6.00006-9

Esangbedo MO, Che A (2016) Grey weighted sum model for evaluating business environment in West Africa. Math Probl Eng 2016(Mcdm). https://doi.org/10.1155/2016/3824350

Esfandiari M, Rizvandi M (2014) An application of TOPSIS method for ranking different strategic planning methodology. Manage Sci Lett 4(7):1445–1448. https://doi.org/10.5267/j.msl.2014.6.022

Estim A, Shaleh SRM, Shapawi R, Saufie S, Mustafa S (2020) Maximizing efficiency and sustainability of aquatic food production from aquaponics systems-a critical review of challenges and solution options. Aquac Stud 20(1):65–72. https://doi.org/10.4194/2618-6381-v20_1_08

Gabus A, Fontela E (1973) Perceptions of the world problematique: communication procedure, communicating with those bearing collective responsibility (DEMATEL Report no.1). Battelle Geneva Research Centre, Geneva, Switzerland

Goddek S, Delaide B, Mankasingh U, Ragnarsdottir KV, Jijakli H, Thorarinsdottir R (2015) Challenges of sustainable and commercial aquaponics. Sustainability (Switzerland) 7(4):4199–4224. https://doi.org/10.3390/su7044199

Government of India (2019) Agriculture and food management. In: Economic Survey 2019-20, Ministry of Finance (Vol. 2) https://www.indiabudget.gov.in/economicsurvey/doc/vol2chapter/echap07_vol2.pdf

Greenfeld A, Becker N, Bornman JF, Angel DL (2020) Identifying knowledge levels of aquaponics adopters. Environ Sci Pollut Res 27(4):4536–4540. https://doi.org/10.1007/s11356-019-06758-8

Greenfeld A, Becker N, McIlwain J, Fotedar R, Bornman JF (2019) Economically viable aquaponics? Identifying the gap between potential and current uncertainties. Rev Aquac 11(3):848–862. https://doi.org/10.1111/raq.12269

Han Y, Deng Y (2018) An enhanced fuzzy evidential DEMATEL method with its application to identify critical success factors. Soft Comput 22(15):5073–5090. https://doi.org/10.1007/s00500-018-3311-x

IBEF (2020) Agriculture in India: information about Indian agriculture & its importance. Indian Brand Equity Foundation https://www.ibef.org/industry/agriculture-india.aspx

Jain P, Baghla K, Aditya R (2020) Effect of corona / COVID19 on the agricultural sector in India. Pharma Innov 9(5):41–45 https://www.thepharmajournal.com/archives/2020/vol9issue5/PartB/9-4-75-293.pdf

Junge R, König B, Villarroel M, Komives T, Jijakli MH (2017) Strategic points in aquaponics. *Water (Switzerland)* 9(3):1–9. https://doi.org/10.3390/w9030182

Kala, M., Mann, G., Rao, N., Mishra, P., Suhag, R., Sinha, R., Kodidala, S. P., Kanwar, S., & Khullar, V. (2018). Annual policy review (Issue April 2018). https://prsindia.org/files/policy/policy_annual_policy_review/APR 2017-18_0.pdf

Karthika BP (2018) A Village finding future in aquaponics. In: Krishijagran https://krishijagran.com/success-story/a-village-finding-future-in-aquaponics/

König B, Janker J, Reinhardt T, Villarroel M, Junge R (2018) Analysis of aquaponics as an emerging technological innovation system. J Clean Prod 180:232–243. https://doi.org/10.1016/j.jclepro.2018.01.037

Labella Á, Dutta B, Rodríguez RM, Martínez L (2021) A linguistic 2-tuple best-worst method. In: The International Workshop on Best-Worst Method. Springer, Cham, pp 41–51

Lee PTW, Wu JZ, Hu KC, Flynn M (2013) Applying analytic network process (ANP) to rank critical success factors of waterfront redevelopment. Int J Shipping Transpt Logistics 5(4–5):390–411. https://doi.org/10.1504/IJSTL.2013.055276

Lin CJ, Wu WW (2008) A causal analytical method for group decision-making under fuzzy environment. Expert Syst Appl 34(1):205–213. https://doi.org/10.1016/j.eswa.2006.08.012

Liou JJH, Yen L, Tzeng GH (2008) Building an effective safety management system for airlines. J Air Transp Manag 14(1):20–26. https://doi.org/10.1016/j.jairtraman.2007.10.002

Love DC, Fry JP, Genello L, Hill ES, Frederick JA, Li X, Semmens K (2014) An international survey of aquaponics practitioners. PLoS One 9(7):1–10. https://doi.org/10.1371/journal.pone.0102662

Mandic K, Delibasic B (2014) Supplier selection using interpolative Boolean algebra and logic aggregation. In: Communications in Computer and Information Science, 443CCIS(PART 2), pp 1–9. https://doi.org/10.1007/978-3-319-08855-6_1

Matthews H (2017) Incentives and barriers to adopting aquaponic and biofloc systems in Canada. Geography Department University of Toronto (Canada). https://www.proquest.com/openview/5fdec4ee094ee14e6603e631b4e3db4b/1?pq-origsite=gscholar&cbl=18750

Monsees H, Kloas W, Wuertz S (2017) Decoupled systems on trial: eliminating bottlenecks to improve aquaponic processes. PLoS One 12(9):1–18. https://doi.org/10.1371/journal.pone.0183056

Mukherjee S (2019) Aquaponics farming gaining ground but limited by costs, product range. In: Business Standard News. https://www.business-standard.com/article/economy-policy/aquaponics-farming-gaining-ground-but-limited-by-costs-product-range-119030800630_1.html

Nagpal S, Kumar A, Khatri SK (2017) Modeling interrelationships between CSF in ERP implementations: total ISM and MICMAC approach. Int J Syst Assur Eng Manage 8(4):782–798. https://doi.org/10.1007/s13198-017-0647-z

Njinga RL, Moyo MN, Abdulmaliq SY (2013) Analysis of essential elements for plants growth using instrumental neutron activation analysis. Int J Agron 2013:1–9. https://doi.org/10.1155/2013/156520

Pitchaimuthu S, Thakkar JJ, Gopal PRC (2019) Modelling of risk factors for defence aircraft industry using interpretive structural modelling, interpretive ranking process and system dynamics. Meas Bus Excell 23(3):217–239. https://doi.org/10.1108/MBE-05-2018-0028

Rezaei J (2015) Best-worst multi-criteria decision-making method. Omega (United Kingdom) 53:49–57. https://doi.org/10.1016/j.omega.2014.11.009

Rosgren, C., & Grahler, K. (2022). Barriers to sustainable innovation in aquaculture: a study of Swedish aquaculture.

Ru D, Liu J, Hu Z, Zou Y, Jiang L, Cheng X, Lv Z (2017) Improvement of aquaponic performance through micro- and macro-nutrient addition. Environ Sci Pollut Res 24(19):16328–16335. https://doi.org/10.1007/s11356-017-9273-1

Savidov NA, Hutchings E, Rakocy JE (2007) Fish and plant production in a recirculating aquaponic system: a new approach to sustainable agriculture in Canada. Acta Horticulturae 742:209–222. https://doi.org/10.17660/actahortic.2007.742.28

Si SL, You XY, Liu HC, Zhang P (2018) DEMATEL technique: a systematic review of the state-of-the-art literature on methodologies and applications. Math Probl Eng 2018(1). https://doi.org/10.1155/2018/3696457

Simanovski S, Pirkebner A (2018) Aquaponics in small-scale farming // Effishent - The Social Startup Using 90 % Less Water. In: Nexus Resource Platform https://www.water-energy-food.org/news/aquaponics-in-small-scale-farming-90-less-water-usage-thanks-to-social-startup-effishent#:~:text=Theaquaponicstechnologycombinesfish,ofthewater-foodNexus

Singh S, Kumar R, Panchal R, Tiwari MK (2021) Impact of COVID-19 on logistics systems and disruptions in food supply chain. Int J Prod Res 59(7):1993–2008

Somerville C, Cohen M, Pantanella E, Stankus A, Lovatelli A (2014) Small-scale aquaponic food production: Integrated fish and plant farming. FAO Fisheries and Aquaculture Technical Paper http://www.fao.org/3/i4021e/i4021e.pdf

Stuiver M, O'hara, S. (2022) Chapter 3: Envisioning the symbiotic city in 2050: two visions of Washington DC and the Netherlands. In: The symbiotic city: Voices of nature in urban transformations. Wageningen Academic Publishers, p 6381

Sunny AR, Islam MM, Rahman M, Miah MY, Mostafiz M, Islam N, Hossain MZ, Chowdhury MA, Islam MA, Keus HJ (2019) Cost effective aquaponics for food security and income of farming

households in coastal Bangladesh. Egypt J Aqua Res 45(1):89–97. https://doi.org/10.1016/j.ejar.2019.01.003

Turnsek M, Joly A, Thorarinsdottir R, Junge R (2020) Challenges of commercial aquaponics in Europe: beyond the hype. Water (Switzerland) 12(1):1–18. https://doi.org/10.3390/w12010306

Turnšek M, Morgenstern R, Schröter I, Mergenthaler M, Hüttel S, Leyer M (2019) Commercial aquaponics: a long road ahead. In: Aquaponics Food Production Systems. https://doi.org/10.1007/978-3-030-15943-6_18

Venkatesh VG, Zhang A, Luthra S, Dubey R, Subramanian N, Mangla S (2017) Barriers to coastal shipping development: an Indian perspective. Transp Res Part D: Transp Env 52:362–378. https://doi.org/10.1016/j.trd.2017.03.016

Vermeulen T, Kamstra A (2013) The need for systems design for robust aquaponic systems in the urban environment. Acta Horticulturae 1004:71–78. https://doi.org/10.17660/actahortic.2013.1004.6

Villarroel M, Junge R, Komives T, König B, Plaza I, Bittsánszky A, Joly A (2016) Survey of aquaponics in Europe. Water (Switzerland) 8(10):3–9. https://doi.org/10.3390/w8100468

Wang CH, Chen JN (2012) Using quality function deployment for collaborative product design and optimal selection of module mix. Comput Ind Eng 63(4):1030–1037. https://doi.org/10.1016/j.cie.2012.06.014

World Bank (2012) India: Issues and Priorities for Agriculture. The World Bank http://www.worldbank.org/en/news/feature/2012/05/17/india-agriculture-issues-priorities

WWF (2013) Wellington Aquaponics : an environmental success Support WWF-Canada Today. Canada Today https://wwf.ca/2013/09/05/wellington-aquaponics-environmental-success-story/

Yang T, Kim HJ (2019) Nutrient management regime affects water quality, crop growth, and nitrogen use efficiency of aquaponic systems. Sci Hortic 256(March):108619. https://doi.org/10.1016/j.scienta.2019.108619

Yep B, Zheng Y (2019) Aquaponic trends and challenges – a review. J Clean Prod 228:1586–1599. https://doi.org/10.1016/j.jclepro.2019.04.290

Zadeh L (1965) Fuzzy sets. Inf Control 8(3):16. https://doi.org/10.1016/S0019-9958(65)90241-X

Zhao J, Bose BK (2002) Evaluation of membership functions for fuzzy logic controlled induction motor drive. In: IEEE 2002 28th Annual Conference of the Industrial Electronics Society. IECON 02 (Vol. 1, pp. 229-234). IEEE

Zhu S, Chen S (2002) The impact of temperature on nitrification rate in fixed film biofilters. Aquacult Eng 26(4):221–237. https://doi.org/10.1016/S0144-8609(02)00022-5

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.

Hindawi
Wireless Communications and Mobile Computing
Volume 2023, Article ID 8068038, 27 pages
https://doi.org/10.1155/2023/8068038

Research Article

Blockchain-Based Secure Localization against Malicious Nodes in IoT-Based Wireless Sensor Networks Using Federated Learning

Gebrekiros Gebreyesus Gebremariam,[1,2] **J. Panda**,[2] **and S. Indu**[2]

[1]*Department of Electronics and Communication Engineering, Raya University, Maychew, 7020 Tigray, Ethiopia*
[2]*Department of Electronics and Communication Engineering, Delhi Technological University, Shahbad Daulatpur, Main Bwana Road, Delhi 110042, India*

Correspondence should be addressed to Gebrekiros Gebreyesus Gebremariam; kiros2004comp@gmail.com

Received 11 November 2022; Revised 9 January 2023; Accepted 13 January 2023; Published 31 January 2023

Academic Editor: Javier Prieto

Copyright © 2023 Gebrekiros Gebreyesus Gebremariam et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Wireless sensor networks are the core of the Internet of Things and are used in healthcare, locations, the military, and security. Threats to the security of wireless sensor networks built on the Internet of Things (IoT-WSNs) can come from a variety of sources. This study proposes secure attack localization and detection in IoT-WSNs to improve security and service delivery. The technique used blockchain-based cascade encryption and trust evaluation in a hierarchical design to generate blockchain trust values before beacon nodes broadcast data to the base station. Simulation results reveal that cascading encryption and feature assessment measure the trust value of nodes by rewarding each other for service provisioning and trust by removing malicious nodes that reduce localization accuracy and quality of service in the network. Federated machine learning improves data security and transmission by merging raw device data and placing malicious threats in the blockchain. Malicious nodes are classified through federated learning. Federated learning combines hybrid random forest, gradient boost, ensemble learning, K-means clustering, and support vector machine approaches to classify harmful nodes via a feature assessment process. Comparing the proposed system to current ones shows an average detection and classification accuracy of 100% for binary and 99.95% for multiclass. This demonstrates that the suggested approach works well for large-scale IoT-WSNs, both in terms of performance and security, when utilizing heterogeneous wireless senor networks for the providing of secure services.

1. Introduction

Technology and network advancements have created a new environment where the Internet of Things (IoT) is flourishing, fostering several positive social and economic changes [1]. The Internet of Things (IoT) are the networked and interconnected systems, devices, applications, and people that process data [2]. The Internet of Things makes it possible for physical objects to communicate with one another and share critical information while also making decisions and completing important tasks. As a result, these objects now have the capacity to see, hear, think, and carry out significant acts [3]. IoT-WSNs are important and popular research areas in recent applications, including military, smart cities, healthcare, and environmental [4]. The rapid growth of the Internet of Things (IoT) is largely due to the contributions of wireless sensor networks (WSNs), which have both practical and academic importance. Ad hoc networks, or WSNs, are made up of a swarm of devices, or "sensor nodes," each of which is equipped with some form of sensing technology. Sensor nodes should be able to self-organize, be deployed at random, be fault-tolerant, and monitor a wide region with high precision. IoT-WSNs are self-organizing networks with different parameters for monitoring and collecting environmental data. The sensor nodes cooperate and evaluate each other for effective data transmission and detection and removal of malicious nodes in the network. The malicious attacks can be internal or external for exploiting information and creating wrong routing information between the sensor node and base station by compromising the beacon node in the cluster region. The malicious node degrades the network performance by

broadcasting wrong information and creating mistrust between the nodes and end users. The trust evaluation process continuously uses threshold and signal strength values between the sensor nodes to detect and analyze the network's malicious nodes.

The integrity of bitcoin networks is ensured by blockchain, a distributed ledger technology [5]. This groundbreaking architecture is now often utilized in a variety of distributed contexts, including healthcare and automotive ad hoc networks (VANET). There are three primary reasons why blockchain technology is well-suited to addressing IoT security concerns, according to a number of studies. As with an IoT network, the data associated with blockchain technology is stored and handled in a decentralized fashion by all of the agents in the network. Distributed blockchain technology offers many advantages over traditional centralized management approaches, including the elimination of a potential single point of failure and a significant reduction in the cost of keeping a large number of Internet of Things devices up and running. Lack of trust between IoT devices is a commonplace because of the energy-limited nature of these devices. As a result, sensitive information stored on these devices is at risk of being exposed or stolen. Due to blockchains unique immutability, it allows parties to exchange information without having to rely on one other's integrity. In the blockchain design, data is saved in blocks, each of which has a hash backup of the data in the preceding block. In many instances, it is also necessary for the devices to be completely untraceable. However, most gadgets have a low setup and are not suited for the complexity of standard encryption methods. With the blockchain architecture of IoT-WSNs, entities connect with one another using transactions as the fundamental unit of information exchange. In the first iteration of the blockchain, every piece of data pertaining to a client's commercial exchange is considered a transaction.

Many fields, including medicine, computer vision, and wireless communication, have benefited from the significant progress made in recent years in machine learning (ML) techniques and especially in deep learning (DL) models. ML also encompasses shallow models [6]. It can provide means for detecting certain forms of attacks without requiring intensive human engagement. Intrusion detection systems (IDS) that employ ML techniques can train themselves by observing and comparing "normal" and "abnormal" flows of traffic. It can be used to detect harmful traffic patterns in a dataset or in an actual environment. Although these approaches have been put to good use for IDS, they often necessitate a centralized organization to process and aggregate data from all users on the network and found that high packet loss rate reduces prediction accuracy at larger network size. In this study, we proposed blockchain-based hybrid federated machine learning techniqe for locating and identifying the origin of assaults in IoT-WSNs, using the immutable and distributed ledger technology blockchain.

1.1. Blockchain Applications in IoT-WSNs. Blockchain is a distributed database for storing transactions' records to provide trust in unknown nodes in the network with the same edge. A blockchain, as its name implies, is a series of linked blocks. The structure for storing bargaining data is a chain of blocks, each of which has its own hash value and contains a specified quantity of transaction data. All the data on the blockchain is secure because each block contains a unique hash value for the one before it [5]. The blockchain is a distributed, unchangeable ledger of all transactions that have been validated by a majority of the network's peers. Distributed ledger is a term for the system of decentralized data storage it describes [7]. Whenever a ledger transaction is performed, it is validated with the approval of the majority of the network users. The most well-known application of blockchain technology in the real world is Bitcoin. Connecting blockchain technology with the Internet of Things has many benefits. For example, a decentralized blockchain storage architecture can synchronize IoT devices and deliver real-time data to each IoT node. However, the data produced by these applications is enormous, leading to problems associated with big data. Therefore, artificial intelligence (AI) functions as an analytical tool and contributes to decision-making, categorization, prediction, and detection in a blockchain-enabled IoT network to solve this problem. Further, the analysis in the current blockchain-based architecture is performed on a cloud server. However, there are limitations associated with using a centralized server, including slow speeds and accuracy, as well as low latency and limited computational storage. Fog computing, a new distributed paradigm that may be able to meet these needs, warrants extensive study. Fog computing, which Cisco first defined in 2012, is a distributed architecture that brings cloud services to the network's periphery. Fog computing's principles offer extensive assistance to blockchain-based applications, including but not limited to location awareness, support for heterogeneity, low latency, mobility, and geodistribution. In addition to these load-balancing and data-gathering benefits, fog computing has found application in other areas. As a result, this research combines fog computing with AI to create a distributed security mechanism that can identify DoS attacks directed at mining pools in a blockchain-enabled IoT network. The integrated secure attack detection model using blockchain, IoT, AI, and fog computing is depicted in Figure 1 for various applications.

Blockchain is an up-and-coming technology that is being implemented in numerous sectors to facilitate security and distributed trust among peer nodes [8]. The following sections cover some literature that has utilized blockchain technology to safeguard IoT applications. For authentication and authorization in IoT-based WSNs, blockchain technology was integrated into IoT. To solve this problem, we propose a blockchain-based decentralized localization technique, a pseudonymous permission management framework that gives users complete authority over their own information while also protecting their privacy. Blockchain models are decentralized access control managers that store access permissions to resources. A new subfield of study is exploring the use of blockchain technology in IoT based-WSNs [9].

Some academics have recently included blockchain in WSNs to solve problems with processing capability, data storage, node failure, and user access management. The mobile edge computing (MEC) server can store encrypted

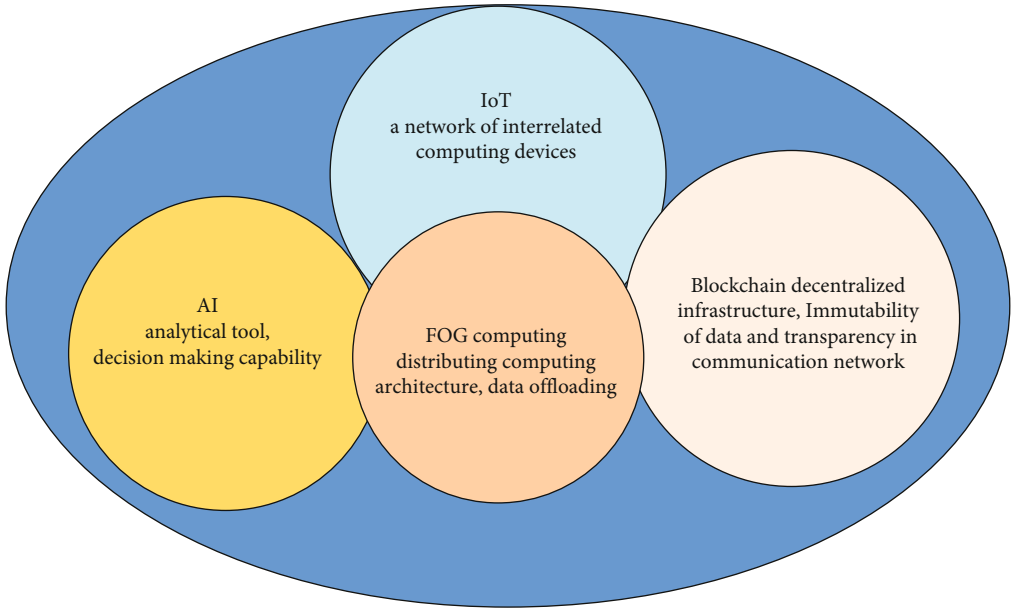


Figure 1: Illustrating of secure IoT-based WSNs showing the use of a detection model incorporating blockchain, IoT, AI, and fog computing [7].

block hashes in memory and delegate computationally difficult mining work to the edge computing nodes. The blockchain uses census algorithms to achieve agreement between unknown nodes using proof-of-work (PoW), proof-of-state, and proof-of-authority (PoA). The blockchain establishes trust in the network by storing the transactions in the block in chronological order consisting of headers and bodies. The headers contain the hash elements, including the transaction root, previous blocks, and timestamp. Blockchain techniques are applied in different domains in IoT-based wireless sensor networks for effective localization of the beacon nodes in the network. Sensor nodes sense and collect data containing precise location and identity to the sink node in the network. However, malicious nodes' presence affects the location and position of sensor nodes by compromising the beacon nodes and sending wrong information to the sink node. The blockchain-based secure localization scheme provides reliable service to the client nodes using the nonrepudiation model.

A sensor node's location can be determined by a global positioning system (GPS) or via other means of precise localization [10]. One approach is to centralize the management of information about the network as a whole, employing mathematical techniques to determine the locations of individual nodes. In WSN, the support vector machine is primarily used for securing nodes and determining their locations. The mobile node localization strategy uses support vector machines and information about network accessibility. It determines the node's position by measuring the strength of the incoming signal or RSSI. Fast and widespread localization is possible with localized SVM, although it is still vulnerable to outliers in the training data. Support vector regression (SVR), a unique data extraction method for localization models, has led to more accurate localization. Without increasing the price of the necessary gear, it enhances localization's precision. Using SVM-based learning with a small number of anchor nodes, we can precisely pinpoint the location of the unknown node. Grid cells of finite size guarantee precision. SVM-based DV-hop works well for large-scale networks. Some AI capabilities are made available to sensor nodes by use of a localization method based on a self-organizing map (SOM). SOM can learn classification on its own without human input. Each sensor node's location is calculated using a self-organizing map. The input layer consists of the spatial coordinates of eight anchor nodes close to the unknown node. After training, the network's output layer provides the unidentified node's unique 2D spatial coordinates. This approach has the drawback that nodes must be uniformly distributed across the monitored area.

This study presents a fog computing-based distributed intrusion detection system (IDS) for mining pools in blockchain-enabled IoT networks, complete with safe planning and simulation, attack detection, and classification. Hybrid federated machine learning algorithms are used to train the suggested IDS. Figure 2 depicts a high-level view of the suggested distributed paradigm. Sensing nodes are the building blocks of the detection system, and it is their job to track and identify any moving objects that come within range. Based on their performance in training and testing, all IoT sensors belong to one of several groups, each with the same capabilities for detection. This cluster sends its data to nearby fog nodes. Since fog nodes act as a gateway or access point for IoT devices, they must incorporate an intrusion detection system as part of their security infrastructure. Incoming traffic is analyzed by IDS, and actions are done based on its results. Assuming a typical influx of data, the transaction is broadcast throughout the mining pool's memory. In a blockchain network hosted in the cloud, miners choose which transactions to mine and create new blocks.

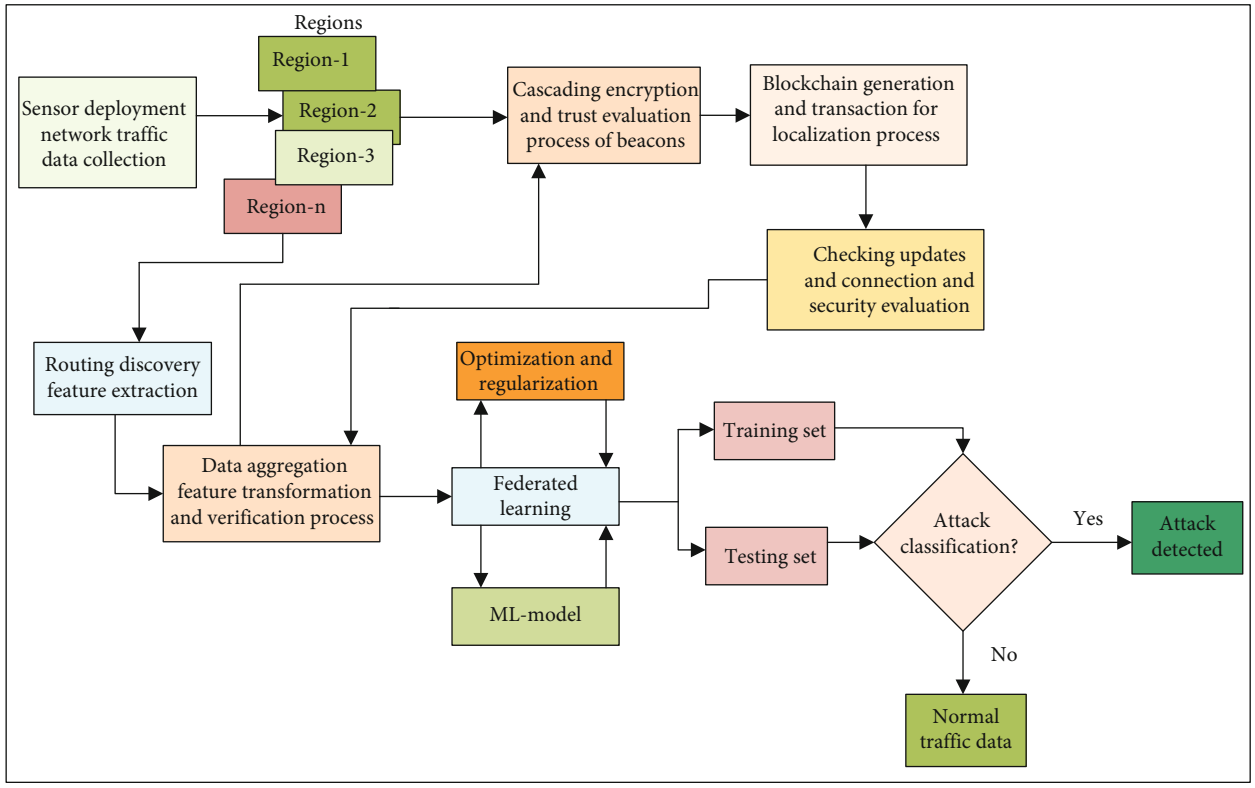


Figure 2: Blockchain-based secure localization detection of attacks for IoT-based wireless sensor networks using federated learning.

To alert the administrator to take corrective action in the event of malicious or invalid transactions, the IDS will sound an alarm. Scalability and detection performances for routing threats in IoT-based WSNs are two advantages of the proposed approach. It functions well in IoT systems that incorporate several wireless sensor networks.

1.2. Motivation and Contribution. The literature reveals a wide variety of security issues and concerns related to the design, development, and implementation of IDS, detection, and classification attacks utilizing blockchain technology in IoT-based wireless sensor networks. As the number of distributed denial of service (DDoS) attacks in the blockchain-IoT ecosystem rises, all IoT networks supported by blockchain will become increasingly at risk. Protecting a blockchain-based Internet of Things network requires a complex distributed security architecture, making sure a security mechanism can deal with the massive amounts of data produced by IoT devices and applies the right analytical tools in a decentralized framework. It is not easy to make an IDS that can tell the difference between safe and dangerous online transactions. Not much is known about how to prevent distributed denial of service (DDoS) attacks against mining pools in blockchain-enabled Internet of Things (IoT) wireless sensor networks once the corresponding model has been implemented. Because of these issues, we presented a hybrid federated machine learning-based technique for detecting and localizing routing attacks in WSNs-IoT by analyzing representative datasets.

In light of these problems, we proposed the use of a secure blockchain enabled by federated learning (FL) in Internet of Things (IoT) wireless sensor networks in order to identify and localize assaults. With FL, devices can work together to learn without transmitting information to a central hub. That is to say, ML/DL can be trained in a distributed fashion, requiring several devices and servers to process data over multiple rounds of training. The two main components of this approach are local learning and model transmission, which together provide for the privacy preservation and cost reduction typical of more traditional centralized machine learning approaches. When using FL, the entire ML/DL model can be refined over time. The FL server picks a subset of clients at the start of each round to take part in the learning process and shares its most recent global model with them. The contributions of the proposed secure localization based on blockchain technique in IoT-based wireless sensor network are bulleted as follows:

(1) Explore blockchain techniques for detecting and localizing malicious nodes in IoT-based wireless senor networks

(2) Enhance detection and localization of malicious nodes in IoT-based wireless sensor networks to improve service quality

(3) Explore the effectiveness of federated learning techniques using performance metrics to evaluate the attack detection performance

(4) Provide secure routing and data transmission in a hierarchical IoT-based wireless sensor network for secure service provisioning

(5) Effective detection and localization of malicious nodes by computing the unknown nodes based on the blockchain technique with the help of beacon node

(6) Secure data aggregation and trust value computation of beacon nodes using a trust management model based on blockchain for malicious node removal and ensuring service provision

(7) Design and plan secure range-free localization processing using blockchain technology-based selection of trust value of the miner beacon nodes

(8) Explore the blockchain secure data aggregation and localization of unknown nodes for malicious node detection and localization using federated learning approaches

The other parts of this paper are organized in the following sections. Section 2 contains previous literature works on secure blockchain-based localization in IoT-based wireless sensor networks. Sections 3 and 4 detail the network models, proposed methods, and simulation techniques of the scheme using blockchain techniques. Sections 5 details simulation results and analysis. Section 6 is about performance evaluation and validation followed by Section 7 containing conclusion and future work.

2. Related Works

Several related works and techniques have been reviewed for detecting and localizing malicious nodes in IoT-based wireless sensor networks. The schemes in the literature include the following:

(i) Secure localization of malicious nodes in low power sensor node networks

(ii) Routing schemes for effective resource consumption and network life time

(iii) Authentication and encryption processes for establishing trust among nodes

(iv) Blockchain-based localization technique for improving quality of services in WSNS

(v) Techniques for achieving optimized and secure data storage in IoT-WSNs

(vi) Secure and optimized models for the next-generation IoT-based WSNs

(vii) Federated machine learning and classification techniques

Kim et al. [1] conducted research on blockchain techniques based on a trust management model for detecting malicious nodes by enhancing relationships among beacon nodes using various evaluation metrics, including closeness, honesty, detection rate, average energy consumption, frequency of interaction, and intimacy. The beacon nodes are selected based on the trust values and broadcast information to the base station, forming a trust evaluation model for detecting and localizing malicious nodes in WSNs. Abubaker et al. [4] presented blockchain-based for detection and localization of malicious nodes in IoT-based wireless sensor networks using federated random forest and support vector machine techniques. They also discussed secure service provisioning using feature evaluation and cascading encryption for detecting and removing malicious nodes in the network. Security and performance metrics, including accuracy, node honesty, and end-to-end delay of the packet transmission, measured the performance of the malicious node detection and secure provisioning approaches. Ming et al. [5] suggested an authenticated group key agreement mechanism using blockchain technology for the Internet of Things. The suggested protocol includes a novel concept called the device manager, whose job is mediating communications between Internet of Things gadgets and blockchain infrastructures. The proposed protocol has been shown to be secure after being subjected to various assaults, as shown by the security analysis. The simulation results demonstrate that the time expenditures of protocol operations are fair and appropriate for IoT settings. Kumar et al. [7] proposed a new distributed intrusion detection system (IDS) leveraging fog computing for detecting DDoS assaults against mining pools in blockchain-enabled IoT networks. Random forest (RF) and a gradient tree boosting system (XGBoost) are trained on fog nodes, and their performances are compared. A real-world IoT dataset is used to evaluate the proposed model's performance, containing the most recent assaults discovered in a blockchain-enabled IoT network. The system for detecting DDoS attacks was connected to a mining pool in a blockchain-enabled IoT network. Three distinct mechanisms power the suggested distributed detection system. First, there is the traffic processing engine, which employs fog nodes to preprocess network data by standardizing its features with the help of the aforementioned Standard Scaler. Goyat et al. [11] examined a secure and novel range-free localization-based blockchain technique in wireless sensor networks against a malicious node in two-dimensional environments. They also discussed the implementation of blockchain technology using beacon nodes' trust value for the mining process of blocks to form a localization process for the unknown nodes with neighbor node list, mobility, residual energy, and repudiation value. The location and precise positions of the unknown nodes are computed using the trusted beacon nodes based on the blockchain-based generated trust values. The malicious nodes broadcast false localization information by compromising the beacon and reporting the false energy information. Awan et al. [12] proposed blockchain-based encryption and trust evaluation model for detection and secure data transmission. The authentication of the sensor nodes and aggregator nodes is in a private and public blockchain, respectively, for detection of the activities of malicious nodes, including transmitting wrong information

of routing and energy by compromising the aggregating node that degrades the network performance by increasing resource consumption and packet loss. The trust values and residual energy of the sensor nodes were computed after authentication based on the detection rate for removing the malicious nodes from the secure routing in the network. Honar Pajooh et al. [13] proposed a multilayer-based distributed and decentralized blockchain security model for a secure IoT model. IoT devices operate under multihop cellular networks using a self-clustering hybrid evolutionary algorithm combining simulated annealing and genetic algorithm approaches by clustering the unknown sensor nodes in the network. The cluster head authenticates and authorizes locally for effective communication between the cluster head and the base station using a private lightweight balanced blockchain with better throughput and network latency. Otoum et al. [14] presented an adaptive framework using integrated blockchain and federated learning using reinforcement support vector federated learning for a secure trust evaluation process in the network. The reinforcement learning occurs at the end devices, and the models were trained for communication with fog and creating a global model. They also evaluated and verified the model through simulation techniques using performance metrics, including trust value, energy consumption, and network lifetime. The model also achieved a detection rate and accuracy of 96% and 93%, respectively, for detecting and localizing malicious nodes using federated learning. She et al. [15] proposed a blockchain trust model framework for detecting malicious nodes in IoT-based wireless sensor networks by constructing blockchain data structure in three-dimensional space using a localization scheme. The detection of malicious nodes is realized using intelligent contact and wireless sensor networks quadrilateral measurement to localize and detect attacks. The sensor nodes are authenticated and recorded in the blockchain network to remove and prevent loss of routing information. Fan et al. [16] proposed a decentralized topology-based blockchain scheme for secure IoT systems during synchronization for detection and analysis attacks using multiple time resources. The scheme solved the problem of time announcement in IoT-based distributed wireless sensor networks using a closed blockchain for recording and broadcasting time for attack detection with adaptive topology for reducing communication overhead. The scheme also avoids centralization using block structure adaptable topology and multiple time sources. Friha et al. [17] presented secure IoT infrastructure using federated learning based on intrusion detection system agricultural applications for data protection producing an improved attack detection model. The scheme utilized a convolutional neural network, recurrent neural network, and deep neural network using benchmark datasets to detect and classify attacks and performance metrics. They also discussed the performance metrics, including recall, precision, F1-score, and accuracy for the multiclass classification federated deep learning model. Javaid [18] examined and designed the Dijkstra algorithm-based routing protocol for secure and efficient communication between the sensor and sink nodes without hole attacks in IoT-based WSNs. The scheme also provided transparency for transaction by the sensor nodes based on the network's blockchain technique for detecting malicious nodes. The transactions were added into blocks using the poof of authority consensus technique to validate the trust model in terms of the energy and detection metrics. Wu and Ansari [19] proposed blockchain-based secure and trust evaluation process for detection of malicious nodes in industrial Internet of Things for information sharing and access control list. The security of access control list is protected and stored by the blockchain. They also designed a voting technique with trust evaluation for access control list for high probability of malicious node detection and authentication in the network. Khan et al. [20] described the implementation of a low-energy, energy-temperature-degree-adaptive, hierarchical routing protocol. One advantage of the protocol is that it reduces the amount of energy used by nodes during data transmission, which in turn increases the lifespan of the network. The cluster heads (CHs) that will be used to carry out routing are chosen by the proposed protocol based on their degree, temperature, and energy. More importantly, the blockchain is used to eliminate the risk of a centralized failure point. Since several nodes are required for a blockchain to function, the CHs and BSs use it to store and record all of the data transactions that take place between them. Blockchain transactions are hashed using the 256-bit secure hashing method (SHA-256). To top it all off, the ETD-LEACH protocol uses a real-time message content validation (RMCV) mechanism to identify rogue nodes before they cause any damage during routing.

Goyat et al. [21] addressed IoT-WSN security issues and trusted architecture of secure localization attacks. In particular, it examines the trust evaluation method and the process of creating a blockchain. Trust values of each beacon node are calculated using several trust metrics, and the weights associated with those values are dynamically changed during the localization process. That is why only the most reliable beacon nodes are chosen during the mining procedure. This two-stage process guarantees an always up-to-date blockchain regarding reliability and the trust values of all beacon nodes. Minoli and Occhiogrosso [22] proposed and implemented a method of protection in the context of a defense-in-depth/castle approach based on blockchain mechanisms (BCMs) playing a role in protecting numerous IoT-oriented applications by becoming part of a security mosaic. A blockchain is an immutable digital ledger that records each completed transaction or piece of data in sequential order in a distributed ledger that outside parties cannot alter. All of the users in the network then have access to these exchanges. Every node in the system keeps the same ledger as all other nodes in the network; the information is saved and published in a public ledger. A full-fledged blockchain-secured network may not always be practical for all IoT applications. The typical limitations of IoT nodes are critical or institutional applications such as smart grids, ITSs, e-health, insurance, and banking that may require nodes with sufficient capabilities to support the necessary peer-to-peer functionality. Ma et al. [23] presented a blockchain-based

secure sharing scheme using ITS as an example for IoV data. It uses smart contracts to enable features like automated registration, speedy authentication, and a reliable sharing mechanism for IoV data. The private information is protected through homomorphic encryption and zero-knowledge proof processing performed by the smart contract and stored as ciphertext on the distributed ledger. We use a tamper-proof Merkle tree-based block to permanently record all IoV data processing and usage procedures and a PBFT consensus mechanism to guarantee the integrity of the entire network's ledger. The IoV chain method maintains the integrity, accessibility, and tamper resistance of the IoV data under the premise of privacy and security, allowing for the data to be traced if and when required.

Yang et al. [24] addressed the issue of centralized anomaly detection in WSNs, and the authors describe a blockchain-based ensemble anomaly detection (BCEAD) system that stores the model of a common anomaly detection method on the distributed ledger. Once a suitable block structure and consensus method have been developed, the scheme changed iteratively to increase detection and classification accuracy. The detection approach is safe from system-level attacks since the blockchain guarantees a trustworthy network environment. In conclusion, the findings demonstrate that BCEAD is superior to competing schemes in terms of performance, cost, and other parameters. Abnormality detection across all network domains is the responsibility of the sink layer. A decentralized ledger stores the isolation forest detection model (tangle). In this study, we also compare and contrast experimental data to show that BCEAD has better detection performance than previous methods. Sarhan et al. [25] presented a federated learning architecture built on a blockchain-based hierarchy to ensure that collaborative IoT intrusion detection is safe and private. The proposed machine learning-based intrusion detection system employs a federated learning architecture with a hierarchical structure to safeguard the confidentiality of the training and company information. There are several potential advantages to using machine learning (ML) skills in the security against IoT cyberattacks. However, the existing framework proposals do not consider data privacy, safe architectures, or scalable deployments of IoT ecosystems. A trusted distributed ledger will manage all business processes and transactions, and the smart contract will ensure that all obligations are met. Hsiao and Sung [26] proposed a method that employed blockchain technology to improve the safety of wireless sensor networks' data transmissions (WSNs). This research combines distributed ledger technology (blockchain) with data transmission to establish a highly reliable WSN infrastructure. Today's wireless network is founded on IoT design principles and uses a blockchain-based approach to ensure the integrity of all transmitted data. The blockchain-based transaction ledger is adapted in this research strategy to store information gathered by sensors. The suggested system collects and analyses sensor data to improve the dependability of the wireless sensing network architecture. To establish a connection between blocks, every blockchain must undergo cryptographic processing and the decryption of public keys. Even more so, before any data can be transmitted in a newly created blockchain, the prior and subsequent sequences of the link must be confirmed. Abbas et al. [27] presented a lightweight blockchain-based authentication solution to store the credentials of standard sensors. Because IoT nodes have a certain amount of time before they die from battery exhaustion, lightweight authentication is achieved by storing only a small number of credentials in the blockchain. Furthermore, a genetic algorithm-enabled software-defined network controller performs the route calculation and is also utilized for on-demand routing to minimize the nodes' energy usage in an IoT network. In addition, we present a route correctness technique to validate the absence of harmful nodes in the suggested path. Abualsauod [28] proposeda peer-reviewed research on UAV IoT frameworks and to analyse it to find answers to problems with the frameworks' overall security and privacy. An optimal solution is provided in the paper for the many security and reliability issues that arise in UAV-enabled IoT applications by integrating different ways based on blockchain technologies. Multiple metrics, including processing speed, latency, accuracy, and overall system usefulness, are used to evaluate and contrast the acquired results.

3. IoT-Based Network Model

The network model comprises various types of wireless sensor nodes operating under different computational processing and clustered into ten regions. Each region is partitioned into clusters, as shown in Figure 3. The nodes are classified into sink nodes, cluster heads, and sensor nodes based on the computational power and data processioning capabilities. The sensor nodes are assumed to be randomly deployed and are unaware of their location and position. At the same time, the beacon nodes and sink nodes are aware of their locations and positions, and they assist the unknown sensor nodes in locating the sensor nodes and identifying the malicious nodes by computing the distance between each node and its position, with each node having its own distinctive identifier [29]. The network model incorporates challenging design requirements and assumptions for efficient performance. The requirements include the following:

(i) Randomly deployed sensor nodes are assumed to be homogeneous, with mobility changing in the topology framework

(ii) Each device in the Internet of Things has its distinct IP address

(iii) The storage and processing power of the cluster's head nodes and base stations, which can be used to implement a smart contract, are described

(iv) The existence of a continuous malicious node in the network makes the framework overconscious

(v) All sink nodes are trustworthy and generate keys to all the nodes in the network

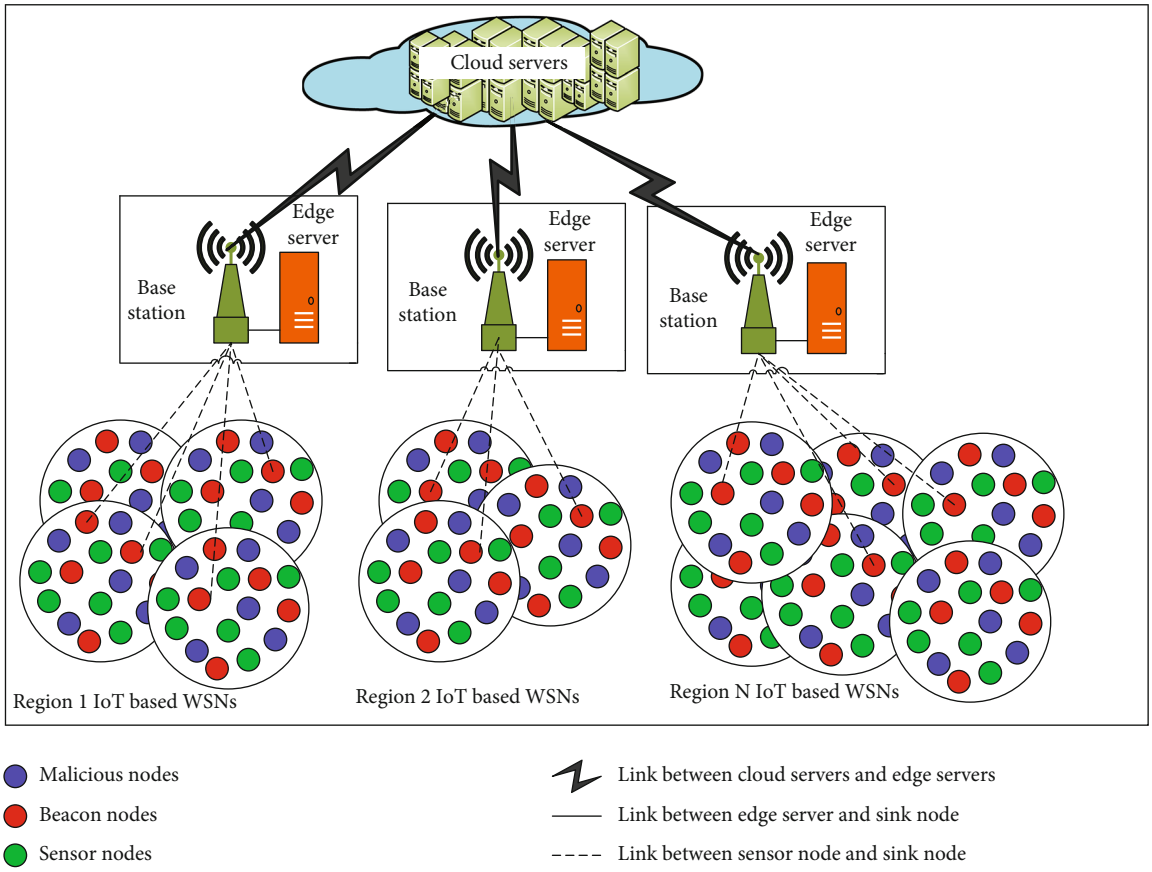


Figure 3: Proposed hierarchical IoT-based wireless sensor network model for secure localization of malicious nodes.

(vi) All sink nodes have high computational and data processing capabilities for data aggregation

(vii) The servers provide reliable and secure storage for the data in the network

The two requirements in IoT-based wireless sensor networks make the sensor nodes broadcast malicious information by gaining control over the other nodes and wrong information about energy and localization that affect the precise localization accuracy. This paper proposed blockchain-based secure localization using a trust evaluation process and cascading encryption for effective network transaction and malicious node detection. The randomly deployed sensor nodes face security threats during the localization process since malicious nodes affect the localization approach's accuracy in IoT-based wireless sensor networks. The blockchain-based technique overcomes this problem through the cascading encryption and effective transaction and trust evaluation process of the beacon nodes. The blockchain receives the trust value list and localizes the malicious nodes using the distributed consensus.

The nodes that send data to the cluster head without knowing their location are unknown and disrupt the normal functioning of the network. To avoid this problem, beacon node having their precise location helps find the accurate location of unknown nodes of their deployment before sending data. The reliable beacon nodes are selected by computing the trust value of beacon nodes without considering the malicious node. The presence of the malicious node is solved by using blockchain-based localization techniques using federated learning without compromising the beacon node and privacy. In this proposed model, each region has its network with multiple clusters for monitoring the environmental conditions. Nodes in the network cooperate to accomplish specified tasks, and the interaction between nodes needs to assure the legitimacy of node identity [30]. One of the crucial ways to safeguard the IoT is through identity authentication. In the network model presented here, two-factor authentication is required to set up encrypted channels between nodes before any communication can occur, or end users can have access to the resources hosted by a node.

4. Proposed System

The proposed system for detecting and localizing malicious nodes, as shown in Figure 2, consists of various phases for identifying and removing attacks in IoT-based wireless sensor networks using hybrid and federated learning-based blockchain technology. The proposed scheme ensures client and service provisioning using hybrid machine federated machine learning models, feature evaluation, and cascading

encryption for effective transaction and localization. The proposed system utilized federated machine learning model for detecting and classifying malicious nodes by training and testing the dataset received from the blockchain trust value lists with suitable data format using preprocessing techniques. In a WSN, the localization process relies on the accuracy of information relayed by beacon nodes; therefore, ensuring their reliability can be difficult [31]. Since IoT-based WSNs' range-free localization method does not rely on any specialized hardware, it is susceptible to mistakes when malicious attacks are present in the network. During the localization procedure, the malicious nodes provide false location data. There is a single point of failure in typical localization strategies for WSNs because they rely on a centralized body. For attack detection and localization, the efficacy of the trust evaluation based on secure blockchain using hybrid federated machine learning, reinforcement learning, and maximum likelihood estimation needs to be put to the test. One application of the blockchain is in the realm of trust administration. Furthermore, just a small set of criteria are employed to determine whether or not to trust a beacon node. For this reason, the trustworthiness of both behaviors and data must be factored into the ultimate trustworthiness assessment. From there, the data is gathered and processed by the blockchain infrastructure that was provided to it by the trusted nodes.

4.1. Sensor Deployment and Routing. The proposed system consists of N number of total sensor nodes deployed with random mobility in two-dimensional environments with B beacon nodes and U unknown nodes. The base stations are located outside the clusters in each region connected to the edge servers. The beacon's network size and communication radius are computed as in the following equations shown below.

$$|N| = |B| + |U|. \tag{1}$$

The transmission range (T_R) of the beacon nodes is computed with the maximum ($T_{\max}$) and minimum ($T_{\min}$) communication ranges as shown below:

$$T_R = (T_{\max} - 1) + \text{random}(0, 1) \times [(T_{\min} - 1) - (T_{\max} - 1)] + 1. \tag{2}$$

Randomly deployed wireless sensor networks face security threats by computing the location estimation and position of unknown nodes in the network. Detecting malicious sensor nodes provides information for calculating the location and enhancing the localization accuracy in WSNs based on beacon nodes. The beacon node is selected from the sensor nodes using the selection three criteria:

(i) The minimum distance of the node from the base station computed using the distance-vector protocol

(ii) The residual energy of the node for activation

(iii) The strength of the received signal from the sensor node

The number of neighbor nodes is close to it. Using the distance vector method, as indicated in equation (3) below, we are able to calculate the distance, denoted by D, that separates any two sensor nodes.

$$D = \sqrt{(u_i - u_j)^2 + (v_i - v_j)^2}. \tag{3}$$

Any node that has a small distance from the base stations (where u and v are the coordinates for nodes i and j, respectively) is probably the cluster head.

Threshold parameters are determined using the multipath model strategy, which is then used to assess the energy expenditure involved in network model communication and activation. In the steady-state phase, each cluster head (CH) communicates with the other nodes to keep their receivers on and respond to its CH [32]. For a data transmission of k bits across a distance D and a threshold distance D_o, we have (4) as the energy required for transmission E_{TX}:

$$E_{\text{TX}} = \begin{cases} k \times E_e + k \times E_f \times D^2, & \text{if } D \le D_o, \\ k \times E_e + k \times E_m \times D^4, & \text{if } D > D_o. \end{cases} \tag{4}$$

For one-bit data transmission, we get where E_{TX} is the energy being sent, E_f is the energy being received, and E_e is the power being lost somewhere along the line [29]. Signal propagation, filtering, modulation, and channel coding all play a role in the amount of energy lost in the process. With a data length of k, we have the following expression for the minimum acceptable transmission distance, d_o:

$$D_o = \sqrt{\frac{E_f}{E_m}}. \tag{5}$$

Energy used by the receiving node for k-bit messages is calculated as

$$E_{\text{RX}(k)} = k \times E_e. \tag{6}$$

4.2. Selection of Miner Node and Localization Process. Methods for computing behavioural trust, data-based trust, and feedback trust, and as well as calculating the overall trust beacon nodes in WSNs, are used. The most trustworthy beacon nodes are picked based on their trust levels when trust is calculated for each node. For mining and trilateration, these beacon nodes are trusted. Our blockchain network's PoA consensus is based on this group of miners. On the sink and beacon nodes, a private blockchain is being used in this network. On the blockchain, the trustworthiness beacon nodes are permanently recorded. In addition, the network's nodes encrypt the trade data and store it on the blockchain. The miner node is the anchor node having its own location with the most significant trusted values for verifying communications and putting the blocks for the next transactions. Figure 1 depicts the architecture of the proposed network model containing block. After this, the translation process begins. Trilateration is used to determine the precise

location of each unknown node. Here, unknown nodes determine their approximate location by calculating the distance between themselves and the three anchor nodes with the highest trust ratings and then utilizing that information to solve for their coordinates using the following equation:

$$D_{\text{tri}} = \begin{cases} \sqrt{(u_a - u_i)^2 + (v_a - v_i)^2} \\ \sqrt{(u_a - u_j)^2 + (v_a - v_j)^2}, \\ \sqrt{(u_a - u_k)^2 + (v_a - v_k)^2} \end{cases} \tag{7}$$

where D_{tri} is the trilateration process for finding the distance of any node a, and (u_a, v_a), (u_i, v_i), (u_j, v_j), and (u_k, v_k) are the positions of the ordinary sensor nodes with respect to i, j, and k beacon nodes.

Finding the nodes' locations and strengths of connections, as well as evaluating their positions, is tasks for which specific algorithms are required. The system is broken down into two types of categories, range-based and range-free localization methods [29]. The second option is more practical financially, but it needs specialized hardware. Accurate positioning and positioning of wireless sensor nodes are evaluated using the received signal strength indicator (RSSI) and the distance vector hop localization techniques. To determine where sensor nodes and cluster leaders are in relation to beacon (base station) nodes, a distance vector localization process is required [33]. Position and separation between the unknown nodes can be calculated and manipulated with the help of the scheme. Beacon nodes in a WSN can be located using the distance vector hop process. The average hop size is used in the distance vector approach to determine the minimal distance. In [34], this algorithm was discovered for the first time. Here are the steps of the range-free distance vector localization scheme [35]:

4.2.1. Routing Initialization. The beacon node transmits a message to all sensor nodes. As soon as it enters the network, the hop count associated with its position data is set to 0 [35]. A node that gets such a message will determine where the beacon node is located, append its own identifier to the message, and then broadcast it to its neighbours while increasing the hop count [36].

4.2.2. Calculating Distance. With the help of the beacon node, we can determine the average hop size and the distance to the unknown node.

4.2.3. Position Estimation. Geometric calculations like triangulation, the polygon technique, and trilateration are frequently employed as the underlying basis for determining the locations of unknown nodes [29]. There are a number of methods that can be used to calculate the separation of two nodes; they include synchronization, radio signal intensity, and the physical features of the carrying wave [37].

4.3. Registration and Authentication Process. Sensor nodes, aggregation nodes, and base stations are the three main categories of WSNs in the proposed network model [12]. Intelligent communications confirm the aggregation, and data processing node exists by using its MAC address and the base station's authentication procedures to verify its identification as in Figure 4. In order to gain access to the model and the public blockchain that is being utilized, the sensor nodes need to be authenticated and given permission [38]. It is essential that the privacy of the nodes be preserved, and in order to do this, our model made use of a layered security architecture. The IoT-WSN framework that has been presented needs to have the capability of establishing trust between different IoT members in order to identify and localize threats. Node placements in wireless sensor networks are shown in Figure 4(a).

In Figures 4(b) and 4(c), we can see that the cluster head (CHs) gathers a large message size, while the sensor nodes (SNs) spend a disproportionate amount of time doing data execution during deployment, registration, and authentication. The cluster heads that are location and position aware are regarded as beacon nodes. The data forwarding, retrieval, processing, and aggregation to the base station are shown in Figure 4(d).

In WSNs, the aggregated node's data and the public blockchain records of certified nodes make it possible to have faith in authentication techniques. After the registration procedure has been finalized, sensor nodes are given permission to join the blockchain, which helps to protect WSNs from outside threats. After being randomly dispersed around the playing field, the sensor nodes eventually form aggregation nodes. As the aggregation nodes communicate with the sensor nodes, they verify and identify sensor nodes via a private blockchain, and the base station verifies the identity of the aggregate node via a public blockchain. Mutual authentication is used for all of the aggregate nodes' internode communications.

4.4. Benchmark Datasets. The evaluation of the proposed system's ability to detect and classify the attack type using machine learning models is conducted against a benchmark consisting of the CICIDS2017 and UNSW-NB15 sample datasets. The dataset was captured and executed using tools for intrusion detection and network attack scenario purposes with seven categories of attacks, as shown in Figure 5(a), with the frequency distribution of the attack samples. The dataset contains normal and common classes of attacks that materialize real-world data, including network traffic analysis. The dataset is also heterogeneous, having diversified attacks with 80 feature sets generated using CICFlowMeter available online.

The UNSW-NB15 is an advanced benchmark dataset for intrusion detection widely used for recent research and many previous works. The raw data traffic packets were generated using the cyber range laboratory IXIA PerfectStorm tool from the Australia Center for cyber security (ACCS) [39], creating hybrid normal and abnormal network traffic packets. The Tool simulates nine classes of attacks and continuously updates the information security vulnerabilities and exposures of captured packets. The dataset consists of 42 attributes, three categorical features, and 39 numerical attributes. The dataset has ten classes of attacks, and the

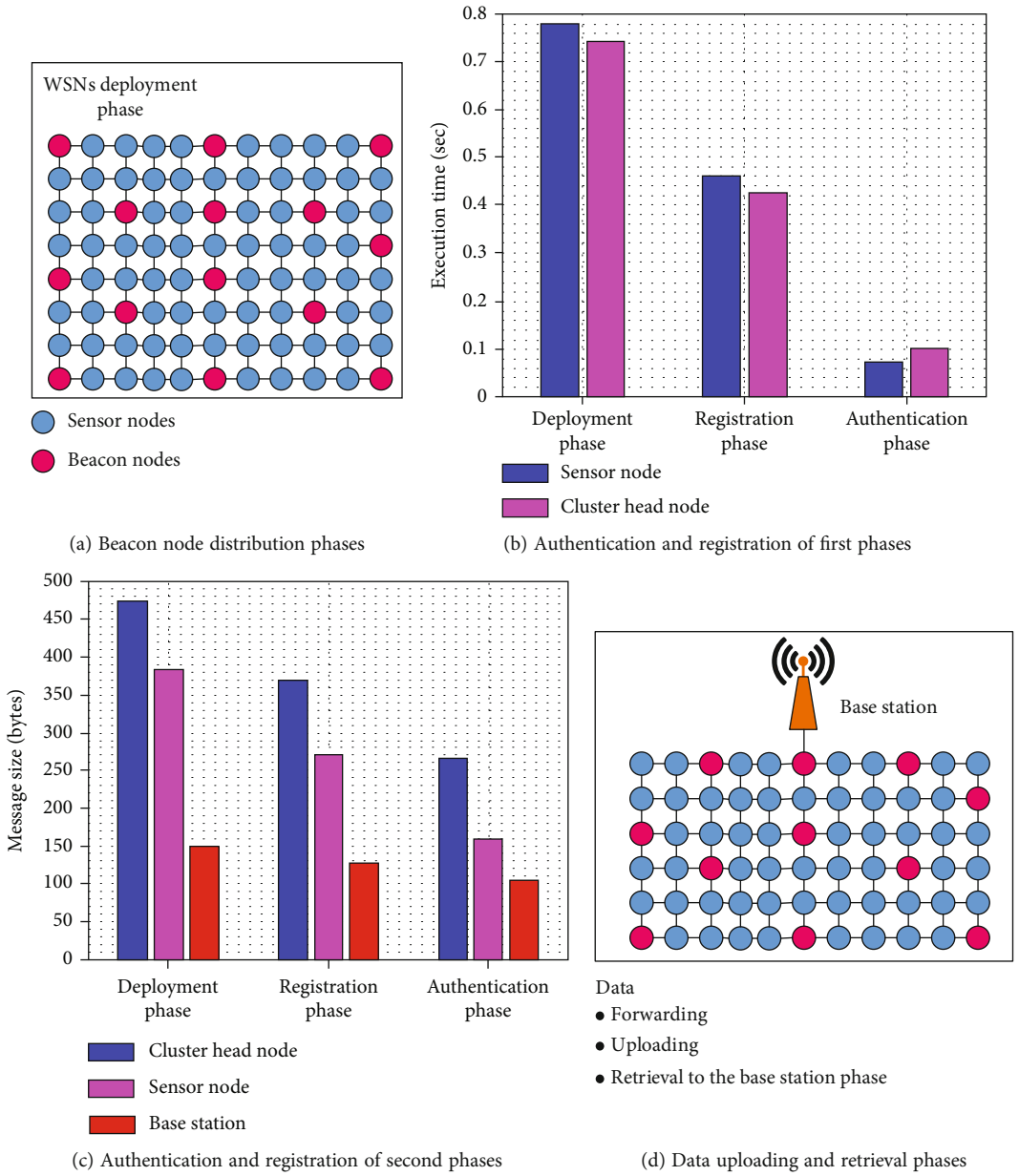
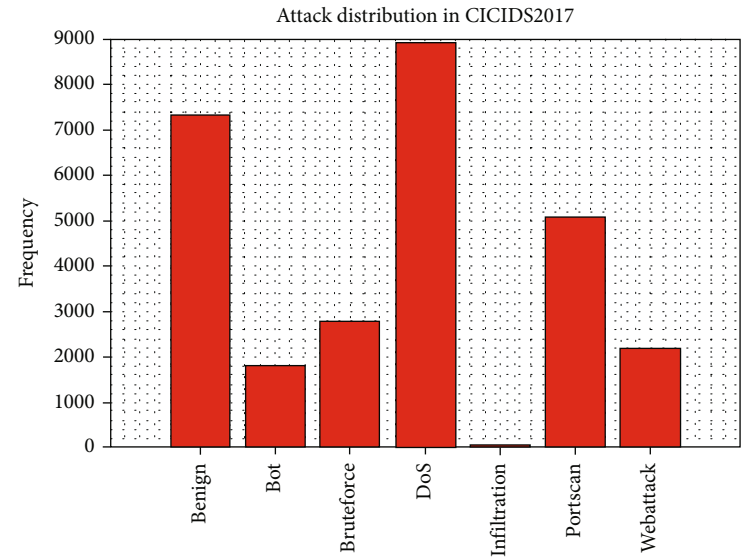


FIGURE 4: Secure data collection and transfer via IoT-WSNs that entail multiple steps, including authentication, registration, and transmission.

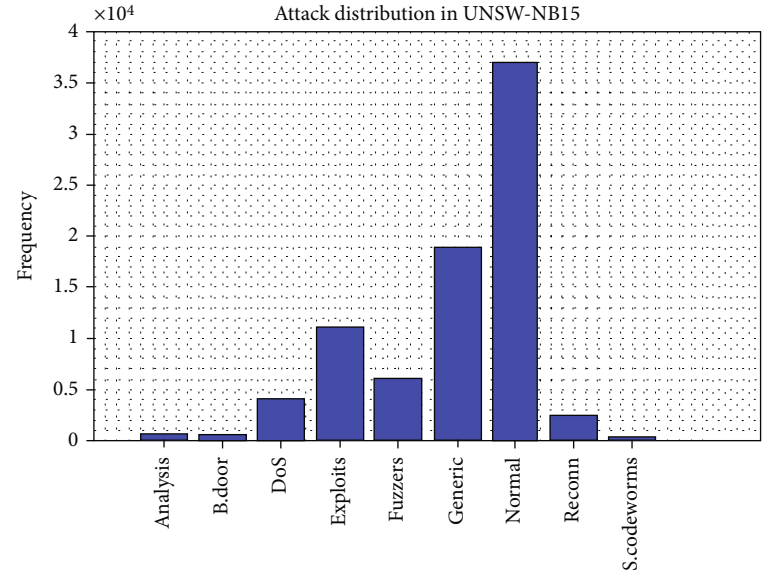
frequency distribution of these attacks is depicted in Figure 5(b). The dataset is divided into 80% training and 20% testing samples for building the predictive models for classification and regression tasks using machine learning models.

4.4.1. Data Preprocessing and Feature Selection. The traffic data can be collected from a local network or the benchmark dataset using the Data collection module [40]. The preprocessing module receives these data and filters them before they are used. These details are also supplied to the cluster and the trust-based safer routing module for more efficient data transmission. The data preprocessing agent makes excellent use of available data preparation methods. Data cleansing, integration, and modification are also undertaken before preprocessing. The data preprocessing step starts with data cleaning and normalization to improve the data quality for training and testing for building the predicting machine learning models. The preprocessing is important for vectorizing, consisting of feature extraction, feature selection, and dimensional reduction [41]. The dataset can be cleaned by removing the duplicated values, replacing the mission data, and removing and fixing unwanted structures from the samples. After cleaning the dataset, it is essential to normalize using the minimum and maximum scaling values as in the following equations:

$$Z_{\text{norm}} = \frac{Z - \min(z)}{\max(z) - \min(z)}, \tag{8}$$



(a) Frequency distribution of attacks in CICIDS2017



(b) Frequency distribution of attacks in UNSW_NB15

Figure 5: Attack occurrences in the CICIDS2017 and UNSW_NB15 benchmark datasets were plotted as a frequency distribution.

where min (z) is the minimum value, and max (z) is the maximum value of the attribute Z, respectively. Z_{norm} is a normalized feature value, and Z is an original feature value [42].

K-means cluster sampling is utilized to improve the classification and detection accuracies of the machine learning model by generating a small K-number of clusters of the original dataset to reduce the training complexity [43]. The K-means sampling technique generates highly representative small groups by removing redundant data to increase efficiency, computational power, and resource. The synthetic minority oversampling process (SMOTE) creates high-quality samples to avoid class imbalance. After the data preprocessing technique, a feature engineering technique is applied to produce sensitive, high-quality features, reduce dimensionality, and remove redundant features by computing the correlation features between features.

4.5. Federated Learning and Security. Currently, the federated cybersecurity (FC) model communicates and collaborates at several levels to provide security solutions for IoT applications. However, privacy and security risks are associated with the standard method of transmitting data within the cluster and across different groups. As a result, federated learning (FL) has emerged as a viable option for exchanging data and information in a method that is both secure and private as in Figure 6. To ensure the security of the IoT network, an FC model that works in tandem with FL to share and discuss data at any level is essential. Most FC strategies based on FL as a cyberdefense mechanism have concentrated

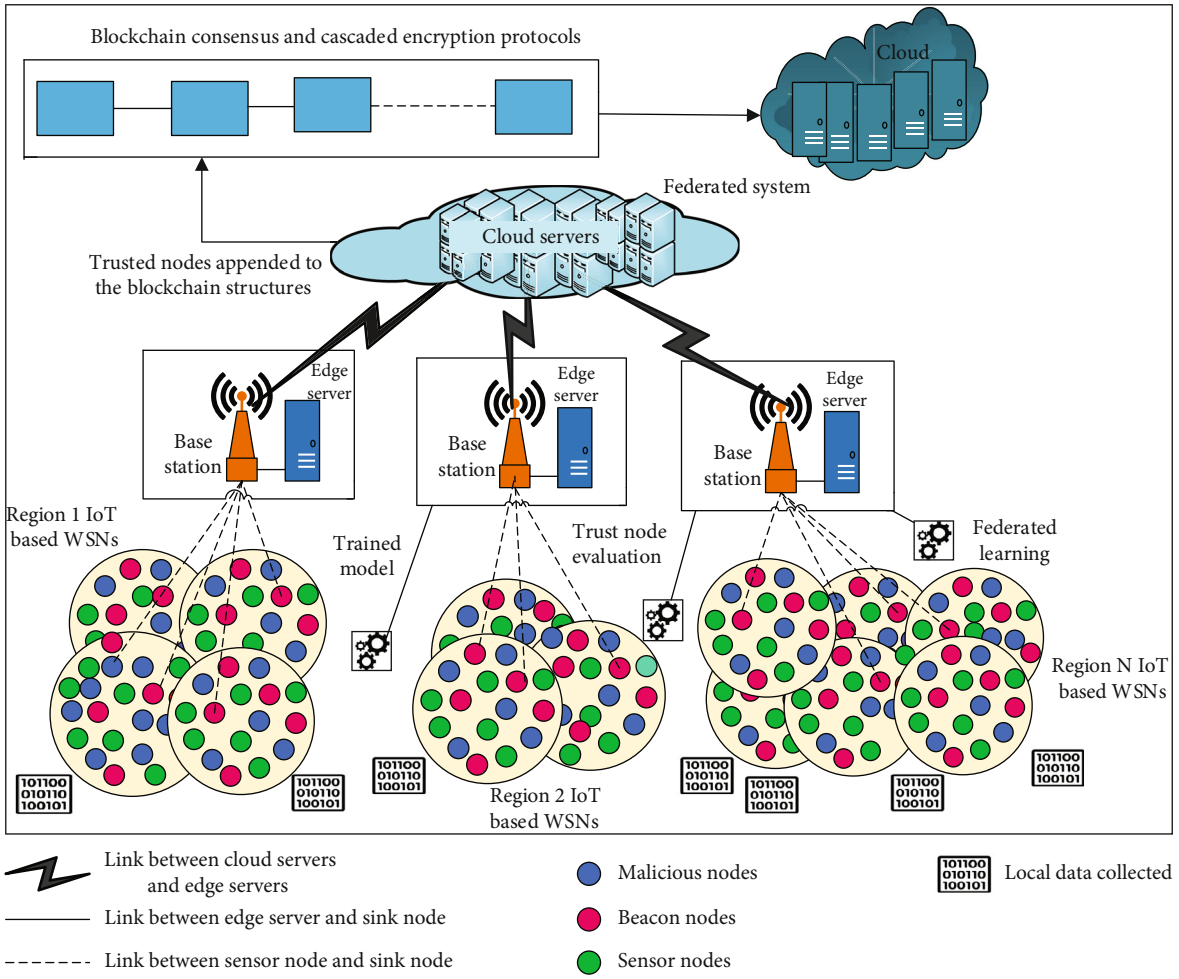


Figure 6: Secure hybrid federated IoT-based WSN model for detection and localization of malicious attacks.

on protecting IoT networks while assuming a unified, worldwide service offering. The method has been developed for a single global model managed by a single service provider; however, it may be readily generalized to a collaborative scenario with several global models managed by numerous service providers. Only a tiny amount of study has developed a federated security model that integrates different international approaches [44]. Machine learning techniques identify resource allocation, security and privacy, communication cost, localization of malicious nodes, and enhancing network lifetime in IoT-based WSNs [45].

Some of the learning techniques utilized in this work include a support vector machine, XGBoost, random forest, and ensemble stacking to improve the performance and evaluate the effectiveness of the proposed system using various evaluation metrics. They are effective and practical techniques for knowledge extraction and finding the relations in wireless sensors that make detecting malicious nodes in the network more accessible. Trusted nodes valuated by the beacon nodes and base station are allowed to enter the block structure as shown in Figure 6 in IoT-based WSNs.

Machine learning models are statistical algorithms with the ability to infer meaningful insights about future data by learning complex patterns in existing data [25]. These ML models are designed using a training procedure in which the learning models glean relevant meanings from the accessible data logs. Training log sets in an IoT ecosystem are produced locally by a wide variety of digital endpoints including smartphones, smart appliances, sensors, and the like. The efficacy of ML-based IDSs is directly impacted by the quality and quantity of data utilized in training and evaluating learning models. The first type, known as localized learning, involves training a model locally using data sets generated by a single endpoint; the second, known as centralized learning, involves training a model centrally using data sets collected from multiple endpoints, and the third, known as federated learning, involves training a model across multiple endpoints using local data sets without exchanging them.

4.5.1. Random Forest. Random forest (RF) is an ensemble learning model that operates by constructing many decision trees for predicting a class [46]. Random forest manages high-dimensional data by judging the importance of the features for solving overfitting and stability problems by reducing the variance. This makes the random forest technique effective for detecting and classifying malicious attacks using

a heterogeneous benchmark dataset in wireless sensor networks. The random forest also handles the missing values and is robust to outliers with less effect for nosy data. The scheme also needs more computational power and more resources due to the construction of many trees. Random forest constructs a forest of several decision trees to predict the model's detection accuracy [47]. It is successful in detecting and classifying attacks using a benchmark dataset.

4.5.2. Gradient Boost. Extreme gradient boosting (XGBoost) is a classification technique for large datasets with a minimum amount of time, making it popular nowadays [48]. The gradient boosting technique uses the extraction of essential features to improve the computational speed and provides a precise output for intrusion detection by reducing memory consumption during the training and testing of the dataset for classification [49]. It is an efficient machine learning technique for optimizing the loss function and computed features, as shown below:

$$\phi(\mathrm{X}) = \sum_{k=1}^{K} f_k(X) \quad f_k \epsilon F, \tag{9}$$

where $\phi(X)$ is the final result of the K sequential classifier and f_k is the decision tree for the K number of iterations in the gradient descent algorithm. The technique also performs parallel computing for getting results in classifying the target results. The XGBoosting improves the gradient descent and regularization strategy for optimizing and managing the process and overfitting factor. The parameters of the classifiers are related as in the following equation shown below:

$$\mathscr{L}(\phi)_{\mathrm{t}} = \sum l(f_{t-1} + f_t) + \Omega(f_t), \tag{10}$$

where $\mathscr{L}(\phi)_t$ is the loss function and $\Omega(f_t)$ is the regularization term to optimize the step size t. The gradient boosting technique uses feature metrics to retrieve the scores according to the features for each attribute.

4.5.3. Ensemble Learning. Ensemble machine learning techniques are classifiers with averaged accuracy to reduce the risk of overfitting and bias from a single classifier [50]. The classification machine learning models are set of tree structures and enhance the overall accuracy. Ensemble techniques are meta-algorithms combining various machine learning techniques into a unified machine learning prediction model for evaluating variation and stacking [51]. The ensemble technique integrates different machine learning results into a single robust model and improves the overall performance. Stacking, bagging, and boosting are the essential ensemble technique concepts. In this scheme, stacking is used for combining and increasing the prediction power of the machine learning models.

4.5.4. Support Vector Machine. The term "support vector machine" (SVM) refers to a family of statistical methods for categorizing network activity under the guidance of an experienced teacher [52]. The fundamental purpose of a support vector machine (SVM) classifier is to build a hyperplane in the feature spaces using a set of vectors called support vectors. Here, we use a dispersed binary classifier for normal and abnormal, which enables the detection of any harmful behavior.

$$w = \sum_{i=1}^{n} a_i y_i x_i \quad \min \left\{ \frac{\|w\|^2}{2} + \mathrm{C} \sum_{i=1}^{n} \varepsilon_i \right\}. \tag{11}$$

$\sum_{i=1}^{n} \varepsilon_i$ is the bounds on the learning vectors, C is the constant that determines the tradeoff between the number of false positives and the margin, and a_i is the Lagrange multiplier. The input and output variables, x and y, are for the demo data set.

Support vector machine (SVM) is a technique for classification, outlier detection, and regression of data samples using hyperplane mapping the data into feature space [47]. It is used for multiclass classification techniques using nonlinear data as a benchmark for evaluation. The SVM technique is used to detect and classify attacks in WSNs to improve the performance and detection accuracy of the system. The SVM has some key features, including more training time, nonlinear data usage, and a higher false alarm rate. It is most beneficial for the classification and prediction of performance evaluation.

4.6. Optimization Techniques. To further enhance the classification of machine learning models for the proposed method on the benchmark dataset, we combine two hyperparameter optimization strategies with Bayesian optimization using a tree based on the Parzen estimation (BO-PTE). Hyperparameters are used by every machine learning method in order to fine-tune the settings and achieve the best possible results. Machine learning performance is enhanced while human labour is reduced thanks to hyperparameter optimization (HPO) [53]. Hyperparameter optimization is also utilized in a black box and global optimization for enhancing functional evaluation. This enables us to describe the black box Bayesian optimization techniques. Bayesian optimization (BO) is a framework for global optimization with expensive black box functions gaining attraction in HPO for deep neural networks. Bayesian optimization is an iterative technique with an acquisition function and probabilistic surrogate model evaluating the decision-making points by applying the Gaussian process. The hyperparameters are handled by tree-based methods, including random forest and tree Parzen estimators (PTE). In this proposed work, Bayesian-based optimization is used with tree Parzen estimators (BO-PTE) for evaluating the best point of evaluation for automatic machine learning.

5. Simulation Results

The validation of the proposed system model is presented using simulation results with different simulation scenarios. The simulation results confirmed that the proposed scheme effectively detects and localizes malicious nodes in IoT-based wireless sensor networks. Our suggested approach employs a

Table 1: Simulation setup for the proposed network model.

Parameter	Values	Parameter	Values
Software	MATLAB	Total sink nodes	10
Number of sensors	1600	Number of clusters	53
Total unknown nodes	1480	Sink position	500, 2000
Protocol type	Clustering and routing	Attacks	Routing
Deployment area	$20000 \times 20000\,m^2$	Mobility	Random
Total beacon nodes	110	Transmission radius	250
Total edge servers	10	Data size	4000 KB

trust evaluation mechanism to pinpoint malicious sensor nodes in the network. In addition, we have included an authentication method to prevent unauthorized access to our network. An energy-efficient and real-time routing system between SNs and BSs is also proposed. Based on authentication, the proposed model is contrasted with the current paradigm. Since authentication's impact is difficult to see in real-time, we measure it in terms of how long a network lasts, how much power it uses, and how fast it can transfer data. In addition, Solidity is employed in the creation of the smart contract. PoW and PoA are used together as part of a consensus process to verify the integrity of the network. The implemented algorithms are compared based on their gas consumption and transaction delay [12].

In our experiments, we used Python 3 and MATLAB as the primary language of code [17]. Popular libraries like NumPy, which facilitates the manipulation of multidimensional arrays and matrices, have been used to implement the proposed strategy. These matrices and arrays are used in the analysis and classification of attacks. Pandas python library provides easy access to advanced data structure modification and analysis tools. Frameworks like TensorFlow and Keras have become increasingly common in deep learning and machine learning. Both supervised and unsupervised ML techniques can be developed with the scikit-learn toolkit. SMOTE was created to oversample people from underrepresented groups systematically. The network planning and simulation of the models were conducted using MATLAB R2021a running on a machine having a Processor Intel(R) Xeon(R) Silver 4214 CPU @ 2.20GHz 2.19 GHz (2 processors), Installed RAM 128 GB (128 GB usable), windows ten 64-bit operating system, and ×64-based processor. Python is also used for data processing and analysis. Table 1 shows the parameters of the simulation setup for network deployment and performs a comprehensive security analysis. Solidity programming language is also employed for creating smart contacts and consensus processes. Both proof-of-work and proof-of-audit are used as network-wide consensus techniques for validation. Measures such as gas consumption, transaction delay, geolocation accuracy, and classification precision are used to evaluate the developed algorithms.

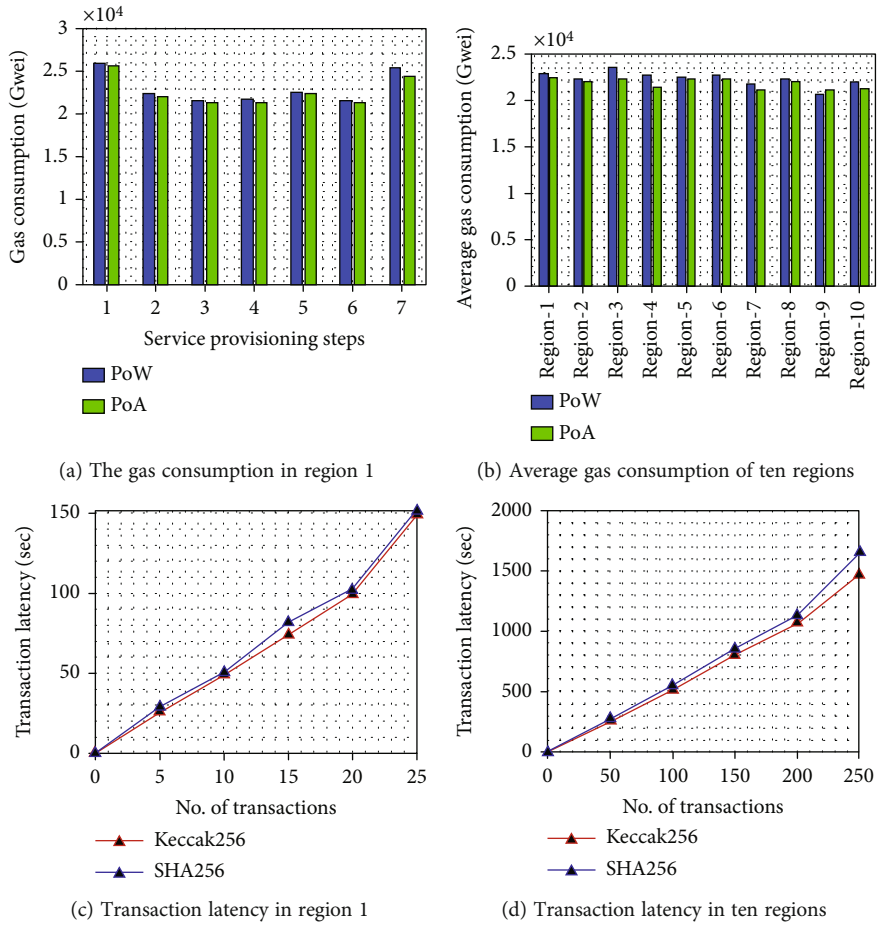
The proposed localization technique uses blockchain for detecting malicious attacks, and the models are deployed into ten independent IoT-based regions for IoT-based WSNs. The PoW and PoA consensus algorithms are used in the region for validating the transactions and adding blocks for detecting malicious nodes and service provisioning. Minor nodes validate the transactions and add blocks for localization of the unknown nodes so that attacks can be removed. This blockchain-based technique enables localizing routing attacks that degrade the network performance and lifetime.

Table 1 also contains the simulation parameters for the models we have proposed. Ten distinct IoT-WSN areas have implemented in this proposed blockchain-based localization for detecting malicious attacks and various simulation parameters for the proposed models. Predesignated nodes are liable for approving transactions and putting blocks onto the blockchain under these models, which use the PoA consensus method Figure 7(a) illustrates how IoT-WSN gas usage varies by power level and area in region 1 IoT-WSNs for service provisioning. For example, Figure 7(b) compares the average gas consumption of the IoT-WSNs in all ten regions to that of PoW and PoA. PoW uses more gas than PoA, as evidenced by these numbers. This is because in proof-of-work (PoW), the miner nodes in charge of validating transactions are chosen through a rigorous mathematical process. On the other hand, in PoA, miners are selected based on their stakes rather than their qualifications. A group of miners is chosen to validate and run blockchain transactions and add new blocks.

5.1. Evaluation Metrics. A trust value is determined for each beacon node to choose the most reliable and secure beacon nodes [4]. According to the trust value calculation, the network has numerous nodes that do not know their specific location. These nodes cannot locate themselves without the help of nearby beacon nodes why they should be the most reliable and trustworthy in a network. The confidence of each anchor node is calculated using the triangulation process and distance vector routing to identify the most reliable beacon nodes. Behavioural, feedback, and trusted data values contribute to the total trust level. Four criteria are used to measure behavioural trust:

(1) Frequency of interaction: anchor nodes' interactions with other anchor nodes are counted as the rate of interactions

(2) Friend nodes and list of neighbors: the ratio of the total amount of one-hope nodes in any beacon node



(a) The gas consumption in region 1

(b) Average gas consumption of ten regions

(c) Transaction latency in region 1

(d) Transaction latency in ten regions

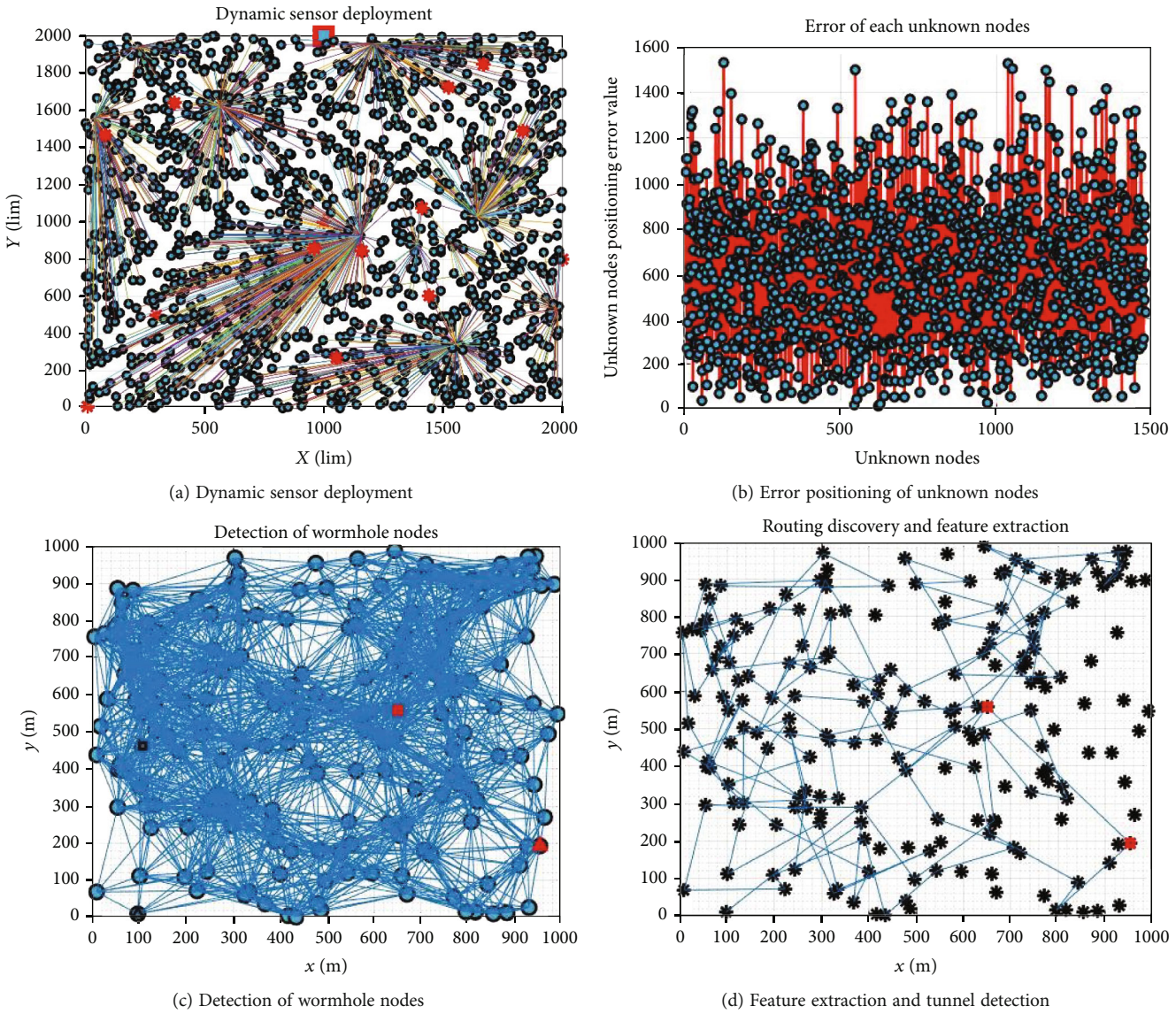
FIGURE 7: The gas consumption and transaction latency of different regions in IoT-based wireless sensor networks.

to the whole number of nodes except for this particular beacon node is an essential statistic for calculating behavioural trust. It is the quickest route from the beacon node to the rest of the network. The number of nodes in the vicinity of anchor node is called its range

5.1.1. Localization Accuracy. The performance of the proposed system is measured using localization error (LE), average localization error (ALE), localization accuracy of the unknown sensor nodes, and the confusion matrix taking the CICIDS2017 benchmark dataset for classification of attacks. The localization error measures how far away the estimated position is from the true position [54]. To facilitate a standard means of comparison, localization errors are normalized to the radio range of sensor nodes. The average localization accuracy measures how precisely the nodes in the study area are located [55]. The average error localization (ALE) [34] is computed as in equations (12)-(14), respectively. The average detection rate, accuracy, precision, and recall are also used to evaluate the effectiveness the proposed system. In order to calculate the ALE, add up the LE of all the unknown nodes and divide by the total number of unknown nodes [29].The LE represents the discrepancy between the estimated and true locations of any undiscovered nodes.

$$\begin{aligned} \mathrm{LE} &= \sqrt{\left(u_i' - u_i\right)^2 + \left(v_i' - v_i\right)^2}, \\ \mathrm{ALE} &= \sum_{i=1}^{n} \frac{\sqrt{\left(u_i' - u_i\right)^2 + \left(v_i' - v_i\right)^2}}{\mathrm{nR}}, \\ \mathrm{ALA} &= \left(1 - \left(\sum_{i=1}^{n} \frac{\sqrt{\left(u_i' - u_i\right)^2 + \left(v_i' - v_i\right)^2}}{\mathrm{nR}}\right)\right) \times 100\%, \end{aligned} \tag{12}$$

where (u'_i, v'_i) and (u_i, v_i) are the real and computed coordinates of the anonymous node I, respectively; n represents unknown nodes, and R represents the communication radius of the network. In Figure 8(b), we can see the deviation of the unknown sensor nodes (b). Using the localization scheme, the positions and deviations of all nodes are calculated. The performance of the proposed scheme is evaluated using the performance metrics including detection rate, false



(a) Dynamic sensor deployment

(b) Error positioning of unknown nodes

(c) Detection of wormhole nodes

(d) Feature extraction and tunnel detection

Figure 8: Dynamic deployment and clustering of sensor nodes for computing unknown node positioning in IoT-based WSNs.

alarm, precision, and recall. Equations (15)-(17) provide a mathematical representation of this:

$$\begin{aligned} \text{Accuarcy} &= \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}, \\ \text{Detection rate} &= \frac{\text{TP}}{\text{TP} + \text{FN}}, \\ \text{F1} - \text{score} &= \frac{2 \times \text{TP}}{2 \times \text{TP} + \text{FP} + \text{FN}}. \end{aligned} \tag{13}$$

Accuracy determines how well the learning model functions are [56]. We put the proposed method to the test by doing experiments and comparing the outcomes to those obtained using an established secure blockchain based on federated hybrid machine learning models. To accomplish this, we must first determine the proportion of attacks that were correctly classified as true positive, trusted nodes that were correctly classified as true negative, false positives that were incorrectly classified as true negative, and false negatives that were incorrectly classified as true positive.

The detection rate is a measure of correctively classified normal traffic and is regarded as a positive value over the total samples in the dataset. The detection rate is the same as the actual positive rate computed in equation (14). The precision is the number of correctly classified attacks over the total number of instances of attacks in the networks.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \tag{14}$$

where true positive (TP) is the number of attacks correctly classified as attacks [57, 58], and false positive (FP) is the number of attacks incorrectly classified in the network [59]. The true negative (TN) is the legitimate nodes organized as

legitimate, and the false negative is the honest nodes incorrectly as malicious nodes in the network traffic.

The detection accuracy is the fraction of trusted sensor nodes that correctly identify malicious nodes relative to the total number of harmful nodes in the network [6]. When compared to the false negative rate (FNR), which is the total number of malicious nodes that are incorrectly classified as legitimate nodes by the suggested model, it is computed by the following equation:

$$\mathrm{FNR} = \frac{\mathrm{FN}}{\mathrm{FN} + \mathrm{TP}}. \tag{15}$$

5.1.2. Intimacy. It is one way to gauge the behavioural trusted values of the anchor nodes. It is the time consumed that anchor nodes spend for communicating. It is computed by using the following formula:

$$\mathrm{In} = \frac{t_i}{t_i + t_a}, \tag{16}$$

where t_a is the time consumed for node to communicate with the a anchor node and t_i is the collaboration time of a particular node with beacon i.

5.1.3. Honesty. Another metric that is used to calculate behavioural trust is honesty. The equation used to calculate it is presented below.

$$h_{\mathrm{o}} = \frac{P_s}{P_n}. \tag{17}$$

There are P_n interactions between source and destination nodes and P_s interactions between source and destination nodes that result in successful outcomes. Truthfulness tells the overall amount of interactions between nodes, both successful and unsuccessful. The interactions take into account end-to-end delay for classifying interactions among nodes. The amount of time it takes, on average, for data packets to be transported from source nodes to sinks, including both valid and faulty data packets, is referred to as end-to-end latency [60]. It is the amount of time required for a data packet to travel from its point of origin to the location where it will be stored permanently.

When malicious nodes are inside the radius of any anchor node, the behavioural trust in those beacons is compromised. Beacon nodes' overall trust is impacted, resulting in an incorrect selection of anchor node and cluster head nodes for accurate localization malicious nodes in WSNs. The number of neighbour nodes, distance error, security, mobility, and localization techniques all plays a role in determining where rogue nodes can be found in WSNs [61].

After calculating behavioural trust, each beacon node's feedback trust is determined. A beacon or an unknown node must supply the other with its coordinates to figure it out. After verifying that the sender beacon node gave it the correct coordinates, this node provides feedback on the sender beacon node. Last, each beacon node's data-based trust is calculated by comparing the actual distance of the beacons to their anticipated distance. Equation (18) gives the Euclidean theorem to calculate the exact distance for estimating distance.

$$D_{\mathrm{ac}} = \sqrt{\left(u_i - u_j\right)^2 + \left(v_i - v_j\right)^2}, \tag{18}$$

where D_{ac} is the actual distance, u_i and v_i denote the positions of the sender anchor node, and u_j and v_j are the positions of the destination anchor node.

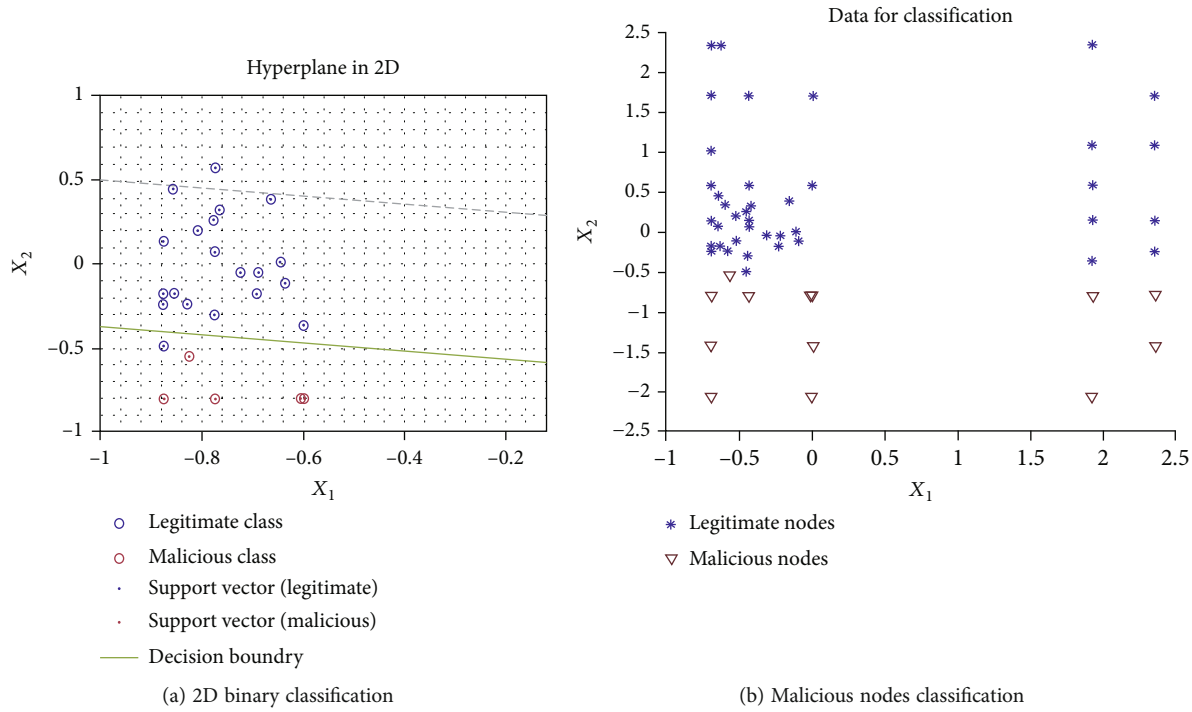
The novelty of the proposed methodology is scalability and chain of trust evaluation using intelligent communication of the wireless sensor nodes within a given region. The security performance of the system is independent for each area; hence, the DoS attacks cannot manipulate and access multiple religions simultaneously.

6. Performance Evaluation and Validation

The evaluation and validation of the proposed system are computed through simulating and classification of hybrid and federated machine learning models. Our simulation results demonstrate the efficacy of the proposed system's federated-based machine learning algorithms for identifying, localizing, and classifying malicious nodes in the simulated network. The proposed model has been simulated to determine its gas consumption, delay, time response, data aggregation, remaining network energy, and dropped packet count.

The practical steps of our model are deployment, as shown in Figure 8(a) and computing in Figure 8(b), for localization error of unknown nodes, registration, authentication, routing calculation, and malicious node detection and classification using hybrid federated machine learning models. The wormhole attacks were detected by computing the distances and position of the sensor nodes marked with red colors as in Figure 8(c). Wormhole tunnels were created and detected using routing discovery and feature extraction between two malicious nodes as depicted in Figure 8(d).The simulation result shows that the proposed system achieved the localization accuracy of the routing attacks is 99.14% by training the extracted features as shown in Figure 8(d). This shows that the localization accuracy of the wormhole attack is effective compare to Kim et al. [1] with localization accuracy of 82.17%.

As a result of this research, the network's malicious nodes have been located and deleted. XGBoost, random forest, extra tree, and ensemble stacking are utilized for evaluating the performance and test the effectiveness of the proposed system utilizing benchmark datasets with various evaluation metrics to classify genuine and malicious nodes. Hyperplanes are used by SVM to classify data points into one of two categories. As a result of this model, a line (hyperplane) is generated that splits the data samples into two distinct categories. A hyperplane is constructed for comparing these data point. The SVM model then determines the class of each data point. SVM's primary goal is to build a decision border that is effective for separating the data samples of each course, as shown in Figure 9(a), for malicious node classification as in Figure 9(b).



(a) 2D binary classification

(b) Malicious nodes classification

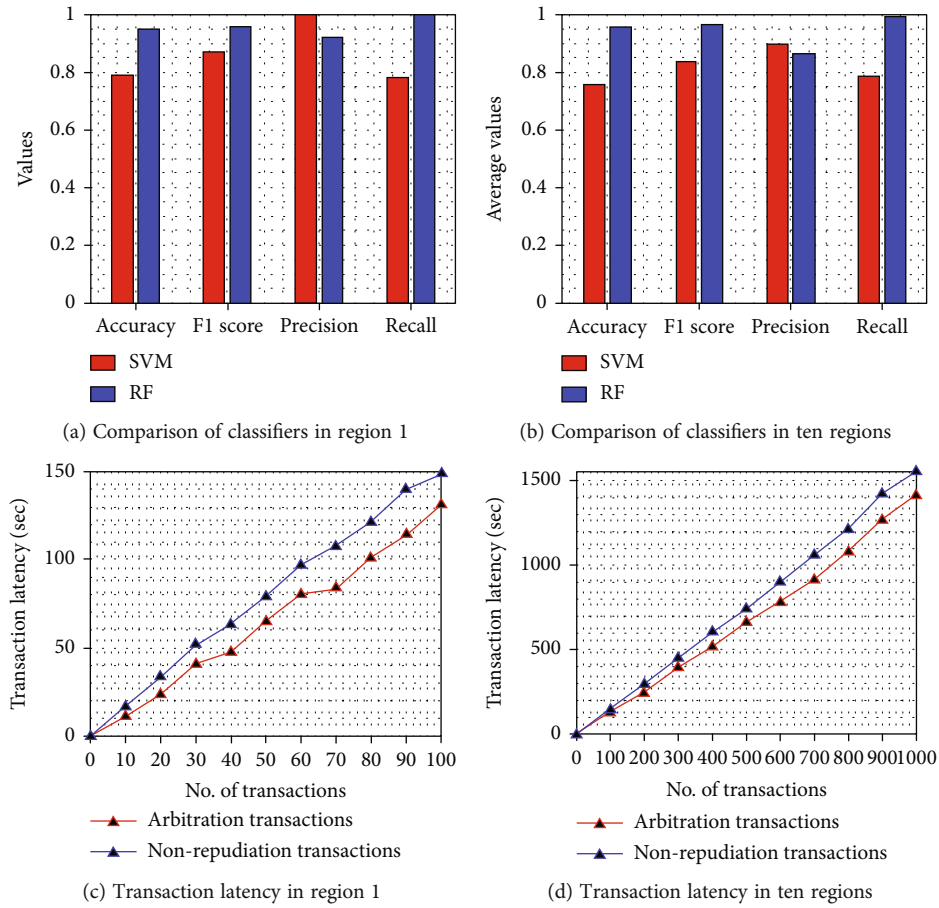
Figure 9: Classification of legitimate nodes from malicious nodes using decision boundary with SVM.

The dataset is imported and labelled when the honesty and end-to-end delay calculations are completed. After that, it is separated into datasets for training and testing. Following the division of the dataset, the model is trained to predict the node categorization based on its predictions. The benchmark dataset demonstrates the SVM model's extensive methodology for detecting malicious nodes.

The results also show that the proposed approach is efficient for malicious node detection using the blockchain technique for ten independent and nonrepudiation regions in IoT-based wireless sensor networks. Preselected network nodes, as depicted in Figure 7(a) and Figure 7(b), in the next generation's IoT-based wireless sensor networks will use the proof-of-authority (PoA) consensus process to validate transactions and add blocks to the block. The majority of consortium blockchains use the PoA consensus technique. To add new blocks to the blockchain, PoA chooses a set of validators. Validators are chosen according to their standing in the system. Since validators can be preselected using PoA, it uses far fewer computing resources than PoW. Contrary to PoW, the PoA consensus technique does not include any form of mining. Proof-of-work (PoW) is a consensus technique utilized by public blockchains like Bitcoin and Ethereum. A cryptographic problem is broadcast over the network to determine a miner in proof-of-work (PoW). The node chosen will be the one that figures out the solution to the challenge quickly and accurately. The chosen miner verifies the transactions and creates the block hash. The hash is then sent across the network to establish an agreement. Figure 7(a) compares the gas consumption in the secure service provisioning method between PoW and PoA in region 1 of an IoT-based wireless sensor network. The gas consumption is a small amount of cryptocurrency, as shown in Figure 7(a). The amount is deducted from the user's account performing the transaction on the Ethereum blockchain [62]. The cryptocurrency is removed from the user account. Figure 7(a) shows the gas consumption of events of the secure service provisioning process in gas units.

Similarly, Figure 7(b) compares the average gas consumption of IoT-based wireless sensor networks between PoW and PoA in all ten regions. PoW uses more gas than PoA, as evidenced by these numbers. This is because in proof-of-work (PoW), the miner nodes in charge of validating transactions are chosen through a rigorous mathematical process. On the other hand, in PoA, miners are selected based on their stakes rather than their qualifications. A group of miners is chosen to validate transactions and add new blocks to keep the blockchain running. Since all transactions depend on a limited number of miners, the PoA process weakens distributed nature of the blockchain. This is partly because PoA makes blockchain networks more centralized. The transaction delay of Keccak-256 and SHA-256 hashing algorithms in region 1 IoT-based wireless sensor networks and across all ten IoT sensor networks is shown in Figures 7(c) and 7(d). The secure hashing method has been expanded, and the Keccak256 hashing algorithm is the result (SHA3). There is a built-in implementation of the Keccak256 algorithm in the Solidity programming language. Keccak256 may hash any data into a hexadecimal value of a predetermined size.

Because the hashing process is irreversible, this hash cannot be used to retrieve the original data. When compared to other hashing algorithms like SHA256 and RIPEMD160, Keccak256 offers a lower cost profile. Hence, we choose to



(a) Comparison of classifiers in region 1

(b) Comparison of classifiers in ten regions

(c) Transaction latency in region 1

(d) Transaction latency in ten regions

Figure 10: Performance compression of classifiers and transaction latency using blockchain technique for detecting and localizing malicious nodes in IoT-based WSNs.

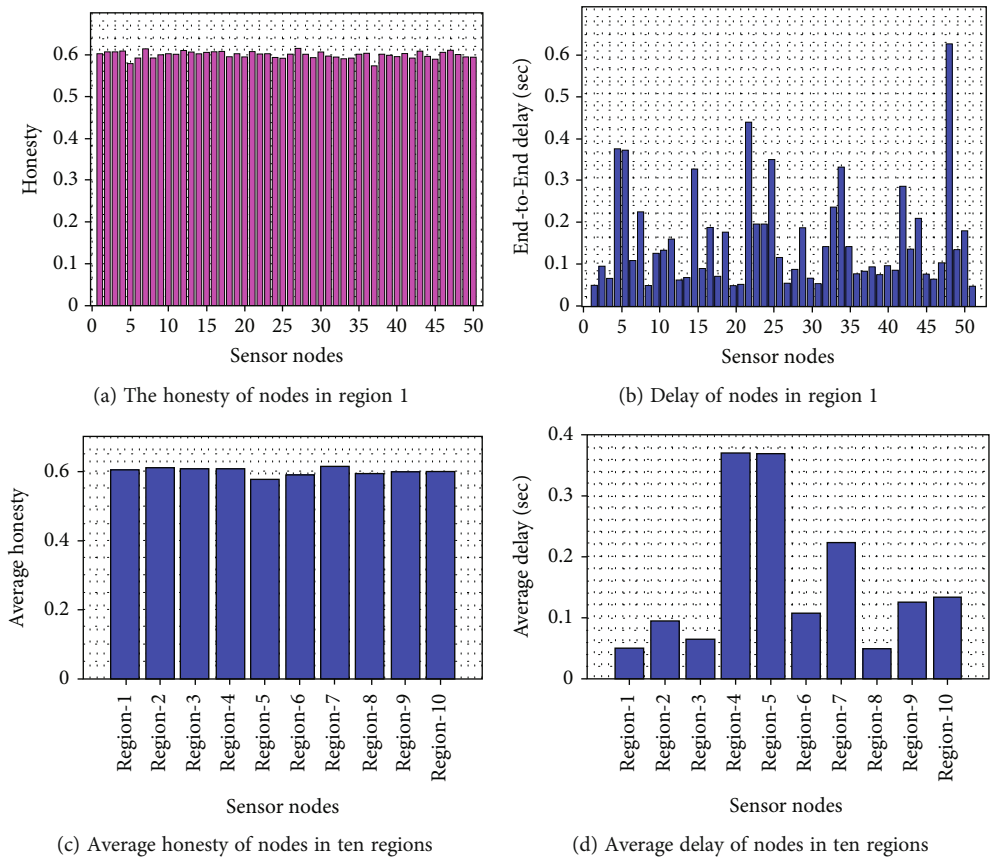
employ it. Network sensors in region 1 of Figure 7 perform 25 transactions per second, while all Figure 7 sensors perform 250 transactions per second (d).

Figures 7(c) and 7(d) depict a network with one and ten areas of IoT sensors, respectively. SHA-256 and Keccak-256 hashing methods behave identically in both figures. It can be deduced from the data that SHA-256 takes almost twice as long as Keccak-256 to execute. SHA-256 lacks an infinite input space, which makes it easier for Keccak-256 to do hashing. A nonrepudiation model based on a support vector machine, XGBoost, random forest, and ensemble stacking classifiers is effectively used for scalable and large networks. The model also provides effective actions for detecting and localizing malicious nodes in the network model using federated machine learning techniques. Linear separability in the SVM hyperplane allows for categorizing data points into two distinct categories.

As a genuine class, the one with the highest honesty value and the least delay from beginning to end is considered the best candidate. However, malicious nodes have a high end-to-end delay and a poor integrity rating. The recall, precision, accuracy, and F1 score of RF and SVM in IoT-based wireless sensor networks are shown in Figures 10(a) and 10(b). IoT wireless sensor networks with rogue nodes can be detected more effectively by using radio frequency identification (RFID) technology. Figure 10(a) shows that SVM has an accuracy of 89%, but RF has an accuracy of 99%.

It shows that support vector machine (SVM) has an average accuracy of 86%, and RF has an average accuracy of 99.3% in all ten regions of interest. According to the findings, RF outperforms SVM in detecting malicious nodes across several regions' networks based on honesty and end-to-end latency. Ensemble learning multiple decision trees are another explanation for RF's better accuracy. Many trees are produced for training and testing to solve the same problem of identifying malicious nodes. Figure 10(a) shows that the F1 score of SVM in a single region is 87.95%, while RF has a 96% F1 score. As depicted in Figure 10(b), SVM and RF have an average F1 score of 84% and 97% in all ten regions of IoT based WSNs, respectively. Because of this, hostile nodes in sensor networks, whether in a single part or over numerous areas, are less likely to be mistaken for legitimate nodes via RF.

In addition, as shown in Figures 10(a) and 10(b), SVM's recall score is lower than RF's because RF is better at predicting harmful nodes. In addition, these results show that SVM's precision is superior to RF's. As a result, SVM can correctly identify the positive observations thanks to its support vector. Our suggested nonrepudiation model was shown to be effective in all ten clusters of the regions of



(a) The honesty of nodes in region 1

(b) Delay of nodes in region 1

(c) Average honesty of nodes in ten regions

(d) Average delay of nodes in ten regions

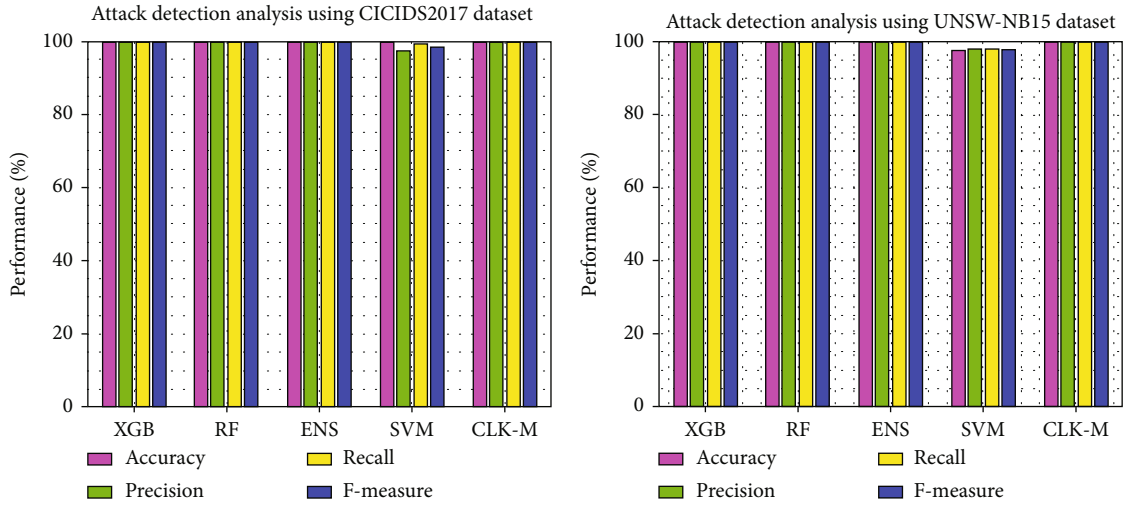
Figure 11: Honesty of wireless sensor nodes and end-to-end delay of nodes in IoT-WSNs.

IoT-WSNs, as shown in Figures 10(c) and 10(d). A varied number of disagreements are reflected in the numbers. Transaction latency is the time it takes a network to process a transaction. Client and service provider latency for arbitration transactions is also depicted. Both the service provider and the customer are found to be vicious in these transactions. According to the diagrams in Figure 10, our suggested system can process a high number of transactions in a short period. As a result, no data is exchanged between the features indicated and the service supplied during these transactions, allowing for a fair comparison. It demonstrates that the arbitration transactions use the bare minimum of network resources including throughput and bandwidth. In addition, the results reveal that the nonrepudiated transactions have relatively low latency. This proves that our suggested nonrepudiation strategy is effective to protect a large number of users over a wide area of IoT-WSNs while being efficient in terms of quality of services and security.

Nodes from region 1 IoT-WSNs are shown to be honest in Figure 11(a). From the region 1 IoT-WSNs, a random sample of fifty nodes is chosen randomly. Shown in the graph are various levels of trustworthiness for each node. For this reason, each node has a varied number of successful interactions. All ten areas' IoT-WSNs' honesty is calculated in the same method, as illustrated in Figure 11(c). Nodes with a high honesty rating are more likely to be legitimate.

This indicates that the node has many interactions and connections with other sensor nodes in WSNs. Honesty is a critical criterion for determining the authenticity of nodes in our networks. Nodes' validity can also be verified by looking at the end-to-end delay, transmission delay, propagation delay, processing delay, and queuing delay and can all be calculated. The malicious nodes are those that have end-to-end delays that are higher than the predefined threshold value. Nodes are not responding quickly enough when other nodes send them messages. Furthermore, their limited processing skills cannot process the data packet, resulting in service denial.

There is a high end-to-end delay with this type of node that broadcasts this information (latency). Figure 11 shows the end-to-end latency of region 1 IoT-WSNs and the average end-to-end delay of all ten regions' IoT-WSNs (Figures 11(b) and 11(d)). Figure 11 depicts the random selection of fifty ordinary nodes from each location. Figure 11(b) shows that the end-to-end delay for each node varies. This is because each node responds to requests in a different period. The malicious nodes in the network are those that have a higher end-to-end delay than a predetermined threshold. Another interesting finding from Figure 11 is that different IoT-WSN region networks have varied range of honesty and end-to-end delay average values, demonstrating the performance of different IoT-WSN region networks in providing secure services.



(a) Hybrid machine learning models using CICIDS2017

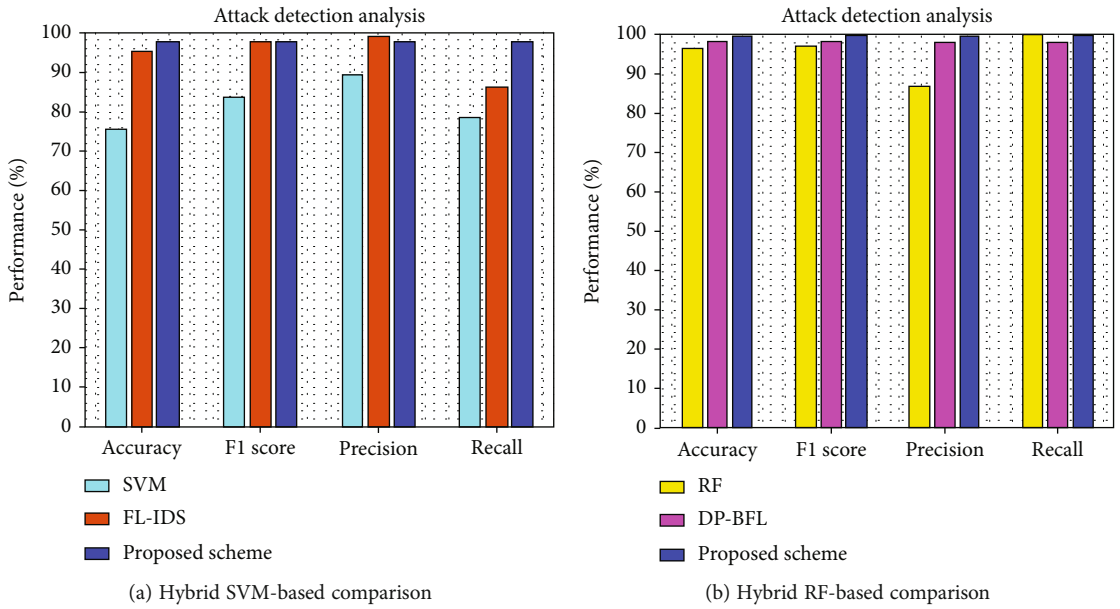
(b) Hybrid machine learning models using UNSW_NB15

Figure 12: The utilization of CICIDS2017 and UNSW_NB15 benchmark datasets for comparison performance of several machine learning models is conducted based on a variety of performance measures.

6.1. Attack Detection Analysis. The performance of the proposed system is evaluated using machine learning models using a benchmark dataset containing a class of attacks in WSNs. The hybrid optimized machine learning also utilizes the same benchmark dataset to evaluate the proposed routing attack localization and detection effectiveness in wireless sensor networks. Figure 12(a) shows the comparative performance of the various machine learning schemes. The performance of the proposed system is further improved with the application of cluster labeling (CL) K-means binary classification techniques.

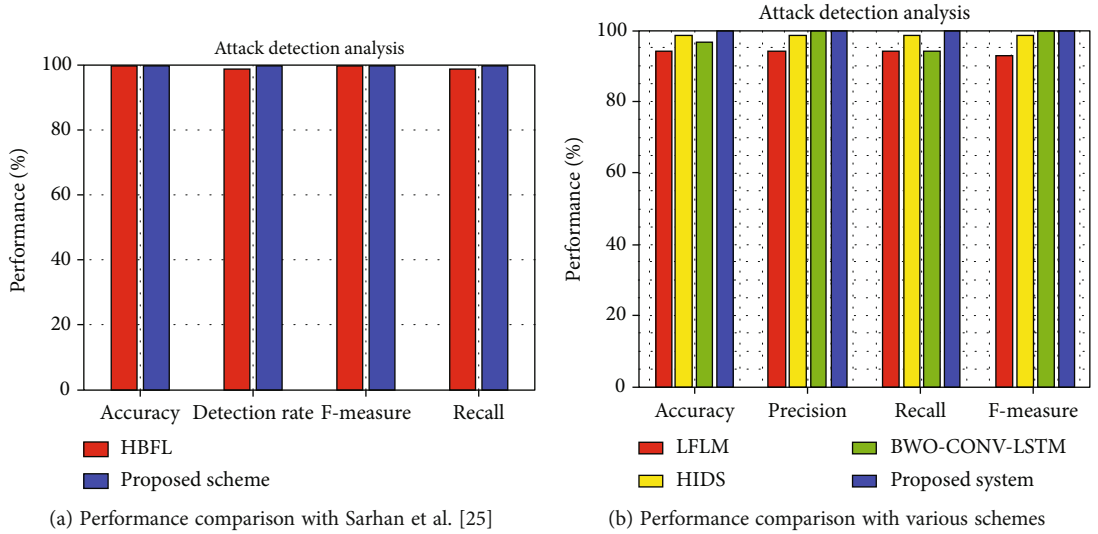
The performance of the proposed scheme is also evaluated using a UNSW_NB15 benchmark dataset with the application of various machine learning models, as shown in Figure 12(b). The binary classification technique using hybrid cluster labeling K-means achieves better classification accuracy of 100% using the benchmark dataset. As shown in Figure 12, the proposed system's performance is adequate for detection and localization attacks in IoT-based wireless sensor networks using benchmark datasets. This can be validated and confirmed by comparing the simulation and experimental results with other research works. Abubaker et al. [4] proposed blockchain-based malicious node detection and localization using federated machine learning using SVM and RF in IoT-based WSNs, as depicted in Figures 13(a) and 13(b). Figure 11 shows that the proposed approach effectively detects and localizes malicious nodes in IoT-WSNs using a hybrid federated machine learning approach. Using the dynamic power of SM 3.0 networks and MNSIT datasets as a benchmark evaluation method, Salim et al. [63] present a differentially privacy blockchain-based explainable FL (DP-BFL) architecture as shown in Figure 13(b). Any Internet-connected device can now contribute data to a globally secure model thanks to this platform. Rehman et al. [64] explored the interplay between blockchain technology and federated learning in healthcare and provides a thorough analysis of the results 5.0. The goal of this research is to build a safe healthcare monitoring system 5.0 by using a blockchain-enabled, federated-learning-based intrusion detection system (FL-IDS) to detect any malicious activity within a healthcare network and allowing doctors to keep tabs on patients via sensors and take preventative measures on a regular basis by foreseeing diseases, as shown in Figure 13(a). The results of the suggested system show that the method is well suited to healthcare monitoring.

The results show that the proposed scheme is better compared to Awan et al. [12] which presented a trust evaluation process based on blockchain for storing the identities of the sensor nodes (SN) and aggregator nodes (AN). The scheme utilized the private and public blockchains for detecting malicious nodes with the detection accuracy of 75% taking malicious nodes 20-80, whereas the proposed technique achieves a detection accuracy of 95%, bringing the number of malicious nodes to 5-30. Otoum et al. [14] proposed a framework using blockchain and federated learning for trust evaluation and security and achieved average detection rate accuracy of 96% and 93%, respectively. Kim et al. [1] achieved average localization detection accuracy and average localization error of 96.5% and 6.97 using secure blockchain based on trust management and evaluation, respectively. Friha et al. [17] designed DNN, CNN, and RNN learning techniques and achieved average detection accuracy of 98.63%, 99.71%, and 99.05% for FELIDS using a benchmark dataset for measuring the detection performance of the system. Kim et al. [41] presented collaborative security system-based federated learning for the internet of things. They achieved a training accuracy of 98.47% using the NSL-KDD benchmark dataset with various classes of attacks for training and testing. This suggests that the proposed scheme is enhanced for detecting and localizing attacks in IoT-based WSNs. The proposed blockchain-based technique for attack detection and localization enhances security. It reduces energy consumption and



(a) Hybrid SVM-based comparison (b) Hybrid RF-based comparison

Figure 13: Performance comparison of the proposed scheme with various evaluation metrics for identification of malicious nodes in IoT-WSNs using benchmark dataset.



(a) Performance comparison with Sarhan et al. [25] (b) Performance comparison with various schemes

Figure 14: Performance comparison of the proposed scheme using various evaluation metrics and benchmark dataset with previous works.

effective data aggregation using the trust values of the beacon nodes for improved network lifetime compared to Ahmed et al. [45]. Mahmood and Jusas [65] proposed blockchain-enabled federated learning for data security and privacy in IoT-WSNs and decentralized nodes and sent global data sharing and achieved average detection accuracy of 95%. The scheme also aggregates the data from the local model to the global model for secure transactions. This shows that the proposed method achieves better classification accuracy. Sarhan et al. [25] proposed a hierarchical blockchain-based federated learning (HBFL) for a secure collaborated intrusion detection system in IoT-based wireless sensor networks as shown in Figure 14(a). The scheme achieves an average detection accuracy of 99.71% using a benchmark dataset UNSW-NB15 containing various classes of attacks. Hassan et al. [66] proposed self-learning and the adaptive protocol to transmit multiuser data automatically by efficiently using channel spectral features and achieved average attack classification accuracy of 98.98 using a nonlinear SVM classification model for IoT-based WSNs. Nasser et al. [67] presented a lightweight federated learning model (LFLM) to learn medical symptoms with high accuracy from data collected by individuals utilizing ambient IoT based-sensors and wearable devices for the COVID-19 response in a collaborative yet private manner with evaluation metrics a shown in Figure 14(b). Roy et al. [68] present a two-layer hierarchical intrusion detection system (HIDS) based on machine learning to successfully detect intrusions

in IoT networks while meeting the IoT resource restrictions. Kanna and Santhi [69] proposed a MapReduce-based black widow optimized convolutional-long short-term memory (BWO-CONV-LSTM) networks for building a hybrid IDS model that efficiently addresses these problems in IoT-based WSNs.

Kumar et al. [56] proposed a federated machine learning-based intrusion detection system for detecting and localizing malicious nodes. They achieved an average detection accuracy of 92.7% using KDD-99 in IoT-based WSNs. The proposed approach is further compared with Mohar et al. [70], and Kumar et al. [71] performed an average localization error of 0.43 with four beacon nodes using a hybrid scalable, secure collaborative localization technique for WSNs. Sajid et al. [72] presented models based on genetic algorithms are presented for identifying malicious nodes, specifically the GA-SVM and GA-DT. After a rogue node has been identified, the Dijkstra algorithm is used to determine the most efficient route through the network. Both GA-DT and GA-SVM are highly accurate, with a 96% and 98% success rate, respectively. The GA-DT and GA-SVM each score 94% and 96%, respectively, when it comes to their accuracy. Nagalakshmi et al. [73] designed and presented detection and localization of black hole attacks based on machine learning models using 114 samples and achieved localization accuracy of 98.01% for WSNs. This demonstrates that the suggested approach performance is more sufficient for identifying and categorising attacks in IoT-based WSNs.

7. Conclusion

The proposed solution utilizes a blockchain-based security technique to detect and localize malicious nodes in hierarchically distributed IoT-WSNs. Using XGBoost and CLK-means machine learning federated classifiers for multiclass and binary classification approaches, the blockchain technique successfully identifies and pinpoints rogue nodes in IoT-based wireless sensor networks. Service delivery without the risk of attack and improved network speed was other goals of this technique, which also made use of feature evaluation and cascading encryption techniques. Results from simulation and classification demonstrate that the proposed system is sufficient for malicious node detection and localization, with average detection accuracies of 99.95% and 100% for XGBoost and CLK-means, respectively, based on the multiclass and binary classification approach with the CICIDS2017 benchmark dataset. Detecting and localizing attacks in IoT-WSNs with the use of a hybrid federated machine learning approach are a novel approach. The outcomes demonstrate that random forest is superior to other methods for classifying diverse network traffic data and detecting malicious nodes. Node honesty, end-to-end delay, and transaction latency are also measured to gauge the effectiveness of the proposed paradigm in delivering services safely and promptly. Since sensor nodes are typically placed in unattended environments where they are susceptible to various routing attacks, more research into how to identify the location of malicious nodes is warranted.

In the future, we want to investigate and develop cutting-edge blockchain-based secure IoT-WSNs with hybrid federated and hybrid machine learning methodologies for massive, safe, and intelligent IoT-WSN deployments in the near future. To locate and identify malicious nodes in IoT-WSNs, we will investigate blockchain-based sophisticated hybrid access control mechanisms. In addition, we will implement cutting-edge multiclass and binary classification evaluation metrics utilizing a wide range of benchmark datasets and hybrid routing protocols.

Data Availability

Data used in this article are available from the corresponding author upon request.

Conflicts of Interest

Authors confirm that they have no financial or personal ties to any parties that might influence the results of this study.

Authors' Contributions

Compiling, writing, and conceptualizing are done by Gebrekiros. Prof. J. Panda, and Prof. S. Indu did the discussion, validation, and evaluation.

References

[1] T. H. Kim, R. Goyat, M. K. Rai et al., "A novel trust evaluation process for secure localization using a decentralized blockchain in wireless sensor networks," *IEEE Access*, vol. 7, pp. 184133–184144, 2019.

[2] A. Sagu, N. S. Gill, and P. Gulia, "Hybrid deep neural network model for detection of security attacks in IoT enabled environment," *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 1, pp. 120–127, 2022.

[3] R. Bharathi, S. Kannadhasan, B. Padminidevi, M. S. Maharajan, R. Nagarajan, and M. M. Tonmoy, "Predictive model techniques with energy efficiency for iot-based data transmission in wireless sensor networks," *Journal of Sensors*, vol. 2022, Article ID 3434646, 18 pages, 2022.

[4] Z. Abubaker, N. Javaid, A. Almogren, M. Akbar, M. Zuair, and J. Ben-Othman, "Blockchained service provisioning and malicious node detection via federated learning in scalable Internet of Sensor Things networks," *Computer Networks*, vol. 204, article 108691, 2022.

[5] C. Ming, C. Xiaoting, D. Wensheng, and G. Jiahui, "A secure blockchain-based group key agreement protocol for IoT," *The Journal of Supercomputing*, vol. 77, no. 8, pp. 9046–9068, 2021.

[6] S. Agrawal, S. Sarkar, O. Aouedi et al., "Federated learning for intrusion detection system: concepts, challenges and future directions," *Computer Communications*, vol. 195, pp. 346–361, 2022.

[7] R. Kumar, P. Kumar, R. Tripathi, G. P. Gupta, S. Garg, and M. M. Hassan, "A distributed intrusion detection system to detect DDoS attacks in blockchain- enabled IoT network," *Journal of Parallel and Distributed Computing*, vol. 164, pp. 55–68, 2022.

[8] O. Cheikhrouhou and A. Koubaa, "BlockLoc: secure localization in the Internet of Things using blockchain," in *2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC)*, pp. 629–634, Tangier, Morocco, 2019.

[9] S. Algarni, F. Eassa, K. Almarhabi et al., "Blockchain-based secured access control in an IoT system," *Applied Sciences*, vol. 11, no. 4, article 1772, 2021.

[10] M. P. Nath, S. N. Mohanty, and S. B. B. Priyadarshini, "Application of machine learning in wireless sensor network," in *2021 8th International Conference on Computing for Sustainable Global Development (INDIACom)*, pp. 7–12, New Delhi, India, 2021.

[11] R. Goyat, G. Kumar, M. K. Rai, R. Saha, R. Thomas, and T. H. Kim, "Blockchain powered secure range-free localization in wireless sensor networks," *Arabian Journal for Science and Engineering*, vol. 45, no. 8, pp. 6139–6155, 2020.

[12] S. Awan, N. Javaid, S. Ullah, A. U. Khan, A. M. Qamar, and J. G. Choi, "Blockchain based secure routing and trust management in wireless sensor networks," *Sensors*, vol. 22, no. 2, pp. 1–24, 2022.

[13] H. Honar Pajooh, M. Rashid, F. Alam, and S. Demidenko, "Multi-layer blockchain-based security architecture for Internet of Things," *Sensors*, vol. 21, no. 3, pp. 1–26, 2021.

[14] S. Otoum, I. A. Ridhawi, and H. Mouftah, "Securing critical IoT infrastructures with blockchain-supported federated learning," *IEEE Internet of Things Journal*, vol. 9, no. 4, pp. 2592–2601, 2022.

[15] W. She, Q. Liu, Z. Tian, J. Sen Chen, B. Wang, and W. Liu, "Blockchain trust model for malicious node detection in wireless sensor networks," *IEEE Access*, vol. 7, pp. 38947–38956, 2019.

[16] K. Fan, S. Wang, Y. Ren et al., "Blockchain-based secure time protection scheme in IoT," *IEEE Internet of Things Journal*, vol. 6, no. 3, pp. 4671–4679, 2019.

[17] O. Friha, M. A. Ferrag, L. Shu, L. Maglaras, K. K. R. Choo, and M. Nafaa, "FELIDS: federated learning-based intrusion detection system for agricultural Internet of Things," *Journal of Parallel and Distributed Computing*, vol. 165, pp. 17–31, 2022.

[18] N. Javaid, "A secure and efficient trust model for wireless sensor IoTs using blockchain," *IEEE Access*, vol. 10, pp. 4568–4579, 2022.

[19] D. Wu and N. Ansari, "A trust-evaluation-enhanced blockchain-secured industrial iot system," *IEEE Internet of Things Journal*, vol. 8, no. 7, pp. 5510–5517, 2021.

[20] A. U. Khan, M. B. E. Sajid, A. Rauf, M. N. Saqib, F. Zaman, and N. Javaid, "Exploiting blockchain and RMCV-based malicious node detection in ETD-LEACH for wireless sensor networks," *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 7281872, 15 pages, 2022.

[21] R. Goyat, G. Kumar, M. Alazab, R. Saha, R. Thomas, and M. K. Rai, "A secure localization scheme based on trust assessment for WSNs using blockchain technology," *Future Generation Computer Systems*, vol. 125, pp. 221–231, 2021.

[22] D. Minoli and B. Occhiogrosso, "Blockchain mechanisms for IoT security," *Internet of Things*, vol. 1-2, pp. 1–13, 2018.

[23] Z. Ma, L. Wang, and W. Zhao, "Blockchain-driven trusted data sharing with privacy protection in IoT sensor network," *IEEE Sensors Journal*, vol. 21, no. 22, pp. 25472–25479, 2021.

[24] X. Yang, Y. Chen, X. Qian, T. Li, and X. Lv, "BCEAD: a blockchain-empowered ensemble anomaly detection for wireless sensor network via isolation forest," *Security and Communication Networks*, vol. 2021, Article ID 9430132, 10 pages, 2021.

[25] M. Sarhan, W. W. Lo, S. Layeghy, and M. Portmann, "HBFL: a hierarchical blockchain-based federated learning framework for a collaborative iot intrusion detection," 2022, https://arxiv.org/abs/2204.04254.

[26] S. J. Hsiao and W. T. Sung, "Employing blockchain technology to strengthen security of wireless sensor networks," *IEEE Access*, vol. 9, pp. 72326–72341, 2021.

[27] S. Abbas, N. Javaid, A. Almogren, S. M. Gulfam, A. Ahmed, and A. Radwan, "Securing genetic algorithm enabled SDN routing for blockchain based Internet of Things," *IEEE Access*, vol. 9, pp. 139739–139754, 2021.

[28] E. H. Abualsauod, "A hybrid blockchain method in Internet of Things for privacy and security in unmanned aerial vehicles network," *Computers and Electrical Engineering*, vol. 99, article 107847, 2022.

[29] G. G. Gebremariam, J. Panda, and S. Indu, "Localization and detection of multiple attacks in wireless sensor networks using artificial neural network," *Wireless Communications and Mobile Computing*, vol. 2023, Article ID 2744706, 29 pages, 2023.

[30] Z. Cui, F. Xue, S. Zhang et al., "A hybrid blockchain-based identity authentication scheme for multi-WSN," *IEEE Transactions on Services Computing*, vol. 13, no. 2, pp. 241–251, 2020.

[31] A. U. Khan, N. Javaid, M. A. Khan, and I. Ullah, "A blockchain scheme for authentication, data sharing and nonrepudiation to secure internet of wireless sensor things," *Cluster Computing*, vol. 112, pp. 1–16, 2022.

[32] Z. Ullah, "A survey on hybrid, energy efficient and distributed (HEED) based energy efficient clustering protocols for wireless sensor networks," *Wireless Personal Communications*, vol. 112, no. 4, pp. 2685–2713, 2020.

[33] S. Messous and H. Liouane, "Online sequential DV-hop localization algorithm for wireless sensor networks," *Mobile Information Systems*, vol. 2020, Article ID 8195309, 14 pages, 2020.

[34] S. Dong, X. G. Zhang, and W. G. Zhou, "A security localization algorithm based on DV-hop against Sybil attack in wireless sensor networks," *Journal of Electrical Engineering & Technology*, vol. 15, no. 2, pp. 919–926, 2020.

[35] A. Hadir, K. Zine-Dine, M. Bakhouya, and J. El Kafi, "An improved DV-hop localization algorithm for wireless sensor networks," in *2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, pp. 330–334, Wuhan, China, 2014.

[36] H. Chen, W. Lou, Z. Wang, J. Wu, Z. Wang, and A. Xi, "Securing DV-hop localization against wormhole attacks in wireless sensor networks," *Pervasive and Mobile Computing*, vol. 16, pp. 22–35, 2015.

[37] F. Khelifi, A. Bradai, A. Benslimane, P. Rawat, and M. Atri, "A survey of localization systems in Internet of Things," *Mobile Networks and Applications*, vol. 24, no. 3, pp. 761–785, 2019.

[38] R. Jeet, S. S. Kang, S. S. Hoque, and B. N. Dugbakie, "Secure model for IoT healthcare system under encrypted blockchain framework," *Security and Communication Networks*, vol. 2022, Article ID 3940849, 11 pages, 2022.

[39] S. M. Kasongo, "An advanced intrusion detection system for IIoT based on GA and tree based algorithms," *IEEE Access*, vol. 9, pp. 113199–113212, 2021.

[40] S. Subramani, *Deep learning based IDS for secured routing in wireless sensor networks using fuzzy genetic approach*, Research Square, 2022.

[41] S. Kim, H. Cai, C. Hua, P. Gu, W. Xu, and J. Park, "Collaborative anomaly detection for Internet of Things based on federated learning," in *2020 IEEE/CIC International Conference on Communications in China (ICCC)*, pp. 623–628, Chongqing, China, 2020.

[42] Y. Kayode Saheed, A. Idris Abiodun, S. Misra, M. Kristiansen Holone, and R. Colomo-Palacios, "A machine learning-based intrusion detection for detecting Internet of Things network attacks," *Alexandria Engineering Journal*, vol. 61, no. 12, pp. 9395–9409, 2022.

[43] L. Yang, A. Moubayed, and A. Shami, "MTH-IDS: a multitiered hybrid intrusion detection system for Internet of Vehicles," *IEEE Internet of Things Journal*, vol. 9, no. 1, pp. 616–632, 2022.

[44] B. Ghimire and D. B. Rawat, "Recent advances on federated learning for cybersecurity and cybersecurity for federated learning for Internet of Things," *IEEE Internet of Things Journal*, vol. 9, no. 11, pp. 8229–8249, 2022.

[45] A. Ahmed, S. Abdullah, M. Bukhsh, I. Ahmad, and Z. Mushtaq, "An energy-efficient data aggregation mechanism for IoT secured by blockchain," *IEEE Access*, vol. 10, pp. 11404–11419, 2022.

[46] P. Roy and C. Chowdhury, "A survey of machine learning techniques for indoor localization and navigation systems," *Journal of Intelligent & Robotic Systems*, vol. 101, no. 3, pp. 1–34, 2021.

[47] G. Kocher and G. Kumar, "Machine learning and deep learning methods for intrusion detection systems: recent developments and challenges," *Soft Computing*, vol. 25, no. 15, pp. 9731–9763, 2021.

[48] S. Pande, A. Khamparia, and D. Gupta, "Feature selection and comparison of classification algorithms for wireless sensor networks," *Journal of Ambient Intelligence and Humanized Computing*, vol. 2021, pp. 1–13, 2021.

[49] D. Upadhyay, J. Manero, M. Zaman, and S. Sampalli, "Gradient boosting feature selection with machine learning classifiers for intrusion detection on power grids," *IEEE Transactions on Network and Service Management*, vol. 18, no. 1, pp. 1104–1116, 2021.

[50] S. Taleb, A. Al Sallab, H. Hajj, Z. Dawy, R. Khanna, and A. Keshavamurthy, "Deep learning with ensemble classification method for sensor sampling decisions," in *2016 International Wireless Communications and Mobile Computing Conference (IWCMC)*, pp. 114–119, Paphos, Cyprus, 2016.

[51] M. Sri Vidya and G. R. Sakthidharan, "Accurate anomaly detection using various machine learning methods for IoT devices in indoor environment," in *2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, pp. 308–316, Palladam, India, 2021.

[52] L. Moulad, H. Belhadaoui, and M. Rifi, "Implementation of a hierarchical hybrid intrusion detection mechanism in wireless sensors network," *International Journal of Advanced Computer Science and Applications*, vol. 8, no. 10, pp. 270–278, 2017.

[53] M. Feurer and F. Hutter, *Hyperparameter optimization*, Springer, 2019.

[54] M. Farooq-I-Azam, Q. Ni, and E. A. Ansari, "Intelligent energy efficient localization using variable range beacons in industrial wireless sensor networks," *IEEE Transactions on Industrial Informatics*, vol. 12, no. 6, pp. 2206–2216, 2016.

[55] S. Karagol and D. Yildiz, "A novel path planning model based on nested regular hexagons for mobile anchor-assisted localization in wireless sensor networks," *Arabian Journal for Science and Engineering*, vol. 47, no. 8, pp. 9833–9848, 2022.

[56] K. P. S. Kumar, S. A. H. Nair, D. Guha Roy, B. Rajalingam, and R. S. Kumar, "Security and privacy-aware artificial intrusion detection system using federated machine learning," *Computers and Electrical Engineering*, vol. 96, article 107440, 2021.

[57] Z. Tan, A. Jamdagni, P. Nanda, X. He, and R. P. Liu, "Evaluation on multivariate correlation analysis based denial-of-service attack detection system," in *SecurIT '12: Proceedings of the First International Conference on Security of Internet of Things*, pp. 160–164, Kollam India, 2012.

[58] Z. Tan, "Detection of Denial-of-Service Attacks Based on By," pp. 1–14, University of Technology Sydney Australia, 2013.

[59] M. Nivaashini and P. Thangaraj, "Computational intelligence techniques for automatic detection of Wi-Fi attacks in wireless IoT networks," *Wireless Networks*, vol. 27, no. 4, pp. 2761–2784, 2021.

[60] C. Lyu, X. Zhang, Z. Liu, and C. H. Chi, "Selective authentication based geographic opportunistic routing in wireless sensor networks for Internet of Things against DoS attacks," *IEEE Access*, vol. 7, pp. 31068–31082, 2019.

[61] B. Kaur and D. Prashar, "Localization in wireless sensor network: techniques, algorithms analysis and challenges," in *2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, pp. 1–7, Noida, India, 2021.

[62] T. A. Alghamdi, I. Ali, N. Javaid, and M. Shafiq, "Secure service provisioning scheme for lightweight IoT devices with a fair payment system and an incentive mechanism based on blockchain," *IEEE Access*, vol. 8, pp. 1048–1061, 2020.

[63] S. Salim, B. Turnbull, and S. Member, "A Blockchain-Enabled Explainable Federated Learning for Securing Internet-of-Things-Based Social Media 3. 0 Networks," *IEEE Transactions on Computational Social Systems*, pp. 1–17, 2021.

[64] A. Rehman, S. Abbas, M. A. Khan, T. M. Ghazal, K. Muhammad, and A. Mosavi, "A secure healthcare 5.0 system based on blockchain technology entangled with federated learning technique," *Computers in Biology and Medicine*, vol. 150, article 106019, 2022.

[65] Z. Mahmood and V. Jusas, "Blockchain-enabled: multi-layered security federated learning platform for preserving data privacy," *Electronics*, vol. 11, no. 10, p. 1624, 2022.

[66] T. Hassan, S. Aslam, and J. W. Jang, "Fully automated multi-resolution channels and multithreaded spectrum allocation protocol for IoT based sensor nets," *IEEE Access*, vol. 6, pp. 22545–22556, 2018.

[67] N. Nasser, Z. M. Fadlullah, M. M. Fouda, A. Ali, and M. Imran, "A lightweight federated learning based privacy preserving B5G pandemic response network using unmanned aerial vehicles: a proof-of-concept," *Computer Networks*, vol. 205, article 108672, 2022.

[68] S. Roy, J. Li, and Y. Bai, "A two-layer fog-cloud intrusion detection model for IoT networks," *Internet of Things*, vol. 19, article 100557, 2022.

[69] P. R. Kanna and P. Santhi, "Hybrid intrusion detection using MapReduce based black widow optimized convolutional long short-term memory neural networks," *Expert Systems with Applications*, vol. 194, article 116545, 2022.

[70] S. S. Mohar, S. Goyal, and R. Kaur, "Localization of sensor nodes in wireless sensor networks using bat optimization algorithm with enhanced exploration and exploitation characteristics," *The Journal of Supercomputing*, vol. 78, no. 9, pp. 11975–12023, 2022.

[71] A. Kumar, "A hybrid fuzzy system based cooperative scalable and secured localization scheme for wireless sensor networks," *International Journal of Wireless & Mobile Networks*, vol. 10, no. 3, pp. 51–68, 2018.

[72] M. B. E. Sajid, S. Ullah, N. Javaid, I. Ullah, A. M. Qamar, and F. Zaman, "Exploiting machine learning to detect malicious nodes in intelligent sensor- based systems using blockchain," *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 7386049, 16 pages, 2022.

[73] T. J. Nagalakshmi, A. K. Gnanasekar, G. Ramkumar, and A. Sabarivani, "Machine learning models to detect the black-hole attack in wireless adhoc network," *Materials Today: Proceedings*, vol. 47, pp. 235–239, 2021.

Check for updates

Research Paper

Can Gratitude Counter Workplace Toxicity? Exploring the Mediating Role of Psychological Capital (PsyCap)

Business Perspectives and Research
1–16
© 2023 K.J. Somaiya Institute of Management Studies and Research
Reprints and permissions:
in.sagepub.com/journals-permissions-india
DOI: 10.1177/22785337221148300
journals.sagepub.com/home/bpr
SAGE

Ayatakshee Sarkar[1], Naval Garg[2], DK Srivastava[3] and BK Punia[4]

Abstract

Gratitude is essential in positive organizational scholarship and is said to flourish and prosper organization members. This article aims to investigate the role of gratitude in transforming a toxic workplace into a collaborative work environment by exploring the mediating role of psychological capital (Psycap). The results reveal that gratitude practiced by the employees at the workplace can counter workplace toxicity and enhance a collaborative work environment, creating harmony in society. The study proposes the institutionalization of a gratitude-based organizational culture that prevents workplace toxicity and supports a collaborative work environment. The study is one of the pioneer studies that tried to unlock the "black box" through which dispositional gratitude can help build workplace harmony and collaboration.

Keywords

Gratitude, psychological capital, workplace toxicity, mediation

Introduction

A workplace where employees spend the maximum day duration can either make or break their sanity. In short, the workplace ensures the workforce's physical, social, and psychological well-being (Carmichael et al., 2016). A counter-productive work environment may have many titles but is popularly known as a toxic work environment or workplace toxicity (Anjum & Ming, 2018; Gupta et al., 2021; Jha & Sud, 2020). "Toxic," as the name suggests, has a lethal connotation, and anything toxic is detrimental to health, the psyche, and the surrounding environment (Lave & Upton, 1987). Although the issue of

[1] XLRI, Jamshedpur, Jharkhand, India

[2] Delhi Technological University, New Delhi, Delhi, India

[3] Tata Institute of Social Sciences, Mumbai, Maharashtra, India

[4] Guru Jambheshwar University of Science & Technology, Hisar, Haryana, India

Corresponding author:
Ayatakshee Sarkar, XLRI, Jamshedpur, Jharkhand 831001, India.
E-mails: ayatakshee@xlri.ac.in; ayatakshee@gmail.com

workplace toxicity has been studied sufficiently by previous researchers, the focus has been on the negative repercussions of toxicity at the workplace. The dysfunctional consequences of a toxic work environment include low morale (Tai et al., 2012), demonstrated incivility, workplace harassment (Clear, 2017), bullying in both explicit and implicit means (Tastan, 2017), high absenteeism and turnover (Too & Harvey, 2012), moral disengagement (Duffy, 2009), low compassion competence (Woo & Kim, 2020), high psychological distress (Abubakar, 2018), employee burnout and emotional exhaustion (Megeirhi et al., 2020), low job productivity (Anjum & Ming, 2018), sabotaging (Abubakar et al., 2018) and vandalism, etc. Given such dire consequences, more studies on prevention become critical.

The present study proposes gratitude as an individual trait that boosts one's tolerance for toxicity. In Buddhism, there is a saying that the virtue of gratitude can sometimes turn poison into medicine. It means that gratitude has tremendous power to change any negative situation or state to a positive one. Even if the change is not absolute, it can help make it bearable for both parties. It also improves the workplace climate, contributing to employee welfare and reducing negativity, mistrust, and hatred. Gratitude helps recognize and acknowledge good in self, others, and the larger environment (Emmons, 2003).

The proposed link between gratitude, Psycap, and workplace toxicity is theoretically premised in the "Broaden and Build theory of positive emotions" and "Amplification theory of gratitude." The theory states that positive emotions like gratitude open up one's range of contemplation, thoughts, introspection, and action tendencies (Fredrickson, 2004), which leads to heightened cognition, creativity, innovation, and flexible actions. This broadened awareness builds up resources to create mindful employees who help to foster harmonious relations for collaborative work culture (Goldberg et al., 2020; Trunnell & Braza, 1995). Based on this theory, the current study proposes that gratitude will build positive Psycap resources, which will help combat workplace toxicity and foster a collaborative work environment. This article focus on three research questions:

1. What is the association between gratitude and the prevention of workplace toxicity?
2. What is the relationship of gratitude in enhancing the Psycap resources of the employees?
3. How do gratitude and the Psycap combat workplace toxicity?

The article briefly describes previous studies on gratitude, Psycap, and toxic work environment, especially workplace bullying and incivility. A model is proposed based on adequate theoretical premises that depict relationships between gratitude, Psycap, and a toxic work environment. Next, the research methodological framework of the study is elaborated. Following the methodology, results and findings are discussed. Finally, the theoretical and practical implications and the study's limitations are explained.

Literature Review and Hypothesis Formulation

Gratitude

McCullough et al. (2002) defined gratitude as "a generalized tendency to recognize and respond with grateful emotion to the roles of other people's benevolence in the positive experiences and outcomes that one obtains" (p. 11). It is conceptualized as disposition or trait, state-such as behavior, affective mood, emotions, religious practice, spiritual intervention, and ethical practice (Watkins, 2014). While understanding the meaning and definition of gratitude, one must understand its nuance, demonstrating the contrasts between gratitude and indebtedness. Indebtedness has an expectation factor that leads to negative emotions of fulfillment, dissatisfaction, and unhappiness; hence it is out of the scope of the study.

Watkins (2014) attempted to comprehend the complex concept of gratitude by adopting Rosenberg's (1998) "level of analysis," in which gratitude is perceived as an affective phenomenon. It could be seen at three levels: emotion, mood, and trait. As an emotion, it manifests either as a temporal emotional feeling or brief "psychological changes". As a practical mood, it is transient; still, it lasts a little longer than emotions, or it may be a personality trait ingrained in human personality and remains stable and unaffected under external circumstances (Watkins, 2014). The present study explores gratitude as an affective trait because traits are relatively more durable, representing the state of being (Jans-Beken et al., 2019). Several authors, such as Watkins (2014), Peterson and Seligman (2004), and Garg and Gera (2019) also investigated the trait facet of gratitude. As an affective trait, it is the general tendency to recognize and respond with grateful emotions to the roles of other people's benevolence in the positive experiences and outcomes that one obtains (McCullough et al., 2002, p. 112). The following section will learn about the toxic work environments (TWE).

TWE

TWE encompasses counter-productive facets, such as workplace bullying, incivility, ostracism, and sexual harassment (Anjum & Ming, 2018). However, in the current study, only two aspects, that is, incivility and bullying, have been taken into consideration because of their prevalence in contemporary socio-cultural contexts. Incivility may be verbal or non-verbal action tendencies that reflect disrespect towards colleagues. Quite often, negative interpersonal behavior with low intensity gets unnoticed, and later on, it may result in the upward spiraling of severe aggression and hostility. Unlike physical attacks, incivility is more implicit and less visible but has a similar intent. Workplace bullying, however, is intended to create a risky environment that impedes growth and performance. Besides, it is also detrimental to employees' self-esteem and performance (Anjum & Ming, 2018; Sarkar & Garg, 2020). An interesting observation is that intimidation is usually done in a lonely space when no other colleagues are around so that an employee, if not self-contained, can react and hampers individual sanity. There have been numerous research on workplace bullying and its detrimental effects. Besides, in bullying behaviors, there is a slight connotation of power. Often, the perpetrator holds a higher position than the victim; therefore, in many cases, it goes unreported because of the fear of losing one's job. Consequently, it spirals up and contributes to workplace toxicity. The following section will learn about Psycap.

Psycap

Psycap, being a composite construct, has its root in the discipline of Positive Scholarship Literature or Positive Organization Behavior (Toor & Ofori, 2010). The concept of PsyCap was initially introduced in Martin Seligman's book "Authentic Happiness" in which he mentioned the "flow-state" at the workplace. According to the author, the flow state is a mental state of intense attention, which builds up PsyCap (Luthans & Youssef-Morgan, 2017). Psycap is predominantly a higher-order "positive state" construct (Luthans et al., 2005). It is defined as an individual's positive psychological state of development and is characterized by (a) having confidence (self-efficacy) to take on and put in the necessary effort to succeed at challenging tasks; (b) making a positive attribution (optimism) about succeeding now and in the future; (c) persevering towards goals and, when necessary, redirecting paths to goals (hope) to succeed; and (d) when beset by problems and adversity, sustaining and bouncing back and even beyond (resilience) to attain success' (Luthans et al., 2007, p. 3). It is often characterized by four dimensions: hope, optimism,

self-efficacy, and resilience (Luthans et al., 2007). Although these four constructs are conceptually distinct, they comprise a synergetic set of mental resources (Hobfoll, 2002). Also, Psycap is studied at both individual and organizational levels. At the individual level, it powers up personal growth and development. It provides resources that trigger productivity, efficiency, and sustainability advantage through improved performance (Luthans et al., 2004). It focuses primarily on strengthening unique resources and strengths (Sarkar & Garg, 2020; Megeirhi et al., 2018). In this study, it is proposed that gratitude builds personal-psychological resources amongst employees, which helps combat toxic work environments and establish a more collaborative work culture.

Gratitude, PsyCap, and TWE

The immediate linkage is theoretically based on the "Moral effect theory of gratitude" (Watkins, 2014) and the "Amplification theory of gratitude" (McCullough et al., 2002); besides, the indirect association is premised on the "Broaden-and-build theory of positive emotion" (Fredrickson, 2004). The moral effect theory explains three moral functions of gratitude, that is, barometer, motivator, and reinforcer. As a barometer, gratitude helps one to realize and appreciate others' contributions to one's life. According to McCullough et al. (2002), an individual experiences gratitude when (a) they receive valued assistance, (b) the benefactor incurs high costs and efforts, (c) unintentional efforts of the donor, and (d) the act is gratuitous, that is, not defined by any role-based existence. As a moral motivator, it motivates one to behave prosocially. The ethical motivation for gratitude is different from that of indebtedness. Often, indebtedness leads to unpleasant and adverse psychological states (Buunk et al., 1993), but gratitude elicits a positive feeling of grace. Finally, as a moral reinforcer, it stimulates prosocial behavior and encourages benefactors to do something good in the future. "An expression of gratefulness in the form of 'thank you' reinforces positivity" (McCullough et al., 2002). Thus, these postulates of moral affect theory support the mutual exchange of goodness among employees. Also, a morally motivated employee is less prone to exhibit toxicity in the workplace. Their acts may break the action-reaction cycle of negative behaviors. Based on these theoretical premises, the following hypotheses were proposed.

H1: Gratitude is negatively associated with workplace bullying.

H2: Gratitude is negatively associated with workplace incivility.

The proposed relationships are also based on the "Broaden and Build theory of positive emotion." Unlike negative emotions, positive ones do not obtain specific action tendencies; rather, "they broaden people's momentary thought-action repertoires and build their enduring personal resources" (Fredrickson, 2004, p. 1369). According to this theory, a positive emotion such as gratitude broadens one's outlook by eliciting new ideas, experiences, and information. A grateful person tends to develop new thinking models and act, enlarging one's cognitive contexts. This broadening builds personal, psychological, and social resources for an individual (McCullough et al., 2002). Fredrickson elaborates on four ways in which gratitude makes enduring resources. First, gratitude-based reciprocal altruism develops and strengthens social bonds and friendships among employees. It builds mutual trust, attachment, and permanent alliance (Emmons & Shelton, 2002). It also promotes groups, civil communities, and organizations. Based on these theoretical premises, we propose that gratitude builds up Psycap resources

(hope, optimism, self-efficacy, resilience) among employees. It restricts violent tendencies and violations of norms, thus facilitating better interpersonal relationships and a collaborative work environment.

Watkin's (2014) amplification theory of gratitude also states that the positive emotions of gratitude augment the employees' psychological resources. It makes an employee more creative, joyous, and harmonious. These resources are a kind of survival strategy that helps live a flourishing life in the future (Watkins, 2014). Based on this theory, it can be inferred that gratitude is an essential component of well-being as it amplifies "the good" in one's life. Hence this study proposes that gratitude strengthens an employee's hope, optimism, resilience, and self-efficacy, which helps combat workplace toxicity. The following model and hypotheses are proposed (Figure 1).

H3: Hope mediates the relationship between gratitude and workplace bullying.

H4: Optimism mediates the relationship between gratitude workplace bullying.

H5: Self-efficacy mediates the relationship between gratitude and workplace bullying.

H6: Resilience mediates the relationship between gratitude and workplace bullying.

H7: Hope mediates the relationship between gratitude and workplace incivility.

H8: Optimism mediates the relationship between gratitude workplace incivility.

H9: Self-efficacy mediates the relationship between gratitude and workplace incivility.

H10: Resilience mediates the relationship between gratitude and workplace incivility.

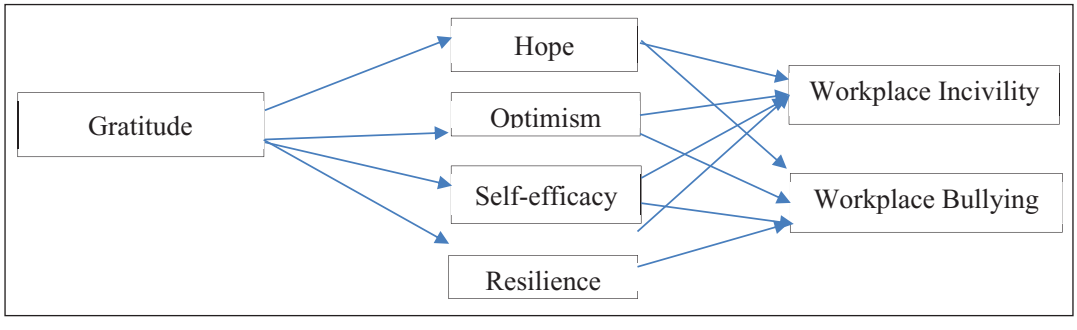


Figure 1. Proposed Mediation Model.

Source: The authors.

Research Methodology

Participants

The study explores the association between gratitude and a toxic work environment (workplace bullying and incivility). It investigates direct and indirect associations between gratitude and workplace bullying and incivility. Indirect linkage is examined by analyzing the mediating role of PsyCap in the association between gratitude and a toxic work environment. The sample comprises 150 employees working in different industries, such as banks (27), insurance (31), pharmaceutical (32), oil & gas (32), and footwear (28). These participants are engaged in MBA- Executive programs in various educational institutes in Delhi-National Capital Region (NCR), India. The study was restricted to the Delhi-NCR region for the convenience of data collection. Authors could easily procure a list of all students enrolled in the MBA-Executive program in this region. Questionnaires were sent to 200 students who were chosen randomly, of which 155 reverted. Five questionnaires were discarded for being incomplete. The respondents were contacted through e-mails, and data was collected with the help of Google Forms. First, data related to gratitude and PsyCap were collected. Information about incivility and bullying was collected after three months to ensure the independent variable's temporal precedence. The demographic details of the respondents are mentioned in Table 1.

Measures

Gratitude

Three-dimensional GRAT-16 scale measures gratitude. It accesses a lack of sense of deprivation (six statements), simple appreciation (six statements), and appreciation for others (four statements). It is a 9-point rating scale with anchors from 1 (strongly disagree) to 9 (strongly agree). A few items included in the above questionnaire were "I could not have gotten where I am today without the help of many people," life has been good to me.

Table 1. Demographic Details.

Variable	Category	Frequency	%
Gender	Male	90	60
	Female	60	40
Age	Up to 30 years	58	37
	31–40 years	50	32
	41–50 years	21	16
	Above 50 years	21	15
Experience	Below 5 years	42	28
	6–10 years	60	40
	11–15 years	18	12
	Above 15 years	30	20

Source: Primary data.

Psychological Capital

PsyCap is captured using Luthans et al.'s (2007) 12-item scale that measures self-efficacy (three statements), hope (four statements), resilience (three statements), and optimism (two statements). It is a 6-point rating scale from strongly disagree (1) to agree (6) strongly. Items included in the questionnaire are—"I feel confident contributing to discussions about the company's strategy," "I can get through difficult times at work because I have experienced difficulty before," and "I am optimistic about what will happen to me in the future as it pertains to work."

Toxic Work Environment

The present study examines two dimensions of the toxic work environment, that is, incivility and workplace bullying. Cortina and Magley (2009) workplace incivility scale (WIS) measures workplace incivility. It is a 7-item scale that captures incivility on a 5-point rating scale ranging from 0 (never) to 4 (most of the time). A few items included in the scale are—"Put you down or was condescending to you," and "Addressed you in unprofessional terms, either publicly or privately." Einarsen et al.'s (1994) Negative Acts Questionnaire-Revised (NAQ-R) measures workplace bullying. It is a 22-item scale that measures bullying on a 5-point rating scale. Examples of items included in the scale are—"Someone withholding information, which affects your performance" and "Spread of gossip or rumour about you."

Data Analysis and Results

Reliability, Validity, Multicollinearity, and Common Method Bias

Reliability is ensured using Cronbach's Alpha (CA) and Composite Reliability (CR) estimates. Hair et al. (1998) recommended a minimum value of 0.70 for CA and CR to confirm the reliability. Hair et al. (2006) also suggested that a threshold value of 0.50 of average variance explained (AVE) establishes convergent validity. AVE values for various parameters are gratitude (0.679), hope (0.553), optimism (0.574), self-efficacy (0.583), resilience (0.598), bullying (0638), and incivility (0.690). All these values are well-above 0.50, which confirms convergent validity. The discriminant validity is accessed with the help of the Fornell–Larcker test. The test demands that AVE values be greater than the squared correlation involving the dimension to confirm discriminant validity. Table 2 reports the result of the Fornell–Larcker test, which confirms the discriminant validity of the PsyCap scale. Multicollinearity is annulled using variance inflation factor (VIF) values. A VIF value <5 is required to suggest the absence of multicollinearity (Kline, 2008). Table 1 reports all study variables' CA values, CR, AVE, and VIF. All these values confirm the threshold values discussed above; hence, validity, reliability, and the absence of multicollinearity are established.

The common method bias could severely impact the study results; thus, procedural and quantitative methods should be employed to deal with it (Podsakoff et al., 2012). Two procedural remedies and one quantitative technique are used in this study. As per the suggestions of Podsakoff et al. (2012), questionnaires with different scales are used (Gratitude—9-point scale, PsyCap—6-point scale, and incivility/workplace bullying—5-point scale. The respondents are promised privacy and strictly academic usage of the data. The Harman one-factor method is also used to assess common method variance. All items of the study are subjected to factor analysis. One single factor does not account for the majority of variation, which negates the presence of common method bias.

The data is investigated in three steps using SPSS 21. The normalcy of collected information is explored with the help of the Kolmogorov–Smirnov test. The significance value is above 0.05, meaning the test is insignificant. It confirms that the data is normal. The mean and standard deviation of gratitude, four dimensions of PsyCap, incivility, and workplace bullying are calculated. Correlation and regression are applied to insight into the interrelationships between the three variables of the study. Further, PROCESS Macro in SPSS is used to calculate bootstrapping estimates that signify mediation.

Results

Table 1 reports mean, standard deviation, reliability, validity, and multicollinearity estimates. The average gratitude score is 5.20, with a standard deviation of 0.845. The most possessed dimension of PsyCap is resilience, with a mean value of 3.69. The mean values of hope, optimism, and self-efficacy are 3.42, 3.49, and 2.89, respectively. The average workplace bullying and incivility values are 1.69 and 1.73, respectively.

Table 2 highlights the correlation matrix of all the variables of this study. It is observed that gratitude is positively and statistically significantly correlated with all four dimensions of PsyCap. Besides, gratitude and the four constructs of PsyCap report a significant negative association with incivility and workplace bullying. These findings support the initial hypothesis of a positive association between gratitude and PsyCap. In other words, it supports the proposed assumption that gratitude builds psychological resources for an individual. Table 3 elucidates regression equations that consider gratitude and the four dimensions of PsyCap as independent variables and incivility and workplace bullying as dependent variables. The results reveal that gratitude significantly predicts incivility and workplace bullying. It explains 5% and 12% variations in incivility and workplace bullying, respectively. Here, a direct linkage between gratitude and a toxic work environment is established. The results also highlight that the four dimensions of PsyCap are significant forecaster of incivility and workplace bullying. PsyCap could also explain the 34% and 28% variations in workplace bullying and incivility. Again, these findings statistically support the proposed hypothesis of the negative linkage of PsyCap and a toxic work environment. The results provide support for the acceptance of H1 and H2.

Table 2. Descriptive Statistics.

Variables	Dimensions	Mean	S. D	CA	CR	AVE	VIF
Gratitude	Gratitude	5.20	0.845	0.844	0.872	0.679	2.637
PsyCap	Hope	3.42	0.683	0.812	0.840	0.553	1.624
	Optimism	3.49	0.760	0.820	0.848	0.574	1.421
	Self-efficacy	2.89	0.653	0.838	0.863	0.583	1.374
	Resilience	3.69	0.732	0.847	0.878	0.598	1.247
Toxic work environment	Bullying	1.69	0.746	0.778	0.792	0.638	-
	Incivility	1.73	0.719	0.765	0.781	0.69	-

Source: Primary data.

Note: CA, Cronbach's Alpha; CR, Composite Reliability; AVE, Average Variance Explained; VIF, Variance Inflation Index.

Table 3. Correlation Matrix.

Variables	1	2	3	4	5	6	7
Gratitude	1						
Hope	0.373**	1					
Optimism	0.221**	0.117	1				
Self-efficacy	0.300**	0.432**	0.243**	1			
Resilience	0.390**	0.269**	0.324**	0.214**	1		
Bullying	−0.350**	−0.347**	−0.348**	−0472**	−0.391**	1	
Incivility	−0.235**	−0.311**	−0.275**	−0.494**	−0.265**	0.488**	1

Source: Primary data.

Note: *Sig. at 0.005, ** Sig. at 0.01.

Table 4. Results of Multiple Regression.

DV	IV	Regression Equation	R^2	Sig.	Hypotheses Support
Bullying	G	*Bullying = 3.302–0.350 Gratitude*	0.12	Yes	H1
Incivility	G	*Incivility = 2.774–0.235 Gratitude*	0.05	Yes	H2
Bullying	Hope optimism self-efficacy resilience	*Bullying = 4.690–0.128 Hope—0.184 Optimism—0.319 Self-efficacy—0.219 Resilience*	0.34	Yes	-
Incivility	Hope optimism self-efficacy resilience	*Incivility = 4.157–0.098 Hope—0.0135 Optimism—0.395 Self-efficacy—0.100 Resilience*	0.28	Yes	-

Source: Primary data.

Note: G = Gratitude, *$p < 0.05$ and ** $p < 0.001$.

The mediating effect of psychological capital is analyzed with the help of Baron and Kenny's (1986) method. However, it is argued that Baron and Kenny's (1986) method provides a fair idea about mediation, but the statistical significance of the mediator could be examined using a Sobel test (Yay, 2017). Also, both these methods assume a normal distribution of data, thus, researchers advocate the bootstrapping process, which is resampling with a replacement for large times such as 5,000 or 10,000 times, to confirm the mediation effect. Table 5 reports the direct and indirect of gratitude on bullying and incivility. In Table 6, all eight possible paths of the mediation model, 0 (Zero) lie between lower and upper levels of the confidence interval, which signifies that hope, optimism, self-efficacy, and resilience mediate the relationship between gratitude and incivility and workplace bullying. These findings support H3–H10. In a nutshell, the results establish a positive association between gratitude and PsyCap, a negative relationship between gratitude and a toxic work environment, and a mediating effect of PsyCap.

Table 5. Results of Mediation Analysis (Baron & Kenny, 1986)

	IV to DV	IV to M	M to DV	IV to DV (with M)
G → H → Incivility	−0.235*	0.412*	−0.237*	−0.174*
G → H → Bullying	−0.350*	0.435*	−0.314*	−0.241*
G → O → Incivility	−0.235*	0.389*	−0.286*	−0.147*
G → O → Bullying	−0.350*	0.589*	−0.374*	−0.228*
G → SE → Incivility	−0.235*	0.485*	−0.334*	−0.184*
G → SE → Bullying	−0.350*	0.511*	−0.475*	−0.223*
G → R → Incivility	−0.235*	0.417*	−0.288*	−0.144*
G → R → Bullying	−0.350*	0.333*	−0.249*	−0.197*

Source: Primary data.

Note: *$p < 0.05$ and ** $p < 0.001$, G = Gratitude; SE = Self Efficacy, H = Hope, O = Optimism, R = Resilience.

Table 6. Results of Mediation Analysis.

Hypothesized Relationships	Sobel Test / Bootstrapping Estimates						Hypotheses Support
	Z-Value	Sig.	Effect	SE	LLCI	ULCI	
G → H → Incivility	−2.606	0.008*	−0.8027	0.0365	−0.1555	−0.0097	H3
G → H → Bullying	−2.685	0.008*	−0.8023	0.0342	−0.1482	−0.0168	H7
G → O → Incivility	2.201	0.004*	−0.442	0.0230	−0.0991	−0.0081	H4
G → O → Bullying	−2.224	0.024*	−0.0556	0.0274	−0.1179	−0.0123	H8
G → SE → Incivility	−2.070	0.030*	−0.1189	0.0339	−0.1858	−0.0520	H5
G → SE → Bullying	−3.026	0.002*	−0.1069	0.0314	−0.1728	−0.0474	H9
G → R → Incivility	−3.325	0.001*	−0.0679	0.0448	−0.1623	−0.0118	H6
G → R Bullying	−3.134	0.001*	−0.1032	0.0512	−0.2146	−0.0145	H10

Source: Primary data.

Note: *$p < 0.05$ and **$p < 0.001$, G = Gratitude; SE = Self Efficacy; H = Hope; O = Optimism; R = Resilience.

Discussion

The present research was conducted to study the complexity of the relationship between gratitude, Psycap, and workplace toxicity. The results support the validity of all the hypotheses: perceived Psycap resources, gratitude, and their negative association with workplace toxicity. However, an interesting finding is the low *r* square value (0.05) predicting incivility. Plausible reasons that explain the counterintuitive finding may be that incivility behaviors are challenging to expect; they are subtle and implicit in nature (Porath & Pearson, 2010). Gratitude is also a rare practice among people. Rather than feeling grateful for the limited resources, people make hue and cry about the situation.

On the other hand, more explicit observable behaviors, such as feeling optimistic, hopeful, and having confidence can directly reduce negative behaviors. This is explained by the higher r square values (0.34, 0.28) predicting the negative relationship between incivility and bullying behaviors. The study's

findings corroborate past studies in which there are mechanisms linking gratitude to well-being, positive coping mechanisms, and controlling negative behaviors (Sun et al., 2019; Wood et al., 2007). Gratitude interventions significantly reduce negative affect. The specific mechanism that would explain the relationship could be positive coping strategies.

Wood et al. (2007) examined the relationship in which grateful people can positively respond and cope with adversaries that may arise from toxic psychological situations. Stress occurs when events are appraised as threatening or harmful (Lazarus & Folkman, 1984), leading to a deadly upward spiral. Grateful people positively appraise the situation and find opportunities in every challenge they encounter. In the process, they also expand their psychological resources to better deal with the negative situation. In eastern philosophy, *Kritajna* (gratitude) refers to the state of consciousness one needs to cultivate (Wong, 2021). This cultivated state of consciousness heightens our awareness. This awareness enables us to respond to a negative situation rather than giving passive or aggressive reactions. The understanding also allows us to see challenges as tools for learning and polishing lives. Therefore, the study indicates the need to inculcate a gratitude-based culture so that employees are psychologically equipped to respond to inevitable stressors.

Theoretical Contributions of the Findings

The present study offers a significant theoretical contribution to the existing pool of knowledge on gratitude literature. First, it adds to the literature on workplace gratitude. Researchers have highlighted an acute shortage of empirical studies on gratitude in organizational settings (Fredrickson, 2004; Garg & Gera, 2019; Watkins, 2014). The current research tries to plug this gap by probing the role of gratitude in combating workplace incivility and workplace-bullying behavior. Second, the present study offers a unique contribution to positive organizational behavior. The results reveal that gratitude is a significant predictor of psychological capital.

The findings answer another theoretical gap highlighted by Avey et al. (2011). The authors stated that "very few studies measured anything about the formation of PsyCap. In other words, few have considered what is 'to the left' of PsyCap (i.e., the antecedents in a theoretical model)" (p. 148). Third, previous research has revealed several consequences of gratitude, such as well-being (Emmons & Mishra, 2011), organizational commitment (Buonocore & Russo, 2013), and enhancing productivity (Di Fabio et al., 2017). The current finding adds to the bucket by demonstrating its unique effectiveness in combating a toxic work environment and building a more collaborative work culture. The recent research is one of the pioneer studies in the workplace domain of gratitude and toxicity. The study also offers an organizational application of the moral effect theory and amplification theory of gratitude. It suggests that gratitude strengthens one's ability to withstand workplace toxicity by amplifying one's and others' goodness.

Practical Implications of the Study

Besides being an economic space, a workplace is also a social-psychological arena. Here employees come together to manifest positive emotions. It is done so that the workplace can be tolerable, mutually agreeable, and an incubator of innovation and creativity. However, triggers of workplace toxicity like bullying and incivility are significant concerns for business leaders and HR practitioners. Organization development

(OD) practitioners have encouraged experimentation with several techniques to combat negative workplace behaviors. One mechanism to reduce incivility costs in a workplace is to build a culture that rejects it (Porath & Pearson, 2010). In this background, the study's findings propose institutionalizing a gratitude-based organizational culture that rejects workplace toxicity and supports a collaborative work environment. Collaborative work culture is beneficial to an organization in many ways. First, it fosters team spirit; second, it nurtures innovation and creativity; third, it builds a competitive advantage. Gratitude interventions and practices help employees acknowledge, appreciate, and notice the organizational resources' contribution to their well-being. A 10-week intervention program found that employees in their experimental group had higher work-related fulfillment and self-efficacy than the control group (Harty et al., 2001). Some of the prevalent methods of gratitude interventions at work are weekly gratitude diaries (Frost et al., 2007), taking a gratitude break to thank someone in meetings (Stegen & Wankier, 2018), gratitude letters (Sergeant & Mongrain,2011), or a random act of kindness and compassion.

A planned and formal gratitude institutionalization program may establish gratitude at three levels-episodic, persistent, and collective (Fehr et al., 2017). The organization may instantly try to cash upon any pro-employee announcements at a serial level. The employer may elicit gratitude among employees through minimal but impactful interventions like giving its employees a "Thank You card" or "appreciation letter" or a personal video message. At a persistent level, the HR manager may encourage employees to practice gratitude interventions, such as gratitude journaling, opposition thinking, introspection, and mindfulness. Such interventions help develop a stable tendency to be grateful for different facets of organizational life. Finally, gratitude could be cherished collectively through organization-wide acceptance and promotion of acts of gratefulness. Various ceremonies like "Thanksgiving Day" and "Appreciation Day" may help. To encourage the persistent experience of gratitude, several MNCs have initiated a program to allow paid leaves to its employee to serve oppressed and poor people.

Similarly, the findings suggest that business leaders may leverage PsyCap interventions to combat the toxic work environment. PsyCap interventions have fostered unique positive work outcomes in the last 20 years (Dello Russo & Stoykova, 2015). Psycap interventions will not only create positive subjective experiences but, at the same time, broaden employees' outlooks and offer strength to employees' character. Apart from countering workplace toxicity, PsyCap interventions promote job performance, job commitment, job engagement, and organizational citizenship behaviors (Avey et al., 2010). Thus, these PsyCap interventions are doubly beneficial for modern organizations. Managers may arrange counseling and psychological therapy sessions through training, open and two-way communication, discussion forums, and special humorous lectures (Abubakar et al., 2018) for employees. These initiatives promote an optimistic perception of work and organization. Abubakar et al. (2018) highlighted the need for psychologists in modern organizations to manage toxicity in the workplace. The authors also suggested different psychological interventions for employees of different generations. Baby boomers require good anger and crisis management training. Generation X employees want to be trusted by their managers. They need training sessions to express their anger, frustrations, and negative feelings towards colleagues and organizations. Generation Y employees require open discussion forums with their colleagues and supervisors. Thus, employees of different generations need different types of psychological assistance (Abubakar et al., 2018).

Limitations and Future Directions

While important, the study results are bound by some limitations. The study has been conducted in Indian settings only. Given the vast cultural differences, the results may vary in non-Eastern locations.

Future researchers may be interested in leading a cross-cultural study to authenticate the present study's findings. Future researchers may also compare the effect of gratitude on workplace toxicity across employees of different genders, generations (baby boomers, X, Y, Z), experiences, and marital statuses. Such studies may use demographical variables as moderators for better understanding. They may also examine the moderating effect of gratitude in the causal linkage of workplace stressors and toxic work behavior. Gratitude could also be investigated as a variable that moderates workplace toxicity's influence on an employee's psychological and mental health. Second, because of the present study's cross-sectional nature, the study could not capture gratitude's long-term benefits. Thus, a longitudinal research design is better suited for claims of causality. Third, the current study focuses only on dispositional gratitude. Future researchers may explore the emotional aspect of appreciation to build up psychological resources and counter workplace toxicity. Fourth, the current study examined the mediating effect of only one variable, that is, PsyCap. Future researchers may explore the mediating impact of spirituality, compassion, and prosocial behavior. Fifth, the study relied on a single source of data. Future researchers strongly suggest using double source information for cross-validation of results. Sixth, the current sample size is 200. Future researchers may think of increasing the sample size. Finally, the possibility of self-reporting bias due to the self-report nature of the questionnaires used is a significant limitation of this study.

Conclusion

The present study explored the association between gratitude, Psycap, and workplace toxicity. Gratitude negatively predicted workplace toxicity and positively indicated Psycap resources. The role of Psycap resources of the employees was found to predict workplace toxicity negatively. It could be recommended that gratitude can help a modern organization in building up a collaborative work culture based on trust, team spirit, compassion, and resiliency.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Ethical Approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The study was partially funded by the ICSSR doctoral research grant received by the first author during her PhD program.

ORCID iD

Ayatakshee Sarkar https://orcid.org/0000-0002-9620-4395

References

Abubakar, A. M. (2018). Linking work-family interference, workplace incivility, gender, and psychological distress. *Journal of Management Development*, *37*(3), 226–242.

Abubakar, A. M., Megeirhi, H. A., & Shneikat, B. (2018). Tolerance for workplace incivility, employee cynicism, and job search behaviour. *The Service Industries Journal*, *38*(9–10), 629–643.

Anjum, A., & Ming, X. (2018). Combating toxic workplace environment. *Journal of Modelling in Management*, *13*(3), 23.

Avey, J. B., Reichard, R., Luthans, F., & Mhatre, K. (2011). A meta-analysis of the impact of positive psychological capital on employee attitudes, behaviours, and performance. *Human Resource Development Quarterly*, *22*, 127–152.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical consideration. *Journal of Personality & Social Psychology*, *51*, 1173–1182.

Buonocore, F., & Russo, M. (2013). Reducing the effects of work-family conflict on job satisfaction: The kind of commitment matters. *Human Resource Management Journal*, *23*(1), 91–108.

Buunk, B. P., Doosje, B. J., Jans, L. G., & Hopstaken, L. E. (1993). Perceived reciprocity, social support, and stress at work: The role of exchange and communal orientation. *Journal of Personality and Social Psychology*, *65*(4), 801.

Carmichael, F., Fenton, S. J. H., Pinilla-Roncancio, M. V., Sing, M., & Sadhra, S. (2016). Workplace health and well-being in construction and retail. *International Journal of Workplace Health Management*, *9*(2), 251–268

Clear, T. (2017). THINKING ISSUES What's driving uber? Values in computing and the 'sharing economy'. *ACM Inroads*, *8*(4), 38–40.

Cortina, L. M., & Magley, V. J. (2009). Patterns and profiles of response to incivility in the workplace. *Journal of Occupational Health Psychology, 14*(3), 272.

Dello Russo, S., & Stoykova, P. (2015). Psychological capital intervention (PCI): A replication and extension. *Human Resource Development Quarterly*, *26*(3), 329–347.

Di Fabio, A., Palazzeschi, L., & Bucci, O. (2017). Gratitude in organizations: A contribution for healthy organizational contexts. *Frontiers in Psychology*, *8*, 2025.

Duffy, M. (2009). Preventing workplace mobbing and bullying with effective organizational consultation, policies, and legislation. *Consulting Psychology Journal: Practice and Research*, *61*(3), 242.

Einarsen, S., Hoel, H., & Notelaers, G. (1994). Negative acts questionnaire—revised. *Work & Stress*.

Emmons, R. A. (2003). Personal goals, life meaning, and virtue: Wellsprings of a positive life.

Emmons, R. A., & Mishra, A. (2011). Why gratitude enhances well-being: What we know, what we need to know. In *Designing positive psychology: Taking stock and moving forward* (Vol. *8*, pp. 248–262). Oxford University Press.

Emmons, R. A., & Shelton, C. M. (2002). Gratitude and the science of positive psychology. In *Handbook of positive psychology* (Vol. *18*, pp. 459–471). Oxford University Press.

Emmons, R. A., & McCullough, M. E. (2003). Counting blessings versus burdens: an experimental investigation of gratitude and subjective well-being in daily life. *Journal of personality and social psychology, 84*(2), 377.

Fehr, R., Fulmer, A., Awtrey, E., & Miller, J. A. (2017). The grateful workplace: A multilevel model of gratitude in organizations. *Academy of Management Review*, *42*(2), 361–381.

Fredrickson, B. L. (2004). Gratitude, like other positive emotions, broadens, and builds. *The Psychology of Gratitude*, pp. 145–166.

Frost, P. J. (2007). *Toxic emotions at work and what you can do about them*. Harvard Business Press.

Garg, N., & Gera, S. (2019). Gratitude and leadership in higher education institutions. *Journal of Applied Research in Higher Education*, *12*(5), 915–926

Goldberg, S. B., Hanley, A. W., Baldwin, S. A., Bernstein, A., & Garland, E. L. (2020). Does mindfulness practice promote psychological functioning or is it the other way around? A daily diary study. *Psychotherapy*, *57*(3), 310–322.

Gupta, V., Makhecha, U. P., & Shaik, F. F. (2021). Era of disruption: Opportunities and challenges for businesses in India. *Business Perspectives and Research*, *9*(2), 192–194.

Hair, J. F., Anderson, R. E., Tatham, R. L., & William, C. (1998). *Black (1998), Multivariate data analysis*.

Harty, B., Gustafsson, J. A., Björkdahl, A., & Möller, A. (2016). Group intervention: A way to improve working teams' positive psychological capital. *Work*, *53*(2), 387–398.

Hobfoll, S. E. (2002). Social and psychological resources and adaptation. *Review of General Psychology*, *6*(4), 307–324.

Jans-Beken, L., Jacobs, N., Janssens, M., Peeters, S., Reijnders, J., Lechner, L., & Lataster, J. (2019). Reciprocal relationships between state gratitude and high-and low-arousal positive affects in daily life: A time-lagged ecological assessment study. *The Journal of Positive Psychology*, *14*(4), 512–527.

Jha, J. K., & Sud, K. (2021). Exploring influence mechanism of abusive supervision on subordinates' work incivility: A proposed framework. *Business Perspectives and Research*, *9*(2), 324–339.

Kline, R. B. (2008). *Becoming a behavioral science researcher: A guide to producing research that matters*. Guilford Press.

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping*. Springer publishing company.

Lave, L. B., & Upton, A. C. (1987). *Toxic chemicals, health and the environment*. *Vol*. *13*, No. 4, pp. 521–522.

Luthans, F., & Youssef-Morgan, C. M. (2017). Psychological capital: An evidence-based positive approach. *Annual Review of Organizational Psychology and Organizational Behaviour*, *4*(1), 339–366.

Luthans, F., Avolio, B. J., Avey, J. B., & Norman, S. M. (2007). Positive psychological capital: Measurement and relationship with performance and satisfaction. *Personnel Psychology*, *60*(3), 541–572.

Luthans, F., Avolio, B. J., Walumbwa, F. O., & Li, W. (2005). The psychological capital of Chinese workers: Exploring the relationship with performance. *Management and Organization Review*, *1*(2), 249–271.

Luthans, F., Luthans, K. W., & Luthans, B. C. (2004). Positive psychological capital: Beyond human and social capital. *Business Horizons*, *47*(1), 45–50.

McCullough, M. E., Emmons, R. A., & Tsang, J. (2002). The grateful disposition: A conceptual and empirical topography. *Journal of Personality and Social Psychology*, *82*(1), 112–127.

Megeirhi, H. A., Kilic, H., Avci, T., Afsar, B., & Abubakar, A. M. (2018). Does team psychological capital moderate the relationship between authentic leadership and negative outcomes: An investigation in the hospitality industry. *Economic research-Ekonomskaistraživanja*, *31*(1), 927–945.

Megeirhi, H. A., Ribeiro, M. A., & Woosnam, K. M. (2020). Job search behaviour explained through perceived tolerance for workplace incivility, cynicism, and income level: A moderated mediation model. *Journal of Hospitality and Tourism Management*, *44*, 88–97.

Peterson, C., & Seligman, M. E. (2004). *Character strengths and virtues: A handbook and classification* (Vol. *1*). Oxford University Press.

Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology*, *63*, 539–569.

Porath, C. L., & Pearson, C. M. (2010). The cost of bad behavior. *Organizational Dynamics*, *39*, 64–71.

Rosenberg, E. L. (1998). Levels of analysis and the organization of affect. *Review of General Psychology, 2*(3), 247–270.

Sarkar, A., & Garg, N. (2020). "Peaceful workplace" only a myth? Examining the mediating role of psychological capital on spirituality and nonviolence behaviour at the workplace. *International Journal of Conflict Management.*

Sheldon, K. M., & Yu, S. C. (2022). Methods of gratitude expression and their effects upon well-being: Texting may be just as rewarding as and less risky than face-to-face. *The Journal of Positive Psychology*, *17*(5), 690–700.

Sergeant, S., & Mongrain, M. (2011). Are positive psychology exercises helpful for people with depressive personality styles?. *The Journal of Positive Psychology*, *6*(4), 260–272.

Stegen, A., & Wankier, J. (2018). Generating gratitude in the workplace to improve faculty job satisfaction. *Journal of Nursing Education*, *57*(6), 375–378.

Sun, P., Sun, Y., Jiang, H., Jia, R., & Li, Z. (2019). Gratitude and problem behaviors in adolescents: The mediating roles of positive and negative coping styles. *Frontiers in Psychology*, *10*, 1547.

Tai, K., Narayanan, J., & McAllister, D. J. (2012). Envy as pain: Rethinking the nature of envy and its implications for employees and organizations. *Academy of Management Review*, *37*(1), 107–129.

Tastan, S. B. (2017). Toxic workplace environment: In search for the toxic behaviours in organizations with a research in healthcare sector. *Postmodern Openings*, *8*(1), 83–109.

Too, L., & Harvey, M. (2012). 'TOXIC' workplaces: The negative interface between the physical and social environments. *Journal of Corporate Real Estate*, *14*(3), 171–181.

Toor, S. U. R., & Ofori, G. (2010). Positive psychological capital as a source of sustainable competitive advantage for organizations. *Journal of Construction Engineering and Management*, *136*(3), 341–352.

Trunnell, E. P., & Braza, J. F. (1995). Mindfulness in the workplace. *Journal of Health Education*, *26*(5), 285–291.

Woo, C. H., & Kim, C. (2020). Impact of workplace incivility on compassion competence of Korean nurses: Moderating effect of psychological capital. *Journal of Nursing Management*, *28*(3), 682–689.

Wood, A. M., Joseph, S., & Linley, P. A. (2007). Coping style as a psychological resource of grateful people. *Journal of Social and Clinical Psychology*, *26*(9), 1076–1093.

Wong, S. (2021). Cultivating sacred moments: Evaluating a pilot program to foster psychospiritual well-being (Doctoral dissertation, Bowling Green State University).

Watkins, P. C. (2014). *Gratitude and the good life*. Springer.

Yay, M. (2017). The mediation analysis with the Sobel test and the percentile bootstrap. *International Journal of Management and Applied Science*, *3*(2), 123–125.

Materials Today Communications 35 (2023) 105473

Contents lists available at ScienceDirect

Materials Today Communications

journal homepage: www.elsevier.com/locate/mtcomm

Carboxy-methyl tamarind kernel gum based bio-hydrogel for sustainable agronomy

Ritu Malik [a], Sudhir G. Warkar [a,*], Reena Saxena [b,*]

[a] Department of Applied Chemistry, Delhi Technological University, Delhi 110042, India
[b] Department of Chemistry, Kirori Mal College, University of Delhi, Delhi 110007, India

ARTICLE INFO

Keywords:
Agronomy
Biopolymer
Superabsorbent Hydrogel
Soil-conditioner
Water-sustenance

ABSTRACT

The innovative superabsorbent bio-tech hydrogels were concocted by free radical polymerization of sodium-methacrylate & carboxymethyl tamarind kernel gum, for application in agronomical procedures as a soil conditioner & water harvester for sustainable agronomy. The structural morphologies of the hydrogel were characterized by techniques viz. Fourier transform infrared spectroscopy, scanning electron microscopy, and thermogravimetric analysis. The hydrogel was evidenced as a biodegradable hydrogel via the soil burial biodegradability test. The swelling behavior of hydrogel was analyzed as a function of temperature, pH, and salt solutions. The formulated hydrogel engrossed 248 ml/g water and 0.5% CMTKG-PSMA hydrogel amended soil revealed augmentation in the maximum water holding capacity (46.5%), porosity (14%), shoot length (43%), available water content (37%). A decrease in bulk density (8%) was also observed. The water intake was Fickian diffusion following Korsmeyer-Peppas Equation. These upshots were validated upon chickpea plants. These hydrogels proffer a promising substitute as a soil conditioner in agronomy.

1. Introduction

Water's pivotal role and exorbitant usage in agronomical practices viz. irrigation, nutrient and pesticide application, crop cooling, and frost control, has opened a new outlook in research to develop innovative technologies for water's sustenance by effectual use with less wastage for global economic balance and food security. Henceforth, various practices are being explored that make the plants accessible to high available water content and lower the frequency of irrigation [1].

Biopolymer-based organic hydrogels are greener and more eco-friendly practices for keeping the soil moistened and nourished. Their greener attributes viz. non-toxicity, biodegradability, and higher water intake and retention capability shield the environment and natural resources and enhance plant growth and crop productivity [2]. The hydrogel amendment of the soil for agronomy is technically, socially, economically, and environmentally sustainable. These organic hydrogels contribute beneficially towards the soil morphologies, density, texture, permeability, evaporation, and infiltration rates of water through it. They boost the water holding capacity, water use efficiency, soil infiltration rate, and permeability. And, at the same time, drive down soil compaction, surface water run-off, soil erosion, leaching, evaporation of water, and seepage into the ground. They make the water content of the soil readily available to the plants, owing to their property of adhesion to the roots.

Scientifically, hydrogels constitute a group of three-dimensional cross-linked polymeric materials, which owing to the presence of hydrophilic functionalities in them, have the potential to entrap or release variable amounts of water/biological fluids reversibly (swell or shrink), in response to the external stimuli (physical/chemical/biochemical - temperature, solvent, pH, electric field, etc.), without altering the core matrix framework of the polymers. Witchterle and Lim prepared hydrogel for the first time in the 1960 s [3]. Various methods of synthesis of hydrogels include graft polymerization [4], free radical polymerization [5], crosslinking [6], frontal polymerization [7], and ionic radiation technique [8]. Cross-linking in hydrogels is usually achieved either through physical linkages viz. Vanderwaal forces or electrostatic interactions or hydrogen bonding [9]; or through chemical linkages via covalent bonds [10] formation during polymerization. Hydrogels formed owing to chemical linkages [11] are preferred in comparison to hydrogels involving physical linkages as the physical linkages disrupt easily due to weaker interactions. Chemically crosslinked Hydrogels [12,13] have been successfully applied in various areas.

* Corresponding authors.
E-mail addresses: sudhirwarkar@gmail.com (S.G. Warkar), rsaxena@kmc.du.ac.in (R. Saxena).

https://doi.org/10.1016/j.mtcomm.2023.105473
Received 6 October 2022; Received in revised form 15 January 2023; Accepted 20 January 2023
Available online 31 January 2023
2352-4928/© 2023 Elsevier Ltd. All rights reserved.

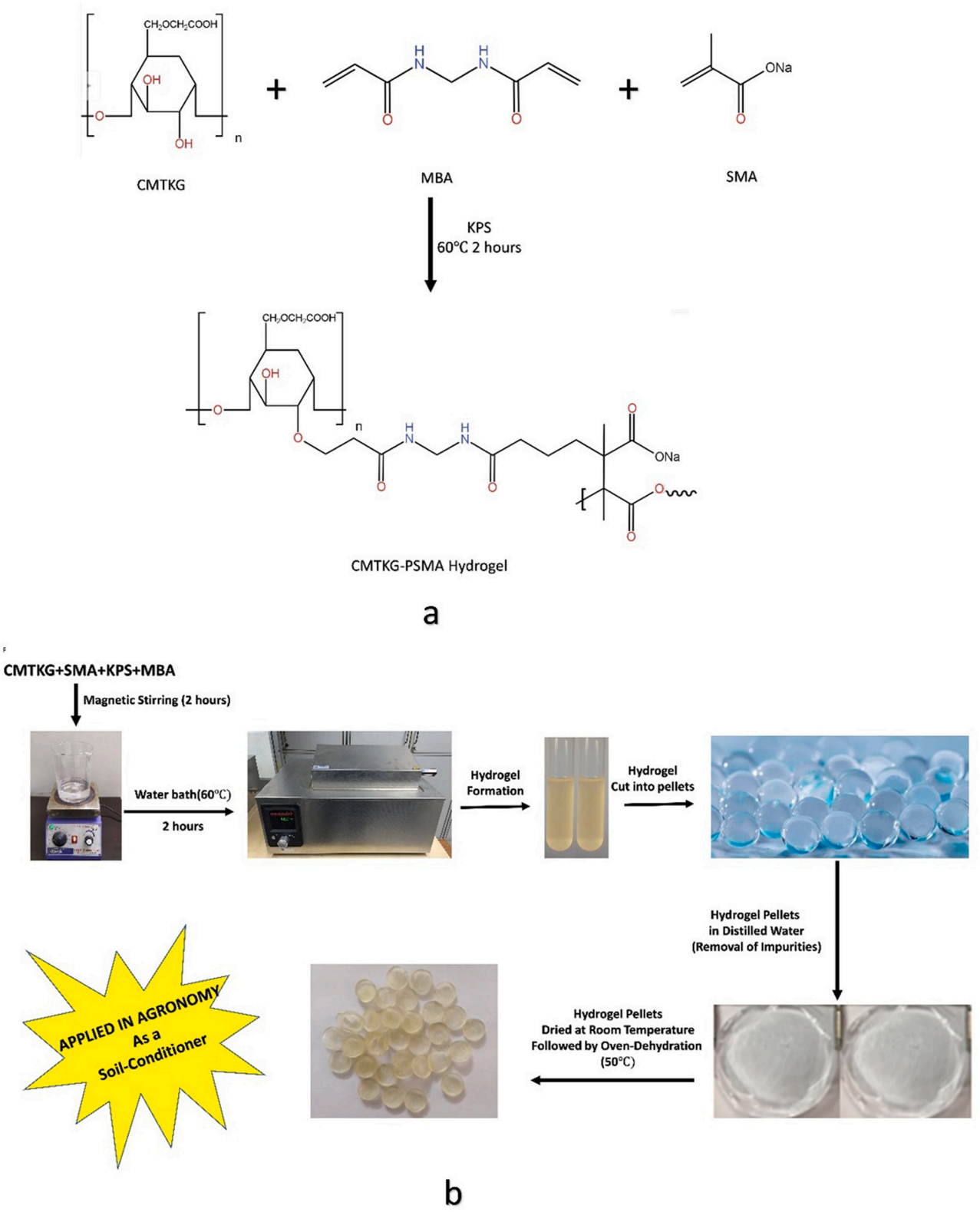


Fig. 1. Synthesis of CMTKG-PSMA Hydrogel. a) Mechanism b) Lay-out of Synthesis.

The hydrogels that can retain a large amount of water with a Swelling Ratio (SR) > 100 of water/biological fluids are termed Superabsorbent hydrogels (SAH) [14,15]. Crosslinking in the SAH is the reason behind the non-disintegration of the 3-dimensional structure of the SAH during swelling. The hydrophilicity of the SAH is due to the presence of hydrophilic functionalities (such as -COOH, $-NH_2$, $-CONH_2$, -CONH-, $-SO_3H$, -OH, etc.), capillary effect, and osmotic pressure [16]. Crosslinking can be attained either in vitro, i.e., during the preparation itself; or in vivo, i.e., after its administration, as in the case of biomedical application.

A low molecular weight crosslinking agent in trace amounts is incorporated along with the polymer into the reaction mixture. Owing to

Table 1
Various formulations for the synthesis of CMTKG-based Hydrogel with their Swelling ratios.

Hydrogel Formulation	Weight of CMTKG (g/10 ml distilled water)	Volume of Methacrylic Acid (ml)	Weight of NaOH (g/ 15 ml)	KPS (g)	MBA (g)	Swelling Ratio in Distilled Water [II]
HG-1	0.4	8.00	4.5	0.1	0.15	54.94
HG-2	0.4	8.00	4.5	0.1	0.1	84.05
HG-3	0.4	8.00	4.5	0.1	0.05	181.31
HG-4	0.1	8.00	4.5	0.1	0.15	66.25
HG-5	0.2	8.00	4.5	0.1	0.1	117.67
HG-6	0.3	8.00	4.5	0.1	0.05	248.88
HG-7	0.1	8.00	4.5	0.2	0.1	86.57
HG-8	0.2	8.00	4.5	0.2	0.1	87.21
HG-9	0.3	8.00	4.5	0.2	0.1	154.99
HG-10	0.1	8.00	4.5	0.05	0.1	104.09
HG-11	0.2	8.00	4.5	0.05	0.1	123.16
HG-12	0.3	8.00	4.5	0.05	0.1	117.84
HG-13	0.1	8.00	4.5	0.3	0.1	103.54
HG-14	0.2	8.00	4.5	0.3	0.1	107.04
HG-15	0.3	8.00	4.5	0.3	0.1	153.87
HG-16	0.1	8.00	4.5	0.1	0.1	100.11
HG-17	0.1	8.00	4.5	0.1	0.05	228.99
HG-18	0.2	8.00	4.5	0.1	0.15	72.14
HG-19	0.2	8.00	4.5	0.1	0.05	240.24
HG-20	0.3	8.00	4.5	0.1	0.15	89.04
HG-21	0.3	8.00	4.5	0.1	0.1	134.84

water thermodynamic compatibility, the linear polymer chains solubilize in water in the absence of crosslinking points. But, in their presence, a swelling force is generated due to the intake of water by the hydrogel, and at the same time, an elastic retractive force also originates due to the uncoiling and relaxation of the polymer. When these two opposing forces i.e., swelling force and elastic retractive force; become equal in magnitude, swelling reaches a steady state in equilibrium, that can be paralleled to the maximum swelling for that matrix [17,18]. When a hydrogel is subjected to some stimulus such as water stress/excess condition or changes in temperature and pH, the swelling degree changes resulting in swelling/shrinking of the hydrogel.

The swelling ratios of the hydrogels are interrelated to the nature of the polymer employed for its fabrication and the solvent in which swelling is done. The swelling studies of the optimized CMTKG-PSMA hydrogel were conducted in distilled water at different temperatures, in different salt solutions, and at various pH.

Once the water is absorbed by a hydrogel, it is released gradually as a response to stimulus, (usually following diffusion gradient and polymer relaxation), which humidifies the soil and enables the plants to cope with the water paucity, thus averting their collapse.

The SAH serves as water/nutrients supplementary capsules, that aid in their prolonged survival when the need arises due to droughts, or intensive crop practices. This SAH mitigates a rain-water harvester, as when they are incorporated into the soil, absorbs rainwater, thereby preventing its loss due to seepage or evaporation, and avails the water to the crop for irrigation when the rain stops. Biopolymer-based hydrogels may also be labeled as synthetic humus [19]. As humus keeps the soil fertile, pest-resistant, and disease-free and enables the plant to thrive well in drought and temperature extremes; so is the SAH. They have hydrophilic functionalities associated with them such as carboxylic acids, which have high electronegative oxygen atom, which forms a hydrogen bond with the electropositive hydrogen atom of a water molecule and thus encapsulates the water molecules inside them.

To deal with the agronomical problems faced, synthetic hydrogels have been introduced into the market. They solve the purpose but, being synthetic, are non-biodegradable, and generate permanent waste, a pitfall for our environment. To overcome this problem, biopolymers or natural polymers are being employed to fabricate biodegradable hydrogels, that are successfully applied as soil conditioners [20] and for nutrient/micronutrient supplementation [21,22]. Chitosan [3], starch [23], alginate [24], cellulose [25] and guar gum [26], carboxymethyl cellulose [27–29] are some of the natural polymers that are continuously being explored for the synthesis of SAH.

Eyeing the prospects of the necessity of a continuous reliant water supply system, that follows a greener approach, low maintenance and favors enhanced growth even under drought conditions; this research implicates the usage of Carboxy-Methyl Tamarind Kernel Gum (CMTKG), a bio-based biopolymer to synthesize an organic hydrogel for dynamic agronomics. CMTKG is classified as a natural biopolymer that is non-toxic, naturally abundant, economical, bio-degradable, and easily chemically derived from Tamarind Kernel gum (TKG). TKG is extracted from the seeds of Tamarindus indica L. It is a polysaccharide (Jellose), composed of glucose, xylose, & galactose units in the ratio of 3:2:1. It has a linear β-(1,4)-D glucose backbone with partial α-D-Xylo-pyranose (C-1) substitution at the O glucopyranosyl (C-6) residues, which are further substituted at C-2 by β-D galactopyranosyl units (C-1). TKG and alkaline mono-chloroacetic acid is subjected to Williamson's etherification (S_N2 mechanism) to form CMTKG, having carboxymethyl groups (-CH_2COOH) at its C-6 position. Its anionic sodium salt was applied for the hydrogel's synthesis. CMTKG resists biodegradation, has a high life span, more swelling power, and high solubility in cold water, and is stable in most solvents as compared to TKG [30].

Due to natural abundant availability and scarcer exploitation of CMTKG in agronomy[31], the fabricated CMTKG-PSMA Hydrogel is unique in application. Manuela et al. synthesized chitosan-based hydrogels with 68 g/g of water absorbance [32], whereas Kasinee Hemvichian et al.[33] reported the synthesis of Superabsorbent polymer exhibiting swelling ratio of 190 g/g of water absorbance. Further, J. Pushpamalar et al.[34] reported biodegradation less than 40% of the synthesized Hydrogel. Eyeing the present research, the novel fabricated CMTKG-PSMA hydrogel revealed 248 g/g of water absorbance with degradation greater than 44% on 45th day of soil burial. In addition to this, minimal addition of CMTKG-PSMA Hydrogel to soil is required for enhancing the water retention by the soil.

Reviewing the literature and identifying CMTKG as a research gap due to its worthy potential aspects and its scarcer exploration in agricultural practices; this research demonstrates pragmatical synthesis, optimization, and characterization of an organic superabsorbent hydrogel of CMTKG with sodium methacrylate for the first time for its application in the agronomical methods of soil conditioning and as a tool for water sustenance via water absorption assessment done by swelling studies, maximum water holding capacity analysis in the soil, available water content determination by pressure plate apparatus and water retention analysis by plantation method.

2. Experimental procedure

2.1. Materials & methods

CMTKG possessing 0.20° of carboxy-methyl substitution was amply aided by Hindustan Gum and Chemicals Limited. Haryana. Methacrylic acid, Potassium persulfate (KPS), and Sodium hydroxide (NaOH) used were of analytical reagent (AR) grade. N, N′-methylenebis(acrylamide) (MBA), was obtained from Merck, Germany. All the experimental solutions were made using distilled water and the pH of the solutions was maintained using HCl and NaOH.

2.2. Synthesis of organic CMTKG-PSMA SAH

The innovative organic superabsorbent hydrogel was synthesized (Fig. 1), using biopolymer CMTKG and sodium methacrylate as precursors. MBA was functional as cross-linking agent and KPS as an initiator. A homogenous mixture of CMTKG, Sodium methacrylate, MBA, and KPS was prepared using a magnetic stirrer. CMTKG was dissolved in 10 ml of distilled water followed by the addition of sodium methacrylate, prepared by adding 8 ml of methacrylic acid to the NaOH solution. After stirring magnetically for half an hour, MBA and KPS were added. Stirring was continued for a further half an hour. Once the homogenous solution was prepared, it was allowed to stand undisturbed for 10–15 mins and then poured into test tubes, to be kept in a water bath, maintained at 65℃, for two and half hours, when the cross-linking initiated and was completed through free radical copolymerization leading to the formation of SAH. The test tubes were broken to get cylindrical-shaped hydrogel, which was cut into slices of approximately equal dimensions and immersed in distilled water for twelve hours to do away with the impurities such as unreacted monomers, initiator, or crosslinker, any acidic or basic impurity, and any side-products owing to side-reactions. The swollen hydrogels were dried at room temperature (20℃) for about fifteen days followed by oven dehydration at a temperature of about 45℃, till weight constancy. About twenty-one Hydrogel formulations were synthesized by varying the biopolymer-CMTKG, cross-linker-MBA, and initiator-KPS, for optimization. These formulations are enlisted in Table 1.

2.3. Swelling assessment

An experimental outline of gravimetric measurement by the continuous method was processed to study the swelling behaviors of the synthesized hydrogels.

The parched dehydrated hydrogels were weighed precisely and immersed in distilled water at different temperatures viz. 28℃, 38℃, 48℃, and 58℃, in various pH solutions of 2.5, 4.2, 6.35, 9, 12, and in saline solutions of 0.9% NaCl solution, 0.9% $MgCl_2$ solution and 0.9% $AlCl_3$ solution. After a defined timeframe, the samples were taken out from the solutions, freed of excess surface water using filter paper, and weighed again. The weighing was repeated till a little decrease is observed in weight. The Equilibrium Swelling Ratio for each sample was computed using the formula specified below [35]:

Equilibrium Swelling Ratio (ESR) = $(W_E - W_D) / W_D$.

Where, ESR = Equilibrium Swelling Ratio.

W_E = Weight of the Hydrogel at equilibrium swelling.

W_D = Weight of the dehydrated Hydrogel.

The technique was repeated thrice with reproducible results.

2.4. Characterization of CMTKG-SMA SAH

2.4.1. Fourier transform infrared spectroscopy (FTIR)

The FTIR spectroscopic characterization technique was performed using a PerkinElmer FTIR Spectrophotometer (Model: Spectrum RX-1) in the range of 400–4000 cm^{-1}.

2.4.2. Thermal analysis

Thermo-gravimetric analysis (TGA) was effectuated employing a Perkin Elmer TGA, having nitrogen (N_2) atmosphere and temperature ranging from 25℃ to 900℃ with 10℃ /min of uniform heating rate.

2.4.3. Scanning electron microscopy (SEM)

SEM analysis was recorded to obtain the surface morphology of the fabricated CMTKG-SMA hydrogel using Scanning Electron Microscope (SEM) (JEOL, Japan Mode; JSM-6610LV).

2.5. Degradation studies

Soil burial degradability [36–39] test of CMTKG-PSMA hydrogel was performed for 60 days. The pre-weighed CMTKG-PSMA hydrogels were buried at a depth of about 5–6 cm in well-moisturized soil. After a specific time-interval and removal of soil, these hydrogels were re-weighed to obtain the degradation percentage based on the dry weight of hydrogel remnants in the soil post-degradation. SEM of the degraded hydrogel was also analyzed to identify the variations in the surface morphology. The residuary dry content (RDC) percentage was calculated implying the formula given below:

RDC % = Weight of the dry content of hydrogel X 100 /Weight of initial dry content of the hydrogel.

2.6. Agronomical applications of (CMTKG-PSMA) hydrogel as a soil conditioner

CMTKG-PSMA hydrogel was assessed for its potential to be exploited as a soil conditioner in dynamics of agronomics via its assimilation into the soil and researching its effects on various physical soil attributes viz. water retention, available water content (AWC), maximum water holding capacity (MWHC), porosity, particle density, bulk density, employing different relevant procedures [26]. The soil sample for experimentation was amassed from Delhi Technological University, Delhi, having a mineral portion of 87% sand and 0.8% of clay and silt. The calculated particle density and bulk density were found to be 2.692 g cm^{-3} and 1.28 g cm^{-3} respectively.

2.6.1. Maximum water holding capacity (MWHC)

The effect of CMTKG-SMA hydrogel on the MWHC of the soil was investigated by preparing three samples of soil with 0.1%, 0.3%, and 0.5% hydrogel along with a control sample. The fixed weighed quantity of each of the properly dried samples was put into a plastic beaker having small holes layered up with filter paper to avoid any errors due to soil loss. All four beakers were kept in a water tub to allow maximum water absorption for about twelve hours, followed by gravitational drainage of extra water. The net weight of water engrossed by the soil was calculated by using the following formula:

MWHC = [(Weight of beaker post water absorption) – (weight of the dried soil + weight of empty beaker + weight of wet filter paper)].

2.6.2. Available water content (AWC) and water retention capacity (WRC)

2.6.2.1. Determination of AWC using pressure plate apparatus. Pressure plate apparatus (Model: 1600 pressure extractor, Santabarbara, CA, USA) was employed to study the effects of CMTKG-SMA hydrogel on the water retention capacity at matric tensions of 0.33 bar and 15 bar, which correspond to the field capacity and permanent wilting point respectively. The three soil samples of 0.1%, 0.3%, and 0.5% hydrogel concentrations along with blank samples were settled one by one, in rubber rings arranged on ceramic plates. The soil was saturated with water for twelve hours. The plates were then subjected to different matric tensions in a pressure chamber, till the water flow from the soil ceases completely. The soil samples were transferred to moisture boxes and weighed accurately, followed by oven-dehydration (100℃ for

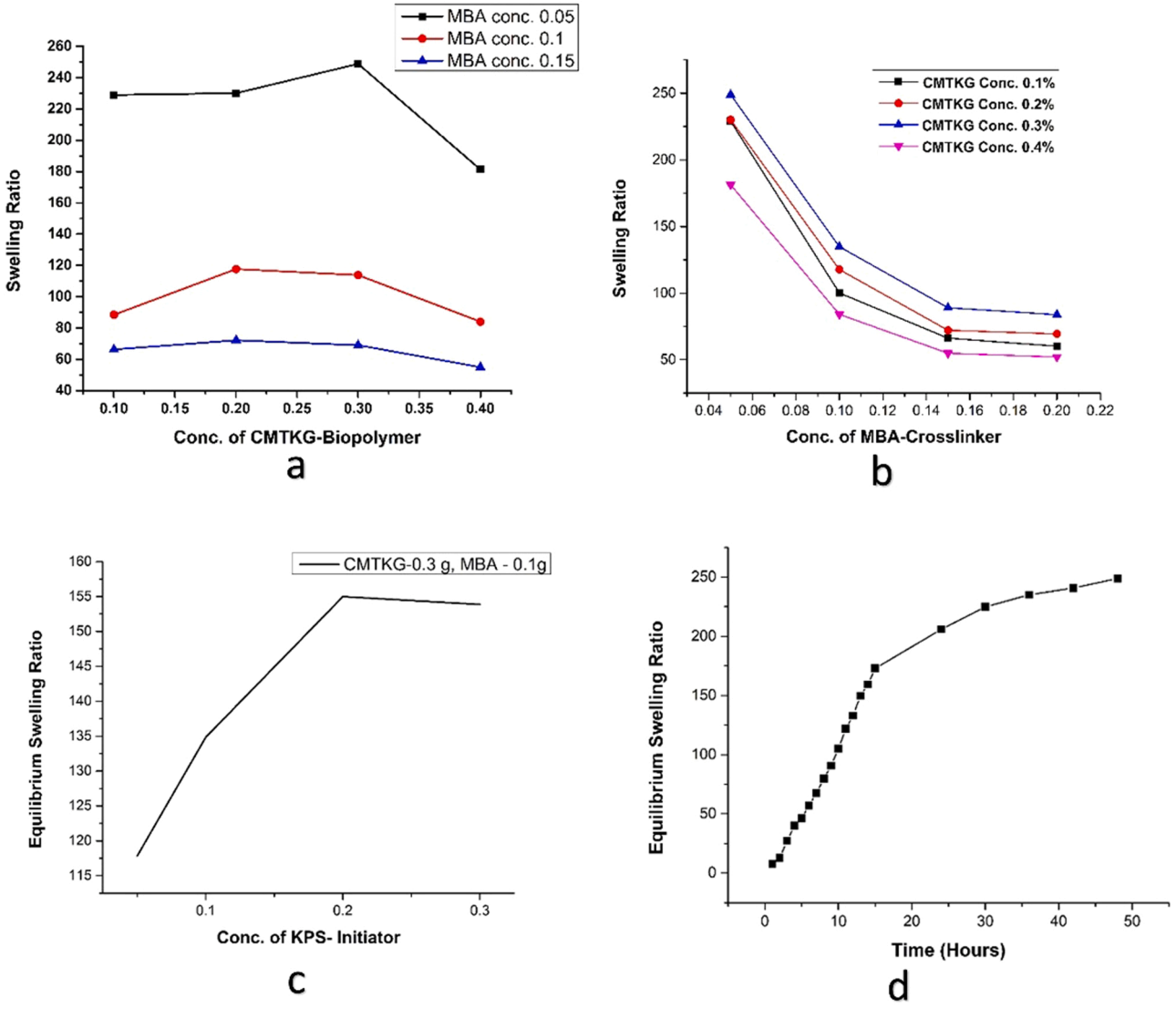


Fig. 2. Effect of various parameters on Equilibrium Swelling. Effect of Biopolymer on Swelling b) Effect of Cross-linker on Swelling c) Effect of KPS-Initiator on Swelling d) Rate of Swelling in Distilled Water with respect to time.

twelve hours) and reweighing after cooling. The percentage of water content (WC) was calculated using the following formula:

$$WC\left(\%\frac{w}{w}\right) = \frac{Swet - Sdry}{Sdry} X100$$

Where WC = Water content of the soil on a weight basis.

S_{wet} = Weight of the wet soil at specific matric tension.

S_{dry} = Weight of the dried soil.

Available water content (AWC).

AWC is the water that can be used up by the plants for their growth and is calculated by subtracting the water content retained at 0.33 matric tension and water content retained at 15 bar matric tension corresponding to field capacity and permanent wilting point respectively, using the following formula:

AWC= WC (0.33 at matric tension) – WC (15 bar matric tension).

2.6.2.2. Determination of water retention capacity by plantation method. To study the effect of CMTKG-SMA on the growth of plants with the limited amount of water mimicking the droughts conditions, chickpea seeds (*Cicer arietinum*) were sown in three soil samples having 0.1%, 0.3%, 0.5% hydrogel concentration in addition to the control sample, taking the same amount of the soil, irrigating with the same amount of water, and maintaining the same environmental conditions. The growth in terms of the days of survival and shoot length was observed for different concentrations of the hydrogel. The procedure was triplicated.

2.6.3. Determination of soil density and porosity

The standard method of calculating the particle density (PD) (employing the pycnometer) and bulk density (BD) was implied [26]. Oven-dried samples (100°C), till weight constancy of different concentrations of hydrogel, and after filtration through a 2 mm sieve were used for experimentation.

The porosity [40] of the different soil samples was also calculated, using the following formula:

Porosity (%) = [1 - $\frac{BD}{PD}$] X 100.

Where BD = Bulk density.

PD = Particle density.

3. Results and discussions

3.1. Mechanism of the synthesis of CMTKG-PSMA hydrogel

The CMTKG-PSMA hydrogel was successfully synthesized using KPS as initiator and MBA as a crosslinker, via free radical polymerization of

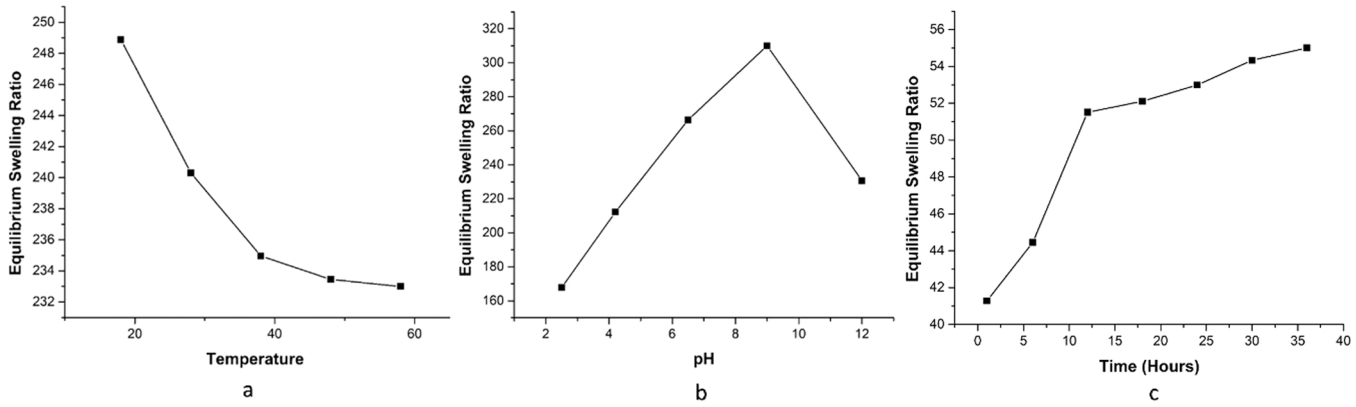


Fig. 3. Variance of Swelling ratios with a) Temperature b) pH c) in saline solution with time.

the sodium salt of methacrylic acid and CMTKG in the water bath at about 60℃ for about two hours. The monomer was incorporated into the biopolymer to enhance the hydrophilicity of the fabricated polymer network. The redox initiator KPS undergoes decomposition to persulphate radical, which abstracts hydrogen from the -OH group of the biopolymer-CMTKG and from the unsaturated double bond of SMA, to generate active radical sites leading to polymerization. As MBA is a bifunctional crosslinking agent, it crosslinked biopolymer CMTKG and PSMA to form a hydrogel framework. The proposed mechanism is depicted in Fig. 1**(a)**. As already discussed, (Table 1), 21 formulations with varied concentrations of biopolymer CMTKG, initiator, and crosslinker were prepared and researched to obtain the best hydrogel to function as a soil conditioner in agronomy. The complete layout of the Synthesis of Bio-Hydrogel is depicted in the Fig. 1**(b).**

3.2. *Effect of biopolymer*

The water absorption was assessed in the different formulations of hydrogels by altering the concentration of biopolymer-CMTKG and maintaining the constancy in the concentration of initiator and concentration of crosslinker. As the concentration of biopolymer is increased, the swelling ratio increases to a particular critical value followed by a further decrease as depicted in Fig. 2**(a)**. The initial increase was observed as we increased the concentration of CMTKG from 0.1 to 0.3 g/10 ml of water. This can be attributed to the increase in the number of hydrophilic groups and simultaneously enhanced repulsion between them due to like charges, which in turn create spaces allowing the water molecules to be absorbed. Further, if the concentration of CMTKG is increased to 0.4 g/10 ml of distilled water, a downfall in the swelling ratio is observed [41]. This is ascribed to an increase in viscosity leading to lesser contact and non-uniform mixing of reactants, resulting in non-formation of the desired hydrogel. Further, an attempt to prepare hydrogel with a CMTKG concentration of 0.5 g/ml of distilled water was unsuccessful due to buildup of high viscosity in solution.

3.3. *Effect of cross-linking agent*

Cross-linking density is the major determinant factor for the observed swelling ratios of CMTKG-PSMA hydrogels following Flory's - Rehner theory [42,43], as can be inferred from Fig. 2**(b)** Higher hydrodynamic 3-dimensional free volume exists at lower values of cross-linking densities as the network is loose-fitting, leading to more retention of water molecules and higher swelling. Keeping other reaction parameters constant, as the concentration of cross-linker-MBA was increased from 0.05 g to 0.15 g, the swelling ratios showed a substantial decrease as anticipated due to a rise in cross-linking density. The hydrogel formulation with 0.04 g MBA was also tried, but only a jelly-like mass was formed, revealing that a critical lower concentration of 0.05 g crosslinker is required for the successful fabrication of hydrogel to entrap water molecules.

3.4. *Effect of initiator*

It was observed that when the concentration of initiator - potassium persulfate (KPS), was increased from 0.05 g to 0.3 g, maintaining the constancy of other reaction constraints, the swelling ratios revealed a linear increase [44] as inferred from Fig. 2**(c)**. The highest swelling ratio (153.87) was observed for hydrogel with an initiator concentration of 0.2 g. But, with a subsequent increase in the concentration of the initiator, the swelling ratio decreases. The probable reason is that as the concentration of free radicals increases, the probability of a reaction between the free radicals increases i.e., they act as scavengers and inhibit the formation of free radicals to initiate the reaction [45].

3.5. *Swelling assessment*

3.5.1. *Effect of temperature on swelling assessment*

The optimized CMTKG-PSMA hydrogel exhibited a swelling ratio of 248.88 in distilled water at 18℃ as depicted in Fig. 2**(d)**), the temperature at which optimization was done. As the temperature increased to 58℃, the swelling ratio dropped to 232.99 as can be seen in Fig. 3 **(a).** At lower temperatures, hydrogen bonding exists between the hydrophilic moiety and water molecules in CMTKG-PMSA hydrogel. The water present at this temperature exists in a bound state. As the enthalpy rises at the higher temperature, hydrogen bonding breaks and the bound water becomes free water and is released from the hydrogel, resulting in the observed lower swelling ratio value [46].

3.5.2. *Effect of pH on swelling assessment*

As the pH of the solutions was increased, -COOH moiety existed as $-COO^-$. This led to increased repulsion between them, allowing more water molecules to enter the hydrogel network. Thus, an increase in swelling ratio was observed till the pH 10 [47]. But, as the pH was further increased, the swelling ratio decreased as NaOH was added to increase the pH and more concentration of Na^+ ions decreased the charge density, thus lowering the swelling ratio of the hydrogel [48], as depicted in Fig. 3**(b).** The swelling ratio at pH 2 was noted to be 167.79, which increased to 309.9 at pH 10 and finally decreased to 230.65 at pH 12.

3.5.3. *Swelling assessment in salt solutions*

The rate of equilibrium swelling was assessed for NaCl and was found to be 55 as shown in Fig. 3 **(c),** i.e., towards the lower side as compared to distilled water, the reason being the decrease in charge density due to surplus Na^+ cations, owing to sodium salt [49]. The higher ionic strength in the surrounding outer medium in saline solutions, reduces

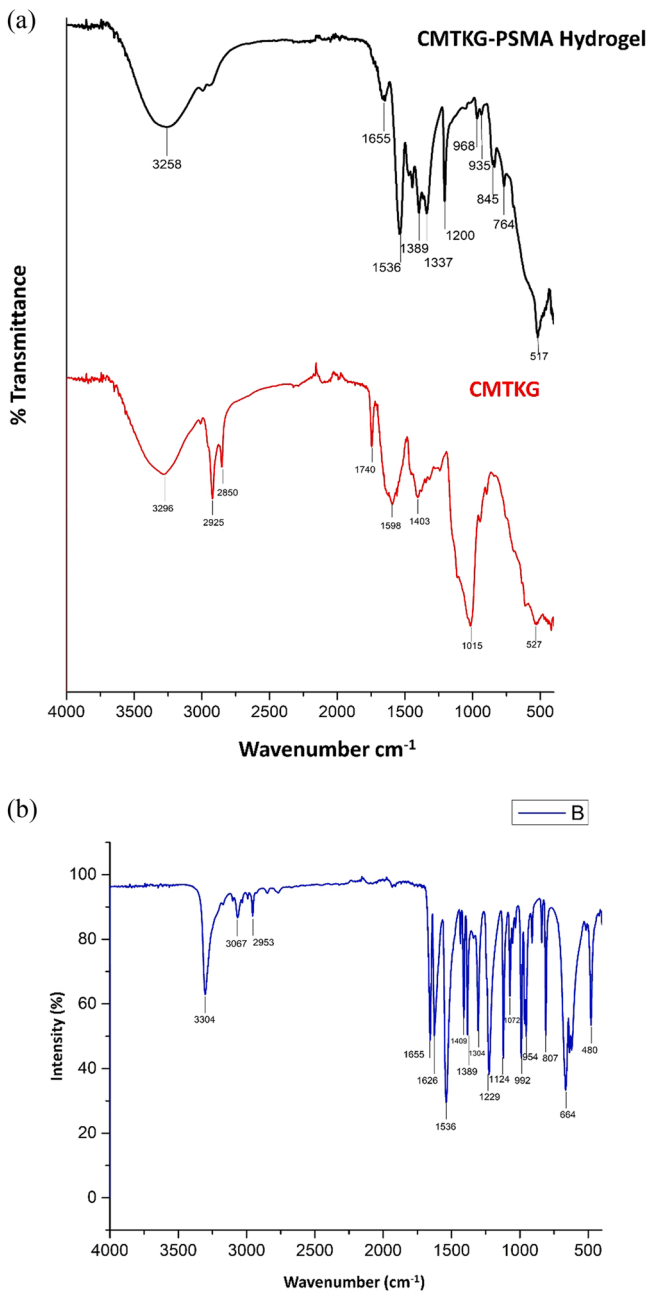


Fig. 4. a) FTIR Spectra of CMTKG & CMTKG-PSMA Hydrogel. Fig. 4b**)** FTIR Spectra of N,N-Methylene Bisacrylamide (MBA).

the osmotic pressure difference between the outer and inner medium, limiting the entry of water inside the hydrogel. Further, the lowering in the swelling ratio was also noted for $MgCl_2$ and $AlCl_3$. As the charge on the cation is increased, more -COO^- chains need to be attached to it for charge neutralization, sterically hindering the water molecules to be entrapped [50]. The observed order for the swelling ratio was found to be: NaCl (55) > $MgCl_2$ (44) > $AlCl_3$.

3.6. Characterization

3.6.1. FTIR spectroscopy

In the FTIR of CMTKG as shown in Fig. 4 **(a)**, a broad band is observed at 3296 cm^{-1} due to the -OH stretch [51], which in the CMTKG-PSMA hydrogel shifts to a lower frequency at 3258 cm^{-1} and is broadened due to intense hydrogen bonding. Further, symmetric, and asymmetric stretch of -COO moiety in CMTKG is observed at 1403 cm^{-1} and 1598 cm^{-1} and in CMTKG-PSMA, at 1389 cm^{-1} and 1536 cm^{-1}.

The intense peak at 1536 in CMTKG-PSMA, can be linked to -C═O asymmetric stretching frequency, which is usually observed at 1558 cm^{-1} in sodium methacrylate, verifying the decrease in conjugation and hence, polymerization [52]. Further, the intensity of CH_2 wagging observed in sodium methacrylate at 890 cm^{-1} is less intensified and observed at 845 cm^{-1}, which is characteristic of the vinylidene group.

In the FTIR spectra of MBA Fig. 4 **(b)**, the intense peaks at 664 cm^{-1} and 1304 cm^{-1} are attributed to the O═C-N and C-N stretch. But, in CMTKG-PSMA, these peaks shift to 764 cm^{-1} and 1337 cm^{-1} respectively, confirming the loss of conjugation due to cross-linking. Further, the peak observed at 1626 cm^{-1} due to C═C stretch, is also less intensified and observed at 1655 cm^{-1}, revealing the cross-linking. The peaks at 1200 cm^{-1} and 517 cm^{-1} are characteristic of the CMTKG skeleton. N-H bending is observed at 1200 cm^{-1}.

3.6.2. Thermal analysis

The probable thermal stability of the CMTKG-PSMA hydrogel is supposed to be more as compared to the CMTKG and the same can be verified by the thermograms as shown in Fig. 5**(a)** and 5**(b).** In both thermograms, three distinct zones of weight loss are revealed [52]. The weight loss in the first stage may be attributed to the removal of moisture; in the second stage, due to degradation of the polymeric backbone viz. degradation of -COOH and -OH moieties; and in the third stage, degradation of the remains. In the TGA Curve of CMTKG-PSMA Hydrogel (Fig. 5**(a)**), further thermal stability is observed after the loss of water of crystallization from 178.63℃- 427.89℃, with negligible weight loss (6.83%), which is not observed in case of CMTKG (Fig. 5**(b)**). The derivative weight curve indicates that the whole weight loss is an endothermic process.

3.6.3. Bio-degradation studies by scanning electron microscopy (SEM)

In the present research, the major problem of the non-biodegradability of synthetic hydrogels was resolved comparatively by fabricating the hydrogels, using natural biopolymer- CMTKG along with sodium methacrylate. Soil burial studies were carried out, which revealed that the hydrogel synthesized can be degraded by micro-organisms existent in soil and all the factors responsible for the rapid growth of microbes such as the significant ratio of carbon and nitrogen [53], oxygen concentration, temperature, humus content, and minerals present in the soil also enhanced process of biodegradation [54] of CMTKG-PSMA. The preliminary bio-degradation rate of the CMTKG-PSMA for the initial 30 days was high and found to be 44% followed by a slower weight loss of about 12% from the 30th day to the 45th day of soil burial. The initial rate of degradation was high as both the organic content (CMTKG) and oxygen present inside the hydrogel were high. But post the 45th day of soil burial, the rate of degradation was observed to be slower, as the anaerobic conditions prevail causing lower microbial activity and there is lesser diffusion of oxygen into the hydrogel because of more water inflow. The hydrogels buried in soil having more microorganisms endured faster biodegradation [55]. This bio-degradation process was verified by the SEM studies (Fig. 6), which revealed a comparative more eroded and more hollow-porous surface morphology of the decomposed hydrogel on the 45th day as compared to decomposed hydrogel on the 60th day as compared to non-degraded one, as can be easily inferred from the Fig. 6.

3.7. Diffusion

The Korsmeyer-Peppas diffusion equation [56] is applied to find the type of diffusion observed for the starting 15 h, as shown in Fig. 7**(a)**

$$F = \frac{Mt}{M\infty} = kt^n$$

As the value of n = 1, as inferred from the graph (Fig. 7 **a),** the

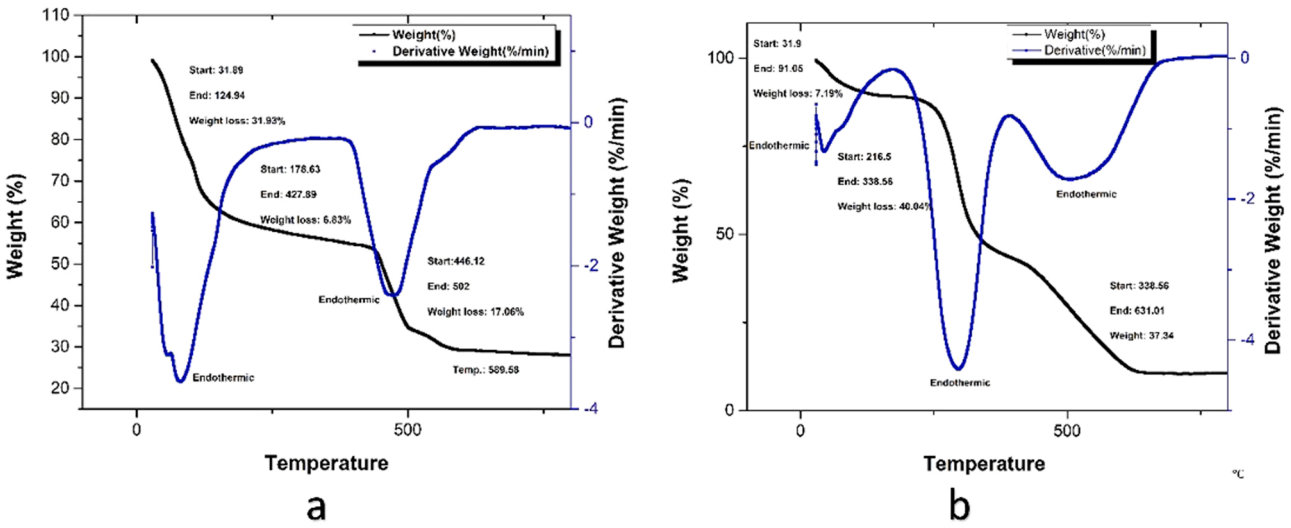


Fig. 5. TGA curves a) CMTKG-PSMA Hydrogel b) TGA Curve of CMTKG.

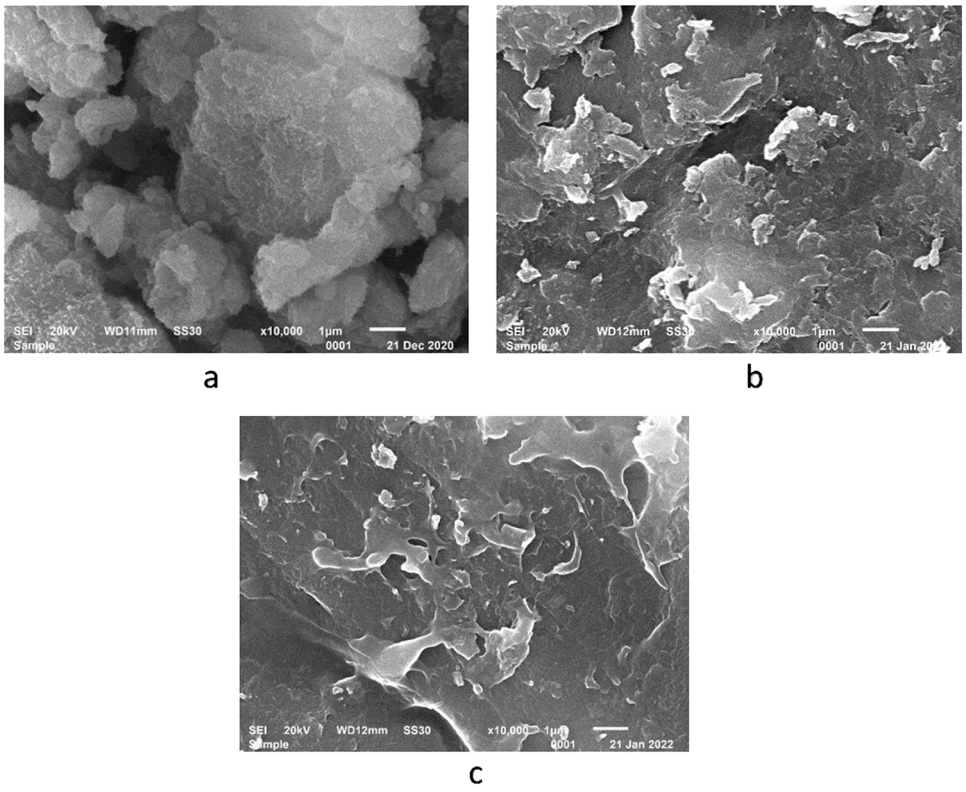


Fig. 6. SEM Studies. Non-degraded CMTKG-PSMA Hydrogel b) On the 45th Day of Decomposition c) On the 60th Day of Decomposition.

migration of water into the hydrogel involves a diffusion mechanism.

3.8. Agronomical applications of CMTKG-PSMA hydrogel

3.8.1. Maximum water holding capacity (MWHC)

An experimental determination of MWHC was conducted by saturating the different samples of the soil with water followed by gravimetric discharge of excess water without the application of any external pressure [26]. The control sample held 40.6327 g of water per 100 g of the soil, i.e., the percentage of MWHC is 40.63%. The different soil samples having 0.1%, 0.3%, and 0.5% CMTKG-PSMA hydrogel held 45.1305 g, 52.2929 g, and 59.5271 g of water per 100 g water respectively showing 11.06%, 28.68%, and 46.5% increase in MWHC percentage as compared to control sample. The concentration of hydrogel in the soil sample and the MWHC of the soil are found to be linearly related as can be inferred from the graph shown in Fig. 7**(b).**

3.8.2. Soil density & porosity

Particle density is significant for the approximation of porosity, air voids, and settling rates of particles in liquids. A pycnometer was employed as it can be filled with precise volume owing to excess water elimination through a capillary opening in the ground-glass stopper, leaving little scope for errors. Particle density was calculated by using the formula [57]:

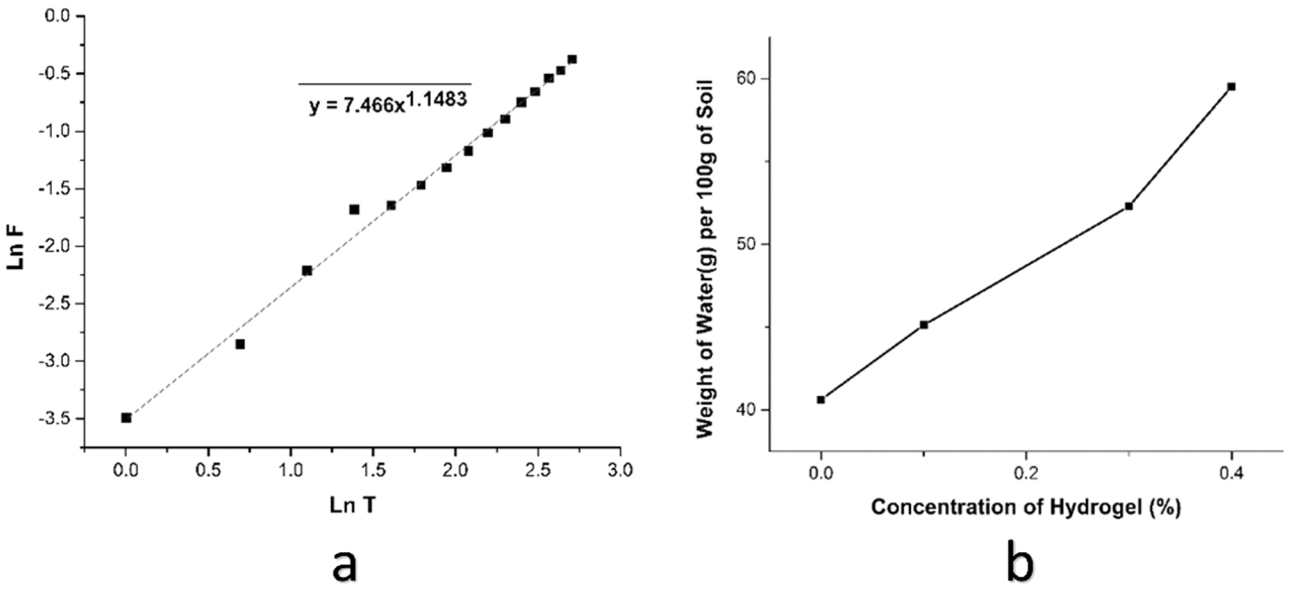


Fig. 7. a) Graph of Ln F vs. Ln T (Non-Fickian Diffusion) **b)** Maximum Water Holding Capacity (MWHC).

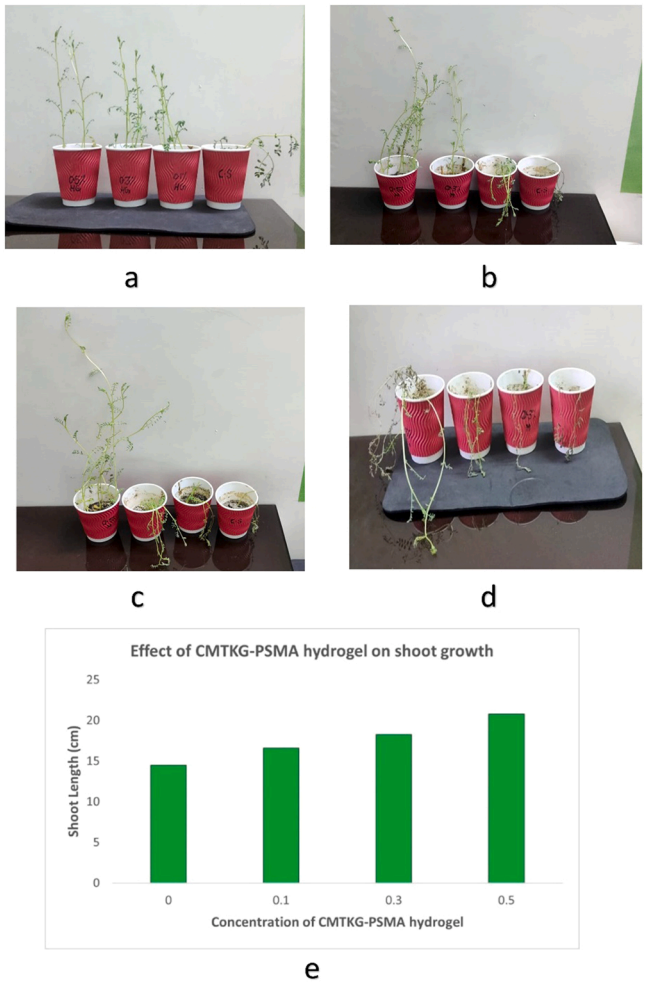


Fig. 8. Effect of Hydrogel on Water Retention Capacity of the Soil. a) After 21 days of plantation b) After 33 days of plantation c) After 40 days of plantation d) After 47 days of plantation e) Effect of Hydrogel on Shoot Growth.

$$\text{Particle density(PD)}, \rho p = \rho w \frac{(Ws - Wa)}{(Ws - Wa) - (Wsa - Ww)}$$

Where W_a = Weight of oven-dried pycnometer.
W_s = Weight of pycnometer plus soil sample.
W_{sa} = Weight pycnometer plus soil sample plus water.
W_w = Weight of pycnometer plus water at room temperature.
ρ_w = Density of water at room temperature.

The particle density of the soil sample was observed to be 2.692 g cm^{-3}.

Bulk density (BD) was noted for various soil samples using the standard procedure [40]. BD of the control sample was found to be 1.28 g cm^{-3} and of the soil samples containing 0.1%, 0.3%, and 0.5% CMTKG-PSMA hydrogel was found to be 1.24 g cm^{-3}, 1.20 g cm^{-3}, 1.18 g cm^{-3} respectively, revealing 3.2%, 6%, and 8% decrease in relation to the control sample.

CMTKG-PSMA hydrogel acts as a humus due to the presence of hydrophilic groups in them, which improves the porosity of the soil [58]. Porosity increased with the increased percentage of CMTKG-PSMA hydrogel and was revealed to be 52.4, 53.9, 55.4, and 60 for the control sample, and samples containing 0.1%, 0.3%, and 0.5% CMTKG-PSMA hydrogel, resulting in 2.9%, 6.7% and 14.5% increase with respect to the control sample respectively. CMTKG-PSMA hydrogel has the competency to decrease the BD and increase the porosity paving the way to more soil aeration which has positive effects on root functionality in plants [58].

3.8.3. *Water retention capacity (WRC) by plantation method*

The effect of CMTKG-PSMA hydrogel on WRC was analyzed by planting chickpeas in four pots containing control soil and soil having 0.1%, 0.3%, and 0.5% CMTKG-PSMA hydrogel. All four pots were subjected to identical conditions of sunlight, water, and temperature. It was observed that post 21 days (Fig. 8) of the plantation, the plant grown in the controlled soil wilted followed by wilting of 0.1%, 0.3%, and 0.5% on the 33rd, 40^{th,} and 47th day of the plantation, respectively (Fig. 8). Thus, it was discovered that the plants grown in CMTKG-PSMA hydrogel-treated soil with its higher concentration, survived for a longer duration under water stress conditions. A trilogy of experiments was done for this research [59].

Further, there was enhanced growth in plants sown with a higher percentage of CMTKG-PSMA hydrogel. On the 25th day of the plantation, the shoot length of plants grown in control soil and soil having

Table2
Percentage Increase in Available Water Content on Increasing the Concentration of Hydrogel.

Soil Sample % CMTKG-PSMA	Water content at 0.33 Bar matric suction (Field capacity) (%)	Water content at 0.33 Bar matric suction (Permanent wilting point) (%)	Available water content (%)
NIL	14.111	8.079	6.032
0.1%	15.246	7.245	7.331
0.3%	17.241	8.539	8.001
0.5%	24.306	16.04	8.266

0.1%, 0.3%, and 0.5% was observed to be 14.5 cm, 16.6 cm, 18.3 cm, and 20.8 cm, and a percentage increase in shoot length was found to be 14.4%, 26.2%, and 43.4% respectively as can be inferred from Fig. 8**(e).**

3.8.4. Available water content (AWC) by Pressure-plate apparatus

Available water content is a crucial paradigm to analyze water content in the soil profile that is readily available for the plants and, for assessing the growth of the plants [60]. As the CMTKG-PSMA hydrogel reveals a positive effect on the AWC characteristic of the soil, its usage under water stress conditions or drought is a boon for agronomy. AWC of different soil samples along with control soil was carried out employing pressure-plate apparatus, in which the different soil samples were subjected to 0.33 Bar matric suction and 15 Bar matric suction, which corresponds to the field capacity (amount of water retained by the soil sample upon saturation followed by excess water drainage) and permanent wilting point (water content of the soil sample below which plants wilt) respectively for each soil sample. The difference between field capacity and permanent wilting point relates to available water content.

The water content of various soil samples at 0.33 Bar, 15 Bar matric suction and AWC has been specified in the Table 2. Percentage increase in the available water content of 0.1%, 0.3%, and 0.5% CMTKG-PSMA containing soil was found to be 21.55%, 32.64%, and 37.04% respectively.

4. Conclusion

Present research successfully fabricated an innovative biodegradable CMTKG-PSMA hydrogel, having 0.3% CMTKG, 0.1% KPS, and 0.05% MBA with an equilibrium swelling ratio of 248.88 mg/L in distilled water at 18℃. The swelling assessment was analyzed considering various parameters such as pH, temperature, and salt solutions. The formulated hydrogel when treated with soil, revealed positive effects on various soil characteristics viz. decrease in bulk density by 8%, an increase in porosity by 14%, and enhanced growth (increased shoot length) in plants. Further available water content was upshifted by 37% with an increased concentration of 0.5% CMTKG-PSMA hydrogel. All these attributes enhanced the water productivity of the plants, bestowing a conducive and safe environment for better growth of the crops in arid and semi-arid regions. This water retention CMTKG-PSMA biohydrogel entails lower application rates and can be applied as a proficient tool for higher water absorption for water sustenance and environment conservation in agronomics.

CRediT authorship contribution statement

Ritu Malik: Conceptualization, Methodology, Data curation, Writing – original draft, Data curation, Investigation. **Reena Saxena**: Software, Validation, Writing – review & editing. **Sudhir G. Warkar**: Supervision, Visualization, Validation, Writing – review & editing.

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.

Data Availability

No data was used for the research described in the article.

Acknowledgements

The authors would like to thank Delhi Technological University, New Delhi, India. Kirori Mal College University of Delhi, New Delhi, India, for infrastructure and financial aid.

References

[1] Khushbu, S.G. Warkar, N. Thombare, Zinc micronutrient-loaded carboxymethyl tamarind kernel gum-based superabsorbent hydrogels: controlled release and kinetics studies for agricultural applications, Colloid Polym. Sci. 2021 2997 299 (2021) 1103–1111, https://doi.org/10.1007/S00396-021-04831-8.
[2] Khushbu, S.G. Warkar, Potential applications and various aspects of polyfunctional macromolecule- carboxymethyl tamarind kernel gum, Eur. Polym. J. 140 (2020), 110042, https://doi.org/10.1016/j.eurpolymj.2020.110042.
[3] S. Nangia, S. Warkar, D. Katyal, A review on environmental applications of chitosan biopolymeric hydrogel based composites, J. Macromol. Sci. Part A Pure Appl. Chem. 55 (2018) 747–763, https://doi.org/10.1080/10601325.2018.1526041.
[4] A. Pourjavadi, G.R. Mahdavinia, M.J. Zohuriaan-Mehr, Modified chitosan. II. H-chitoPAN, a novel pH-responsive superabsorbent hydrogel, J. Appl. Polym. Sci. 90 (2003) 3115–3121, https://doi.org/10.1002/app.13054.
[5] Y. Yue, X. Sheng, P. Wang, Fabrication and characterization of microstructured and pH sensitive interpenetrating networks hydrogel films and application in drug delivery field, Eur. Polym. J. 45 (2009) 309–315, https://doi.org/10.1016/j.eurpolymj.2008.10.038.
[6] G.B. Marandi, K. Esfandiari, F. Biranvand, M. Babapour, S. Sadeh, G. R. Mahdavinia, pH sensitivity and swelling behavior of partially hydrolyzed formaldehyde-crosslinked poly (acrylamide) superabsorbent hydrogels, Wiley Online Libr 109 (2008) 1083–1092, https://doi.org/10.1002/app.28205.
[7] Q.-Z. Yan, W.-F. Zhang, G.-D. Lu, X.-T. Su, C.-C. Ge, Frontal copolymerization synthesis and property characterization of starch-graft-poly(acrylic acid) hydrogels, Chem. A Eur. J. 11 (2005) 6609–6615, https://doi.org/10.1002/chem.200500554.
[8] H.A. Abd El-Rehim, E.-S.A. Hegazy, H.L. Abd El-Mohdy, Effect of various environmental conditions on the swelling property of PAAm/PAAcK superabsorbent hydrogel prepared by ionizing radiation, J. Appl. Polym. Sci. 101 (2006) 3955–3962, https://doi.org/10.1002/app.22904.
[9] A.R. Fajardo, M.B. Silva, L.C. Lopes, J.F. Piai, A.F. Rubira, E.C. Muniz, Hydrogel based on an alginate-Ca2+/chondroitin sulfate matrix as a potential colon-specific drug delivery system, RSC Adv. 2 (2012) 11095–11103, https://doi.org/10.1039/c2ra20785k.
[10] W.E. Hennink, C.F. van Nostrum, Novel crosslinking methods to design hydrogels, Adv. Drug Deliv. Rev. 64 (2012) 223–236, https://doi.org/10.1016/j.addr.2012.09.009.
[11] Y. Yue, X. Wang, J. Han, L. Yu, J. Chen, Q. Wu, J. Jiang, Effects of nanocellulose on sodium alginate/polyacrylamide hydrogel: mechanical properties and adsorption-desorption capacities, Carbohydr. Polym. 206 (2019) 289–301, https://doi.org/10.1016/J.CARBPOL.2018.10.105.
[12] Y. Yue, X. Wang, Q. Wu, J. Han, J. Jiang, Highly recyclable and super-tough hydrogel mediated by dual-functional TiO2 nanoparticles toward efficient photodegradation of organic water pollutants, J. Colloid Interface Sci. 564 (2020) 99–112, https://doi.org/10.1016/J.JCIS.2019.12.069.
[13] Y. Yue, S. Shen, W. Cheng, G. Han, Q. Wu, J. Jiang, Construction of mechanically robust and recyclable photocatalytic hydrogel based on nanocellulose-supported CdS/MoS2/Montmorillonite hybrid for antibiotic degradation, Colloids Surf. A Physicochem. Eng. Asp. 636 (2022), 128035, https://doi.org/10.1016/J.COLSURFA.2021.128035.
[14] R. Bhattacharyya, S.K. Ray, B. Mandal, A systematic method of synthesizing composite superabsorbent hydrogels from crosslink copolymer for removal of textile dyes from water, J. Ind. Eng. Chem. 19 (2013) 1191–1203, https://doi.org/10.1016/j.jiec.2012.12.017.
[15] M.R. Guilherme, F.A. Aouada, A.R. Fajardo, A.F. Martins, A.T. Paulino, M.F. T. Davi, A.F. Rubira, E.C. Muniz, Superabsorbent hydrogels based on polysaccharides for application in agriculture as soil conditioner and nutrient carrier: a review, Eur. Polym. J. 72 (2015) 365–385, https://doi.org/10.1016/j.eurpolymj.2015.04.017.
[16] J.M. Rosiak, F. Yoshii, Hydrogels and their medical applications, Nucl. Instrum. Methods Phys. Res. Sect. B Beam Interact. Mater. At. 151 (1999) 56–64, https://doi.org/10.1016/S0168-583X(99)00118-4.

[17] A. Blanco, G. González, E. Casanova, M.E. Pirela, A. Briceño, Mathematical modeling of hydrogels swelling based on the finite element method, Appl. Math. 04 (2013) 161–170, https://doi.org/10.4236/am.2013.48a022.

[18] H. van der Linden, J. Westerweel, Temperature-sensitive hydrogels, in: Encycl. Microfluid. Nanofluidics, Springer, US, 2008, pp. 2006–2009, https://doi.org/10.1007/978-0-387-48998-8_1538.

[19] A. Hüttermann, L.J.B. Orikiriza, H. Agaba, Application of superabsorbent polymers for improving the ecological chemistry of degraded or polluted lands, CLEAN Soil, Air, Water 37 (2009) 517–526, https://doi.org/10.1002/clen.200900048.

[20] L.Z. Linan, A.C.M. Cidreira, C.Q. da Rocha, F.F. de Menezes, G.J. d M. Rocha, A.E. M. Paiva, Utilization of acai berry residual biomass for extraction of lignocellulosic byproducts, J. Bioresour. Bioprod. 6 (2021) 323–337, https://doi.org/10.1016/J.JOBAB.2021.04.007.

[21] Y. Zhang, X. Tian, Q. Zhang, H. Xie, B. Wang, Y. Feng, Hydrochar-embedded carboxymethyl cellulose-g-poly(acrylic acid) hydrogel as stable soil water retention and nutrient release agent for plant growth, J. Bioresour. Bioprod. 7 (2022) 116–127, https://doi.org/10.1016/J.JOBAB.2022.03.003.

[22] R. Malik, R. Saxena, S.G. Warkar, Biopolymer-based biomatrices – organic strategies to combat micronutrient deficit for dynamic agronomy, ChemistrySelect 7 (2022), e202200006, https://doi.org/10.1002/SLCT.202200006.

[23] H. Ismail, M. Irani, Z. Ahmad, Starch-based hydrogels: present status and applications, Int. J. Polym. Mater. 62 (2013) 411–420, https://doi.org/10.1080/00914037.2012.719141.

[24] B. Song, H. Liang, R. Sun, P. Peng, Y. Jiang, D. She, Hydrogel synthesis based on lignin/sodium alginate and application in agriculture, Int. J. Biol. Macromol. 144 (2020) 219–230, https://doi.org/10.1016/j.ijbiomac.2019.12.082.

[25] C. Demitri, F. Scalera, M. Madaghiele, A. Sannino, A. Maffezzoli, Potential of cellulose-based superabsorbent hydrogels as water reservoir in agriculture, Int. J. Polym. Sci. (2013) (2013), https://doi.org/10.1155/2013/435073.

[26] N. Thombare, S. Mishra, M.Z. Siddiqui, U. Jha, D. Singh, G.R. Mahajan, Design and development of guar gum based novel, superabsorbent and moisture retaining hydrogels for agricultural applications, Carbohydr. Polym. 185 (2018) 169–178, https://doi.org/10.1016/j.carbpol.2018.01.018.

[27] K. Hemvichian, A. Chanthawong, P. Suwanmala, Synthesis and characterization of superabsorbent polymer prepared by radiation-induced graft copolymerization of acrylamide onto carboxymethyl cellulose for controlled release of agrochemicals, Radiat. Phys. Chem. 103 (2014) 167–171, https://doi.org/10.1016/j.radphyschem.2014.05.064.

[28] Y. Zhou, S. Fu, L. Zhang, H. Zhan, Superabsorbent nanocomposite hydrogels made of carboxylated cellulose nanofibrils and CMC-g-p(AA-co-AM, Carbohydr. Polym. 97 (2013) 429–435, https://doi.org/10.1016/j.carbpol.2013.04.088.

[29] W.B. Wang, A.Q. Wang, Preparation, swelling and water-retention properties of crosslinked superabsorbent hydrogels based on guar gum, Adv. Mater. Res 96 (2010) 177–182, https://doi.org/10.4028/www.scientific.net/AMR.96.177.

[30] H. Prabhanjan, Studies on modified tamarind kernel powder. Part I: preparation and physicochemical properties of sodium salt of carboxymethyl derivatives, Starch Stärke 41 (1989) 409–414, https://doi.org/10.1002/star.19890411102.

[31] Khushbu, S.G. Warkar, Potential applications and various aspects of polyfunctional macromolecule- carboxymethyl tamarind kernel gum, Eur. Polym. J. 140 (2020), 110042, https://doi.org/10.1016/J.EURPOLYMJ.2020.110042.

[32] M.M. Iftime, G.L. Ailiesei, E. Ungureanu, L. Marin, Designing chitosan based eco-friendly multifunctional soil conditioner systems with urea controlled release and water retention, Carbohydr. Polym. 223 (2019), 115040, https://doi.org/10.1016/J.CARBPOL.2019.115040.

[33] K. Hemvichian, A. Chanthawong, P. Suwanmala, Synthesis and characterization of superabsorbent polymer prepared by radiation-induced graft copolymerization of acrylamide onto carboxymethyl cellulose for controlled release of agrochemicals, Radiat. Phys. Chem. 103 (2014) 167–171, https://doi.org/10.1016/J.RADPHYSCHEM.2014.05.064.

[34] J. Pushpamalar, S.J. Langford, M.B. Ahmad, Y.Y. Lim, K. Hashim, Eco-friendly smart hydrogels for soil conditioning and sustain release fertilizer, Int. J. Environ. Sci. Technol. 15 (2018) 2059–2074, https://doi.org/10.1007/S13762-017-1598-2/FIGURES/6.

[35] Khushbu, S.G. Warkar, N. Thombare, Controlled release and release kinetics studies of boron through the functional formulation of carboxymethyl tamarind kernel gum-based superabsorbent hydrogel, Polym. Bull. (2021), https://doi.org/10.1007/S00289-021-03634-9.

[36] P. Rizzarelli, C. Puglisi, G. Montaudo, Soil burial and enzymatic degradation in solution of aliphatic co-polyesters, Polym. Degrad. Stab. 85 (2004) 855–863, https://doi.org/10.1016/J.POLYMDEGRADSTAB.2004.03.022.

[37] T. Kijchavengkul, R. Auras, M. Rubino, E. Alvarado, J.R. Camacho Montero, J. M. Rosales, Atmospheric and soil degradation of aliphatic–aromatic polyester films, Polym. Degrad. Stab. 95 (2010) 99–107, https://doi.org/10.1016/J.POLYMDEGRADSTAB.2009.11.048.

[38] R.L. Shogren, W.M. Doane, D. Garlotta, J.W. Lawton, J.L. Willett, Biodegradation of starch/polylactic acid/poly(hydroxyester-ether) composite bars in soil, Polym. Degrad. Stab. 79 (2003) 405–411, https://doi.org/10.1016/S0141-3910(02)00356-7.

[39] K. Sharma, V. Kumar, B.S. Kaith, V. Kumar, S. Som, S. Kalia, H.C. Swart, A study of the biodegradation behaviour of poly(methacrylic acid/aniline)-grafted gum ghatti by a soil burial method, RSC Adv. 4 (2014) 25637–25649, https://doi.org/10.1039/C4RA03765K.

[40] Manual on soil, plant and water analysis., (n.d.). https://www.cabdirect.org/cabdirect/abstract/20073118818 (accessed September 27, 2021).

[41] P. Goyal, V. Kumar, P..S.-C. Polymers, undefined 2007, Carboxymethylation of tamarind kernel powder, Elsevier. (n.d.). https://www.sciencedirect.com/science/article/pii/S0144861706004772?casa_token=2yJZTHny8QoAAAAA:NLbHCbIz3KhkhrM9mIFi2lXWL9YO0S7LdowGcXBXgdsQ9gmbwu8ZUR3CGtV2s5nM5iU7rQR3–3Y0 (accessed August 24, 2022).

[42] K.S. Soppimath, A.R. Kulkarni, T.M. Aminabhavi, Chemically modified polyacrylamide-g-guar gum-based crosslinked anionic microgels as pH-sensitive drug delivery systems: preparation and characterization, J. Control. Release 75 (2001) 331–345, https://doi.org/10.1016/S0168-3659(01)00404-7.

[43] F.T. WALL, Principles of Polymer Chemistry. Paul J. Flory. Cornell Univ. Press, Ithaca, New York, 1953. 688 pp. Illus. $8.50, Science (80–.). 119 (1954) 555–556. https://doi.org/10.1126/SCIENCE.119.3095.555-A.

[44] R. Singhal, R.S. Tomar, A.K. Nagpal, Effect of cross-linker and initiator concentration on the swelling behaviour and network parameters of superabsorbent hydrogels based on acrylamide and acrylic acid, Int. J. Plast. Technol. 131 (13) (2009) 22–37, https://doi.org/10.1007/S12588-009-0004-4.

[45] E.M. Ahmed,; Magdy, A.H. Zahran,; Fatma, S. Aggor,; S A Abd Elhady, S.S. Nada, Synthesis and Swelling Characterization of Carboxymethyl Cellulose-g-Poly (Acrylic acid-co-Acrylamide) Hydrogel and their Application in agricultural field, (n.d.).

[46] N. Satarkar, J.H.-A. biomaterialia, undefined 2008, Hydrogel nanocomposites as remote-controlled biomaterials, Elsevier. (n.d.). https://www.sciencedirect.com/science/article/pii/S1742706107001171?casa_token=f-gPBia6Q_QAAAAA:g-0nIkt4t9zQv5WNV0j1UMX7KuXZ4poBpmeIsUyWgkn71tUeC8xzn7EbBH4MyXCkpkyn93c8TOg (accessed August 9, 2022).

[47] J. Elliott, M. Macdonald, J. Nie, C.B.- Polymer, undefined 2004, Structure and swelling of poly (acrylic acid) hydrogels: effect of pH, ionic strength, and dilution on the crosslinked polymer structure, Elsevier. (n.d.). https://www.sciencedirect.com/science/article/pii/S0032386103011911?casa_token=NKKtF944vncAAAAA:QrCTi2NQWhPtMKn0k1KfmiA8nUluEpsMwL1_IXTIWouIaowVcJakZUCtqyS7lQqmG8OTpdR0yMM (accessed August 9, 2022).

[48] P. Liu, M. Zhai, J. Li, J. Peng, J.W.-R.P. and Chemistry, undefined 2002, Radiation preparation and swelling behavior of sodium carboxymethyl cellulose hydrogels, Elsevier. (n.d.). https://www.sciencedirect.com/science/article/pii/S0969806×01006491 (accessed August 9, 2022).

[49] A. Pourjavadi, P. Eftekhar Jahromi, F. Seidi, H. Salimi, Synthesis and swelling behavior of acrylatedstarch-g-poly (acrylic acid) and acrylatedstarch-g-poly (acrylamide) hydrogels, Carbohydr. Polym. 79 (2010) 933–940, https://doi.org/10.1016/j.carbpol.2009.10.021.

[50] H. Namazi, M. Hasani, M. Yadollahi, Antibacterial oxidized starch/ZnO nanocomposite hydrogel: Synthesis and evaluation of its swelling behaviours in various pHs and salt solutions, Int. J. Biol. Macromol. 126 (2019) 578–584, https://doi.org/10.1016/j.ijbiomac.2018.12.242.

[51] J.H. Trivedi, Synthesis, characterization, and swelling behavior of superabsorbent hydrogel from sodium salt of partially carboxymethylated tamarind kernel powder-g-PAN, J. Appl. Polym. Sci. 129 (2013) 1992–2003, https://doi.org/10.1002/APP.38910.

[52] E.S. Rufino, E.E.C. Monteiro, Characterisation of lithium and sodium salts of poly (methacrylic acid) by FTIR and thermal analyses, Polym. (Guildf.) 41 (2000) 4213–4222, https://doi.org/10.1016/S0032-3861(99)00630-8.

[53] H. Nie, M. Liu, F. Zhan, M. Guo, Factors on the preparation of carboxymethylcellulose hydrogel and its degradation behavior in soil, Carbohydr. Polym. 58 (2004) 185–189, https://doi.org/10.1016/J.CARBPOL.2004.06.035.

[54] Khushbu, S.G. Warkar, A. Kumar, Synthesis and assessment of carboxymethyl tamarind kernel gum based novel superabsorbent hydrogels for agricultural applications, Polym. (Guildf.) 182 (2019), 121823, https://doi.org/10.1016/j.polymer.2019.121823.

[55] P. Cen, L. Xia, Prod. Cell. Solid-State Ferment. (1999) 69–92, https://doi.org/10.1007/3-540-49194-5_4.

[56] E. Karadag Ae, D. Saraydin, Y.C.Ë. Aldiran, O.G. Èven, Swelling Studies of Copolymeric Acrylamide/Crotonic Acid Hydrogels as Carriers for Agricultural Uses, (n.d.). https://doi.org/10.1002/(SICI)1099–1581(200002)11:2.

[57] G.R. Blake, Particle density, Encycl. Earth Sci. Ser. (2008) 504–505, https://doi.org/10.1007/978-1-4020-3995-9_406.

[58] N.C. Womack, I. Piccoli, C. Camarotto, A. Squartini, G. Guerrini, S. Gross, M. Maggini, M.L. Cabrera, F. Morari, Hydrogel application for improving soil pore network in agroecosystems. Preliminary results on three different soils, CATENA 208 (2022), 105759, https://doi.org/10.1016/J.CATENA.2021.105759.

[59] F.F. Montesano, A. Parente, P. Santamaria, A. Sannino, F. Serio, Biodegradable superabsorbent hydrogel increaseswater retention properties of growing media and plant growth, Agric. Agric. Sci. Procedia 4 (2015) 451–458, https://doi.org/10.1016/J.AASPRO.2015.03.052.

[60] Q. de Jong van Lier, E.A.R. Pinheiro, L. Inforsato, Hydrostatic equilibrium between soil samples and pressure plates used in soil water retention determination: consequences of a questionable assumption, Rev. Bras. Ciência Do Solo 43 (2019), https://doi.org/10.1590/18069657RBCS20190014.

Constraints on the Parameterized Deceleration Parameter in FRW Universe

Amine Bouali,[1,*] Himanshu Chaudhary,[2,†] Ujjal Debnath,[3,‡] Tanusree Roy,[3,§] and G.Mustafa[4,¶]

[1] *Laboratory of Physics of Matter and Radiation, Mohammed I University, BP 717, Oujda, Morocco,*
[2] *Department of Applied Mathematics, Delhi Technological University, Delhi-110042, India,*
[3] *Department of Mathematics, Indian Institute of Engineering Science and Technology, Shibpur, Howrah-711 103, India,*
[4] *Department of Physics, Zhejiang Normal University, Jinhua 321004, People's Republic of China,*

Confirmation of accelerated expansion of the universe probed the concept of dark energy theory, and since then, numerous models have been introduced to explain its origin and nature. The present work is based on reconstructing dark energy by parametrization of the deceleration parameter in the FRW universe filled with radiation, dark matter and dark energy. We have chosen some well-motivated parametrized models 1-3 in an attempt to investigate the energy density in terms of deceleration parameters by estimating the cosmological parameters with the help of different observational datasets. Also, we have introduced a new model 4 for the parametrization of the deceleration parameter. Then we analyzed the cosmography parameters using the best-fit values of the parameters. Using the information criteria, we have examined the viability of the models.

arXiv:2301.12107v1 [gr-qc] 28 Jan 2023

I. INTRODUCTION

The concept of cosmic acceleration was probably one of the most promising discoveries in the modern cosmology paradigm. Recently, two independent research works involving distant supernovae suggested that in the present epoch, the universe is undergoing an accelerated expansion [1, 2]. This phenomenon has been favorably explained later by the existence of an energy component with massive negative pressure and comprising nearly 70% of the universe. This is known as "dark energy" (DE). The nature of this is still unidentified. Synchronizing with the observed data, many DE models have been proposed so far. Among them, the ΛCDM model is widely accepted as supposedly it 'best accommodates' the observations but also it comes with some disadvantages like a fine-tuning problem, coincidence problems and so on [3–5]. To overcome these drawbacks, alternative DE models have been explored like a quite favorable phantom, k-essence, Chaplygin gas, etc., [6] for possible explanations of the origin and nature of the dark energy. However, prior to the accelerated phase, the universe had gone through a decelerated phase in the early epoch where the effects of dark energy were absent or subdominant. It is believed that density perturbations occurred in this epoch which played a key role in the structural formation of the universe. So, to cover the entire evolution, we would have to employ a cosmological model which would simultaneously describe the accelerated and decelerated phases. Earlier, the definition of cosmology was addressed by Sandage[7] as a search for two simple but fundamental cosmographic parameters: Hubble parameter (H_0), which determined the expansion rate and a small correction q_0 due to gravity, known as deceleration parameter (DP), responsible for slowing down the expansion. Though the inclusion of 'dark energy' has completely changed the scenario, but still, any practical aspect of cosmological evolution is tightly bound to DP. It is defined as $q = -\frac{a\ddot{a}}{\dot{a}^2}$ where $a(t)$ is the scale factor of the universe. $q > 0$ ($\ddot{a} < 0$) indicates a decelerating universe and vice versa. While HP describes the linear part of the time dependence of the scale factor, the non-linear correction term q_0 opens up possibilities like the presence of local instabilities or the existence of chaotic regimes [8]. Moreover, the dynamics of observable galaxy number variation can be determined through DP. Like tenergyr cosmological models, DE has been subjected to numerous modifications to be better fitted with observational data.

Parametrization of DP as a function of scale factor a or redshift z can be accounted as a suitable approach to it. Limitations to such parametric assumptions are: (i) Most of the parametrizations diverge in the distant future, and some of them are only valid at low redshift limit ($z << 1$) [9, 10]; (ii) For prior parametric assumptions true nature of the dark energy can be misleading; (iii) In non-parametric models, evolution can be directly deduced from observational data avoiding parametric assumptions [11–15]. However, it can help to improve the efficiency of future cosmological surveys. So in pursuit of understanding the transition from decelerated to accelerated phase, the parametrization approach can be proved fruitful. Recently Capozziello reconstructed a divergence-free form of DP starting with Pade polynomials and analyzed the corresponding observational data [16]. A logarithmic parametrization of DP was proposed in ref. [17], and the constraints were obtained by using type Ia supernova, BAO and CMB data sets. Motivated by these ideas, we have adopted some well-motivated parametrizations of DP to reconstruct dark energy and, consequently, Hubble parameter $H(z)$ in terms of redshift z. Mainly, well-

* a1.bouali@ump.ac.ma
† himanshuch1729@gmail.com
‡ ujjaldebnath@gmail.com
§ tanusreeroy1995@gmail.com
¶ gmustafa3828@gmail.com

established parametrized models have been introduced for dark energy equation of state and constrain the model parameters by observational data analysis. In the study of the generalized holographic dark energy model, some well-known parametrization type models have been considered [18, 19]. Till now, some authors have assumed some possible forms of parametrization of deceleration parameter [20–25]. The main advantage of the parameterization deceleration parameter is that the final outcome may not depend on any particular gravitational theory. In the ref. [26, 27], the authors have introduced the analogous parametrized deceleration parameter instead of parametrized dark energy equation of state for the well-established models. Here we have considered the analogous parametrized deceleration parameter for the well-established models and constrains the model parameters by Markov Chain Monte Carlo (MCMC) data analysis. We compare all the models with the ΛCDM model (which is the base model) to examine which model is more viable than others.

The main focus of the work is to constrain the model parameters using recently released data. Here, in particular, we have chosen to use the updated astronomical datasets: the measurements of Hubble parameter from the differential evolution of cosmic chronometers (CC); Pantheon datasets from Type Ia Supernovae sample comprising 1048 measurements; 17 uncorrelated baryons acoustic oscillation (BAO) data. New constraints on DP have been provided by jointly analyzing the above datasets and implementing Monte Carlo Markov Chain (MCMC) method. The era of modern cosmology promotes the study of kinematic quantities, vastly known as "Cosmography" or "Cosmo-kinetics". The very idea of it is observationally driven and completely independent of any prior assumption of the gravity theory or elected cosmological model. Cosmography presents itself with a compelling advantage as it simply follows the symmetry principles and direct observation- without involving Einstein equations (Friedmann equations). Consequently, we can steadfastly avoid some arguable speculations regarding 'dark energy', 'dark matter' and others. While pure cosmography does not envision the scale factor $a(t)$ itself but the history of its evolution can be inferred to some extent. Dunajski and Gibbons [28] have studied the constraints on the cosmographic parameters like the deceleration, jerk, and snap parameters for different dark energy models. The parametrization of these quantities is discussed in [29–32]. Shafieloo, Kim and Linder [22] have discussed the non-parametric reconstruction of these quantities.

The organization of the paper is as follows: In section II, we consider the basic equations of the FRW model. The Hubble parameter is written in terms of the deceleration parameter. Section III deals with the data descriptions like cosmic chronometric datasets, Pantheonic datasets and BAO datasets. In section IV, we consider parametrized deceleration parameter models like models 1, 2, 3 and 4. In section V, we fit the models with $H(z)$ and SNe-IA datasets. In section VI, we discuss the cosmography parameters. In section VII, we analyze the detailed description of the model parameters. In section VIII, we present the information criteria for our models. Finally, the results are presented in section IX.

II. BASIC EQUATIONS OF FRW MODEL

We have considered a spatially flat, homogeneous, isotropic FRW universe with line element

$$ds^2 = -dt^2 + a^2(t)\left[dr^2 + r^2\left(d\theta^2 + sin^2\theta d\phi^2\right)\right] \quad (1)$$

$a(t)$ being the scale factor.
The energy-momentum tensor of the fluid reads as

$$T_{\mu\nu} = (\rho + p)u_\mu u_\nu + p g_{\mu\nu} \quad (2)$$

where ρ and p are the energy density and pressure density of the fluid respectively. The fluid 4-velocity $u^\mu = \frac{dx^\mu}{ds}$ satisfies the relation $u^\mu u_\mu = -1$.

For the FRW Univere, the Friedmann equations in Einstein's gravity are given by

$$H^2 = \frac{8\pi G}{3}\,\rho \quad (3)$$

and

$$\dot{H} = -4\pi G(\rho + p) \quad (4)$$

where, $H = \dot{a}/a$ is the Hubble parameter and overhead dot denotes derivative with cosmic time t. Considering that the universe is filled with fluid matter of total energy density ρ and total pressure p, it obeys the energy conservation equation

$$\dot{\rho} + 3H(\rho + p) = 0 \quad (5)$$

We start with the prediction that the universe is composed of matter content comprising radiation, dark matter (DM) and dark energy (DE). So, ρ and p consist of densities and pressures of radiation, DM and DE. So $\rho = \rho_r + \rho_m + \rho_d$ and $p = p_r + p_m + p_d$. Now assume that radiation, DM and DE follows the conservation equation separately so that

$$\dot{\rho}_r + 3H(\rho_r + p_r) = 0, \quad (6)$$

$$\dot{\rho}_m + 3H(\rho_m + p_m) = 0 \quad (7)$$

and

$$\dot{\rho}_d + 3H(\rho_d + p_d) = 0 \tag{8}$$

For radiation, $p_r = \frac{1}{3}\rho_r$, so from equation (6) we obtain $\rho_r = \rho_{r0}a^{-4}$. Since the DM follows negligible pressure (i.e., $p_m = 0$), so from equation (7) we obtain $\rho_m = \rho_{m0}a^{-3}$.

Let us consider the deceleration parameter

$$q = -1 - \frac{\dot{H}}{H^2} \tag{9}$$

The corresponding deceleration parameter for DE is given by [**?**]

$$q_d = -1 - \frac{\dot{H}_d}{H_d^2} \tag{10}$$

where H_d is the Hubble expansion rate of the dark energy term. So from equations (3) and (4), we can write

$$H_d^2 = \frac{8\pi G}{3}\ \rho_d \tag{11}$$

and

$$\dot{H}_d = -4\pi G(\rho_d + p_d) \tag{12}$$

Using the field equations (11) and (12) and the energy-conservation equation (8), the fluid energy density becomes

$$\rho_d = \rho_{d0}\ e^{\int \frac{2(1+q_d)}{1+z}dz} \tag{13}$$

where ρ_{d0} represents the present value of the density parameter and z is the redshift parameter described as $1 + z = \frac{1}{a}$ (presently, $a_0 = 1$). Defining $\Omega_{r0} = \frac{8\pi G\rho_{r0}}{3H_0^2}$, $\Omega_{m0} = \frac{8\pi G\rho_{m0}}{3H_0^2}$ and $\Omega_{d0} = \frac{8\pi G\rho_{d0}}{3H_0^2}$, from equation (3), we obtain the Hubble parameter as

$$H^2(z) = H_0^2\left[\Omega_{r0}(1+z)^4 + \Omega_{m0}(1+z)^3 + \Omega_{d0}\ e^{\int \frac{2(1+q_d)}{1+z}dz}\right] \tag{14}$$

where $\Omega_{d0} = 1 - \Omega_{r0} - \Omega_{m0}$.

III. DATA DESCRIPTION

In this section, we will constrain our model parameters by using three types of dataset. The CC datasets consist 31 measurements, Pantheon dataset consists 1048 measurements and 17 uncorrelated BAO measurements to obtain the best fit value of our model parameters. We have implemented the Markov Chain Monte Carlo (MCMC) [33] and implemented with the open source package Polychord [34] and GetDist [35]. The total χ^2 function of the combination CC + BAO + Pantheon and define as

$$\chi^2 = \chi^2_{CC} + \chi^2_{SN} + \chi^2_{BAO}.$$

A. Cosmic Chronometric (CC) datasets

We consider the compilation of 31 measurements of CC lying between the redshift range $0.07 \leq z \leq 1.965$. The underlying principle for these measurements was proposed in [36], by relating the Hubble parameter with redshift z, and cosmic time t

$$H(z) = -\frac{1}{1+z}\frac{\mathrm{d}z}{\mathrm{d}t}$$

The χ^2 function for these measurements, denoted by $\chi^2_{\rm CC}$, is

$$\chi^2_{\rm CC} = \sum_{i=1}^{31}\frac{\left[H^{th}(z_i) - H^{obs}(z_i)\right]^2}{\sigma^2_{H^{\rm obs}(z_i)}}, \tag{15}$$

where $H^{\rm th}(z_i, k, \alpha, h)$ represent the theoretical value obtained from our cosmological model, $H^{\rm obs}(z_i)$ and represent the observed value of hubble parameter with standard deviation $\sigma^2_{H^{\rm obs}(z_i)}$. (to see more and rundown all measurements see [37])

B. Pantheon datasets

The Pantheon datasets comprises 1048 measurements of type Ia supernovae from five different sub-samples SNLS, SDSS, PSI, low- z, and HST in the redshift range of $0.01 < z < 2.3$ [38] . The χ^2 function of the Pantheon data is given as

$$\chi^2_{\rm Pan} = \Delta\mu C^{-1}_{\rm Pan}\Delta\mu^T, \tag{16}$$

where $\Delta\mu = \mu^{\rm obs}_i - \mu^{\rm th}$. Where $\left(\mu^{\rm obs}_i\right)$ represented as observed distance modulus and evaluated as

$$\mu^{\rm obs}_i = \mu_{B,i} + \mathcal{M}, \tag{17}$$

$\mu_{B,i}$ represents the observed peak magnitude at maximum in the rest frame of the B band for redshift z_i, while $\mathcal{M}$ represents nuisance parameter. The theoretical distance modulus evaluated as

$$\mu^{\rm th} = 5\log_{10} D_L + \mathcal{M}, \tag{18}$$

where

$$D_L = (1 + z_{\rm hel})\int_0^{z_{\rm cmb}}\frac{H_0 dz}{H(z)}, \tag{19}$$

with $z_{\rm hel}$ is heliocentric and $z_{\rm cmb}$ is CMB rest frame redshifts. The covariance matrix is measured as $C_{\rm Pan}$ =

$C_{\rm sys}$ $+ D_{\rm stat}$. Where $C_{\rm sys}$ is systematic covariance matrix and $D_{\rm stat}$ stands for diagonal of covariance matrix of the statistical uncertainty and calculated as

$$D_{\rm stat\ ,ii} = \sigma^2_{\mu_{B,i}}. \tag{20}$$

The description and the systematic covariance matrix together with $\mu_{B,i}, \sigma^2_{\mu_{B,i}}, z_{\rm cmb}$, and $z_{\rm hel}$ for the i th SnIa are mentioned in [39].

C. Uncorrelated Baryon Acoustic Oscillations (unCor BAO)

We picked 17 uncorrelated BAO measurements from the greatest collection of BAO dataset of (333) measurements since adopting the entire catalogue of BAO might lead in a very significant error owing to data correlations, therefore we opted a small dataset to minimise inaccuracies. Transverse BAO studies contribute measurements of $D_H(z)/r_d = c/H(z)r_d$ with comoving angular diameter distance.[40] [41].

$$D_M = \frac{c}{H_0} S_k\left(\int_0^z \frac{dz'}{E\left(z'\right)}\right), \tag{21}$$

where

$$S_k(x) = \begin{cases} \frac{1}{\sqrt{\Omega_k}} \sinh\left(\sqrt{\Omega_k} x\right) & \text{if} \quad \Omega_k > 0 \\ x & \text{if} \quad \Omega_k = 0 \\ \frac{1}{\sqrt{-\Omega_k}} \sin\left(\sqrt{-\Omega_k} x\right) & \text{if} \quad \Omega_k < 0. \end{cases} \tag{22}$$

We also consider the angular diameter distance $D_A = D_M/(1+z)$ and the $D_V(z)/r_d$. which is combination of the BAO peak coordinates and r_d is the sound horizon at the drag epoch. Finally we can obtain "line-of-sight" (or "radial") observations directly the Hubble parameter

$$D_V(z) \equiv \left[z D_H(z) D_M^2(z)\right]^{1/3}. \tag{23}$$

IV. PARAMETERIZED DECELERATION PARAMETER

Most simplest parametrization of q which contains two parameters can be taken as

$$q(z) = q_0 + q_1 \mathcal{X}(z) \tag{24}$$

where q_0 and q_1 are constants and $\mathcal{X}(z)$ is a function of redshift z. In search of satisfactory solutions to the cosmological puzzles, many forms of $\mathcal{X}(z)$ has been suggested. As mentioned earlier, most of them were inadequate in explaining future evolution scenarios. So the persuasion of an ideal divergence-free parametrization of DP is still relevant. The well-known parametrized equation of state parameter models has been introduced by several authors, and the corresponding analogous of these models for parametrized deceleration parameters have been introduced in [26, 27]. Here, we have adopted the analogous of some well-known parametrized models for the deceleration parameter, which contains two unknown parameters and calculated the corresponding Hubble parameter in terms of redshift z.

A. Model 1 (Wetterich type)

The Wetterich model for parametrized equation of state parameter has been studied in [42, 43]. The analogous Wetterich type parametrization of deceleration parameter has been introduced in [26, 27] and is given by

$$q_d(z) = \frac{q_0}{1 + q_1 1og(1+z)} \tag{25}$$

where q_0 and q_1 are constants. Then the energy density will be given by

$$\rho_d = \rho_{d0}\ (1+z)^2\ \left\{1 + q_1\ log(1+z)\right\}^{\frac{2q_0}{q_1}} \tag{26}$$

From equation (14), we obtain

$$H^2(z) = H_0^2 \left[\Omega_{r0}(1+z)^4 + \Omega_{m0}(1+z)^3 + (1 - \Omega_{r0} - \Omega_{m0})\,(1+z)^2 \left\{1 + q_1 \log(1+z)\right\}^{\frac{2q_0}{q_1}}\right] \tag{27}$$

B. Model 2 (Barboza-Alcaniz type)

The Barboza-Alcaniz model for parametrized equation of state parameter has been studied in [44]. The analogous Barboza-Alcaniz type parametrization of deceleration parameter has been introduced in [26, 27] and is given by

$$q_d(z) = q_0 + q_1 \frac{z(1+z)}{1+z^2} \tag{28}$$

where q_0 and q_1 are constants. Then the energy density will be

$$\rho_d = \rho_{d0}\ (1+z)^{2(1+q_0)}\ \left(1+z^2\right)^{q_1} \tag{29}$$

From equation (14), we obtain

$$H^2(z) = H_0^2 \left[\Omega_{r0}(1+z)^4 + \Omega_{m0}(1+z)^3 + (1 - \Omega_{r0} - \Omega_{m0})\,(1+z)^{2(1+q_0)}\ \left(1+z^2\right)^{q_1}\right] \tag{30}$$

C. Model 3 (CPL type)

The famous Chevallier-Polarski-Linder (CPL) model for parametrized equation of state parameter has been studied in [29, 30]. The analogous CPL type parametrization of deceleration parameter has been introduced in [26, 27] and is given by

$$q_d(z) = q_0 + q_1 \frac{z}{1+z} \tag{31}$$

where q_0 and q_1 are constants. Subsequently, the energy density (13) becomes

$$\rho_d = \rho_{d0}\ (1+z)^{2(1+q_0+q_1)}\ e^{\frac{2q_1}{1+z}} \tag{32}$$

From equation (14), we obtain

$$H^2(z) = H_0^2 \left[\Omega_{r0}(1+z)^4 + \Omega_{m0}(1+z)^3 + (1-\Omega_{r0}-\Omega_{m0})\,(1+z)^{2(1+q_0+q_1)}\ e^{\frac{2q_1}{1+z}}\right] \tag{33}$$

D. Model 4

Here we propose a new parametrized model for deceleration parameter and is given as in the form:

$$q_d(z) = q_0 + q_1 \frac{1+z}{2+z} \tag{34}$$

where q_0 and q_1 are constants. Subsequently, the energy density (13) becomes

$$\rho_d = \rho_{d0}\ (1+z)^{2(1+q_0)}(2+z)^{2q_1} \tag{35}$$

From equation (14), we obtain

$$H^2(z) = H_0^2 \left[\Omega_{r0}(1+z)^4 + \Omega_{m0}(1+z)^3 + (1-\Omega_{r0}-\Omega_{m0})\,(1+z)^{2(1+q_0)}(2+z)^{2q_1}\right] \tag{36}$$

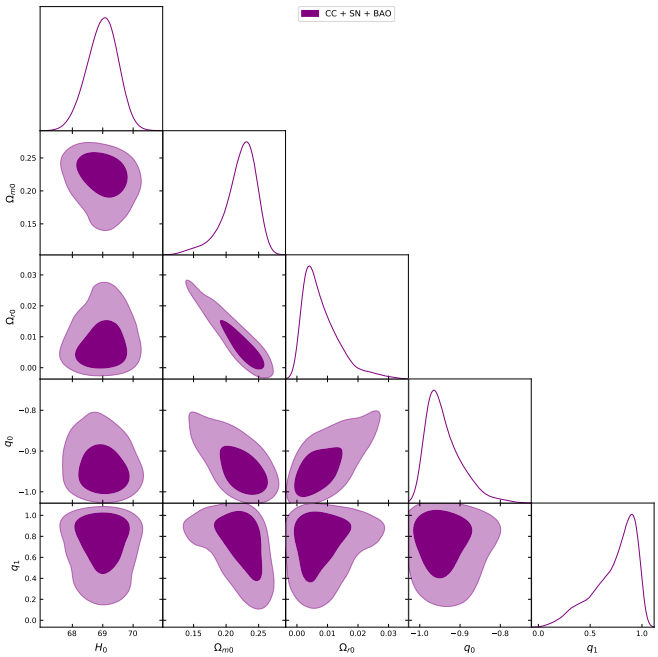


FIG. 1. The above figure shows the MCMC confidence contours at 1σ and 2σ obtained from CC+SNIa+BAO dataset for Model 1 (Wetterich type).

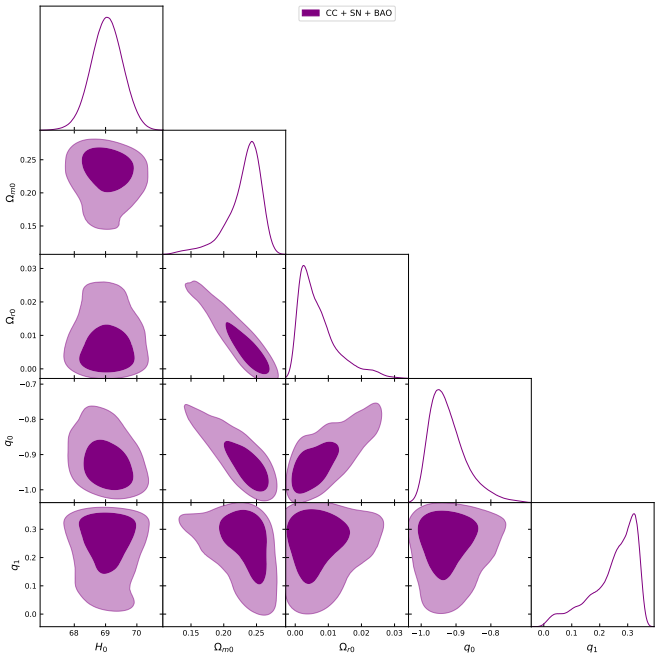


FIG. 2. The above figure shows the MCMC confidence contours at 1σ and 2σ obtained from CC+SNIa+BAO dataset for Model 2 (Barboza-Alcaniz type).

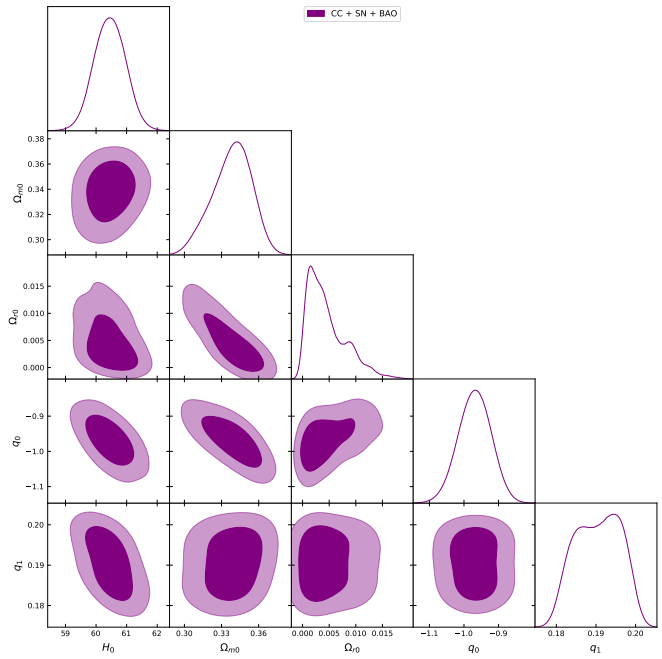


FIG. 3. The above figure shows the MCMC confidence contours at 1σ and 2σ obtained from CC+SNIa+BAO dataset for Model 3 (CPL type).

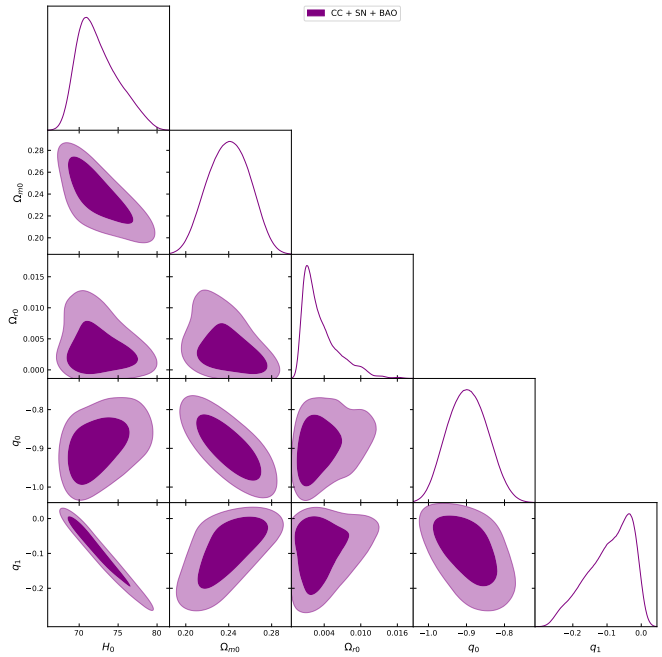


FIG. 4. The above figure shows the MCMC confidence contours at 1σ and 2σ obtained from CC+SNIa+BAO dataset for Model 4.

MCMC Results		
Model	Parameters	Best fit Value
ΛCDM Model	H_0	$69.854848^{+1.259100}_{-1.259100}$
	$\Omega_{\rm m0}$	$0.268654^{+0.012822}_{-0.012822}$
	Ω_Λ	$0.724585^{+0.009373}_{-0.009373}$
Model 1	H_0	$68.990853^{+0.502428}_{-0.502428}$
	$\Omega_{\rm m0}$	$0.221900^{+0.022982}_{-0.022982}$
	$\Omega_{\rm r0}$	$0.007835^{+0.005651}_{-0.005651}$
	q_0	$-0.941550^{+0.043352}_{-0.043352}$
	q_1	$0.740557^{+0.220903}_{-0.220903}$
Model 2	H_0	$69.061167^{+0.498457}_{-0.498457}$
	$\Omega_{\rm m0}$	$0.231719^{+0.025169}_{-0.025169}$
	$\Omega_{\rm r0}$	$0.006670^{+0.005154}_{-0.005154}$
	q_0	$-0.924723^{+0.052163}_{-0.052163}$
	q_1	$0.251461^{+0.084597}_{-0.084597}$
Model 3	H_0	$60.450021^{+0.511464}_{-0.511464}$
	$\Omega_{\rm m0}$	$0.338454^{+0.015533}_{-0.015533}$
	$\Omega_{\rm r0}$	$0.004485^{+0.003527}_{-0.003527}$
	q_0	$-0.968857^{+0.046034}_{-0.046034}$
	q_1	$0.190554^{+0.006576}_{-0.006576}$
Model 4	H_0	$71.392060^{+2.608372}_{-2.608372}$
	$\Omega_{\rm m0}$	$0.240129^{+0.021352}_{-0.021352}$
	$\Omega_{\rm r0}$	$0.003402^{+0.002740}_{-0.002740}$
	q_0	$-0.897355^{+0.056271}_{-0.056271}$
	q_1	$-0.091362^{+0.073319}_{-0.073319}$

TABLE I. Summary of the MCMC results using CC + Pan + BAO dataset.

V. OBSERVATIONAL, AND THEORETICAL COMPARISONS OF THE HUBBLE AND DISTANCE MODULUS FUNCTIONS

Once we have the free parameters of the Model (1-4), we can compare the model predictions to the observational data and the *Lambda*CDM model with error bands, respectively.

A. Comparison with the Hubble data points.

First, We consider the comparison of the Models (1 - 4) with the 31 (CC) data points and the ΛCDM model with 1 σ and 2 σ error bands . The comparison findings are shown in Figure 5, 6, 7 and 8. The Figure shows that all model fit with (CC) dataset quite well.

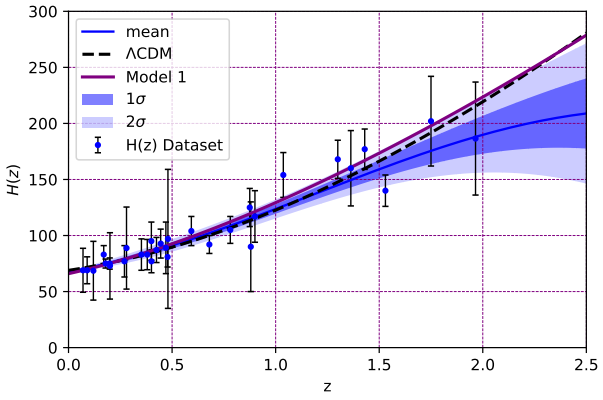


FIG. 5. The figure show that the theoretical curve of Hubble function $H(z)$ of Model 1 and ΛCDM models against the CC datasets.

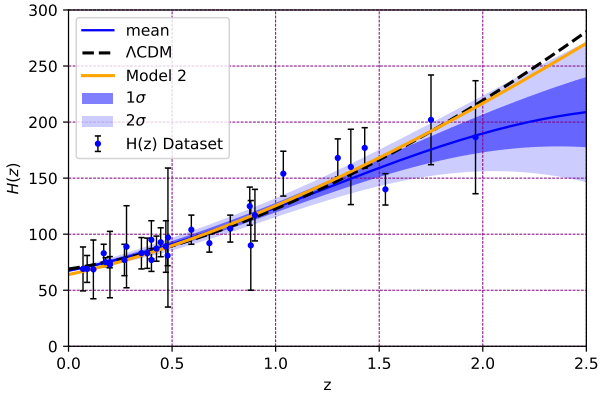


FIG. 6. The figure show that the theoretical curve of Hubble function $H(z)$ of Model 2 and ΛCDM models against the CC datasets.

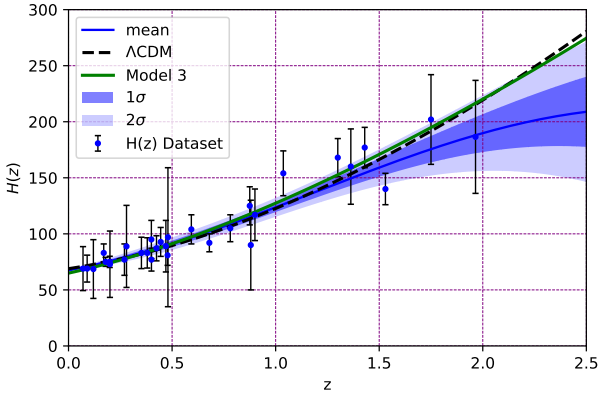


FIG. 7. The figure show that the theoretical curve of Hubble function $H(z)$ of Model 2 and ΛCDM models against the CC datasets.

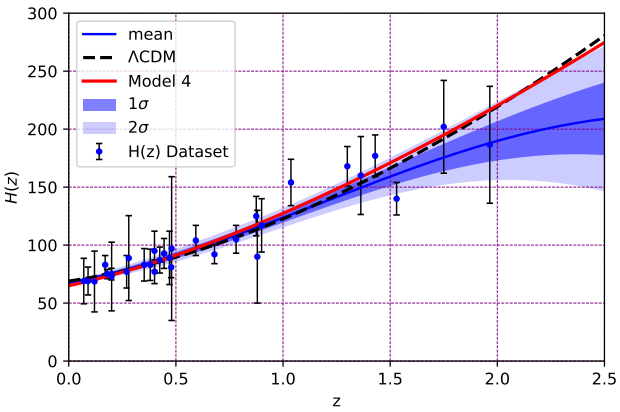


FIG. 8. The figure show that the theoretical curve of Hubble function $H(z)$ of Model 4 and ΛCDM models against the CC datasets.

1. *Comparison with the Pantheon data.*

We now compare the $\mu(z)$ distance modulus function of Models (1-4) with the Pantheon data. From Fig. 9,10,11 and 12 one can see that all Models fit with the Pantheon dataset, very well.

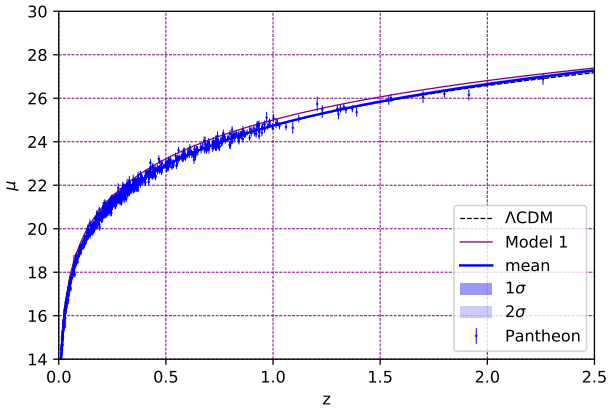


FIG. 9. The figure show that the theoretical curve of distance modulus $\mu(z)$ of Model 1 and ΛCDM models against the Pantheon data.

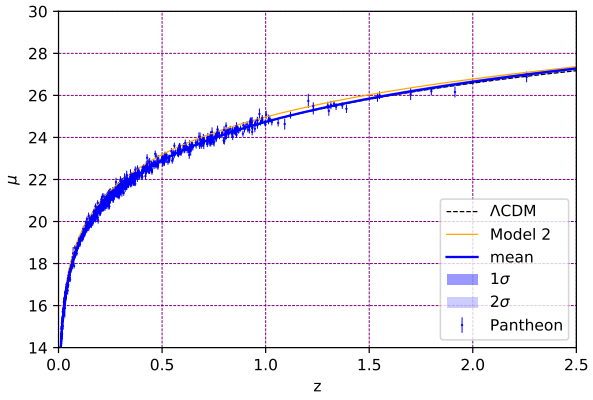


FIG. 10. The figure show that the theoretical curve of distance modulus $\mu(z)$ of Model 2 and ΛCDM models against the Pantheon data.

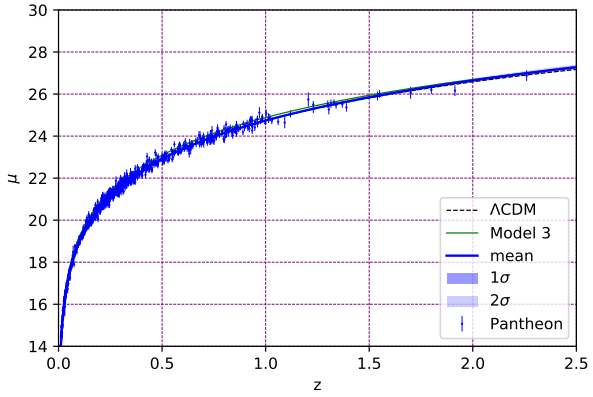


FIG. 11. The figure show that the theoretical curve of distance modulus $\mu(z)$ of Model 3 and ΛCDM models against the Pantheon data.

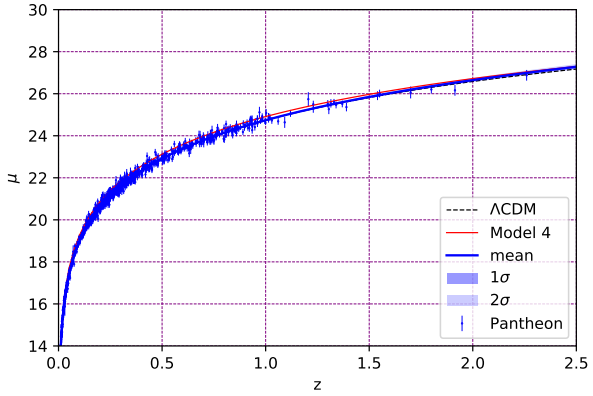


FIG. 12. The figure show that the theoretical curve of distance modulus $\mu(z)$ of Model 4 and ΛCDM models against the Pantheon data.

VI. COSMOGRAPHY PARAMETERS

To study the early evolution and late evolution of the universe, some other parameters named cosmographical parameters can be analyzed. The cosmographical parameter like jerk (j), snap (s) parameters are [45–48]

$$j = \frac{\dddot{a}}{aH^3} = (1+z)\frac{dq}{dz} + q(1+2q), \tag{37}$$

$$s = \frac{a^{(4)}}{aH^4} = -(1+z)\frac{dj}{dz} - j(2+3q) \tag{38}$$

$$\tag{39}$$

So the cosmographical parameters contains the higher-order derivatives of the deceleration parameter q. The 'jerk' parameter is considered to have a very useful feature that is for standard ΛCDM model, j always takes the value unity which helps us assess the deviation regarding different dark energy models. Sahni et al. and Alam et al. analyzed the importance of the jerk parameter j for discriminating various dark energy models. We have explored the evolution of such kinematical quantities with respect to the redshift for the involved parametric models.

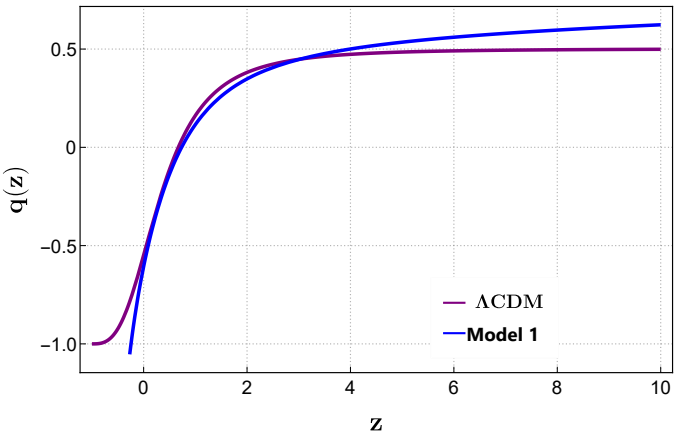


FIG. 13. Evolution of deceleration parameter of Model 1 with respect to redshift.

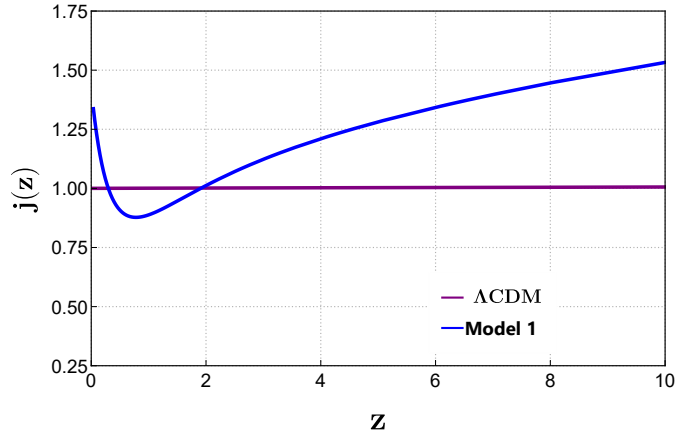


FIG. 14. Evolution of jerk parameter of Model 2 with respect to redshift.

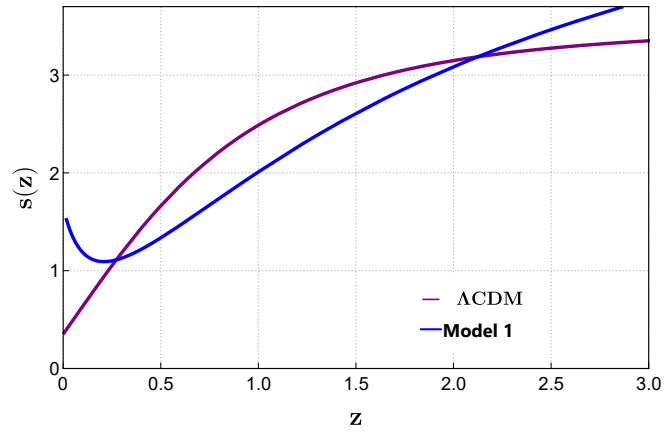


FIG. 15. Evolution of snap parameter with of Model 3 with respect to redshift.

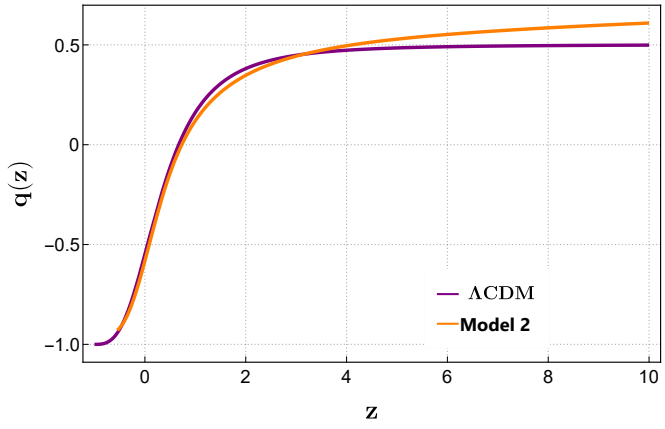


FIG. 16. Evolution of deceleration parameter of Model 1 with respect to redshift.

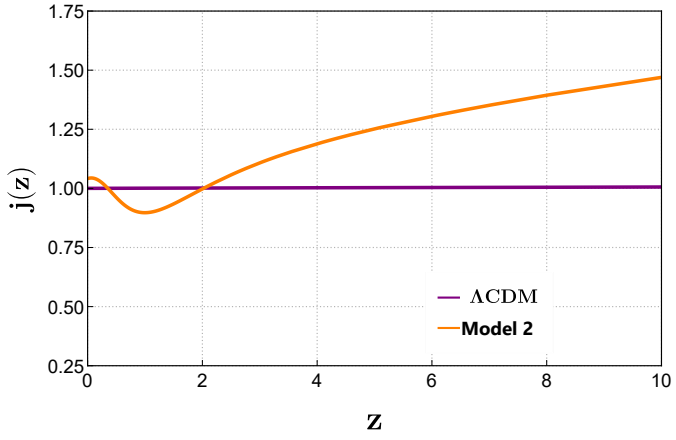


FIG. 17. Evolution of jerk parameter of Model 2 with respect to redshift.

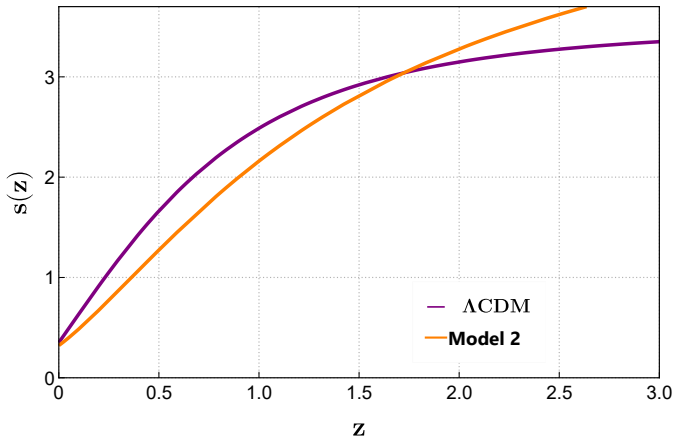


FIG. 18. Evolution of snap parameter with of Model 2 with respect to redshift.

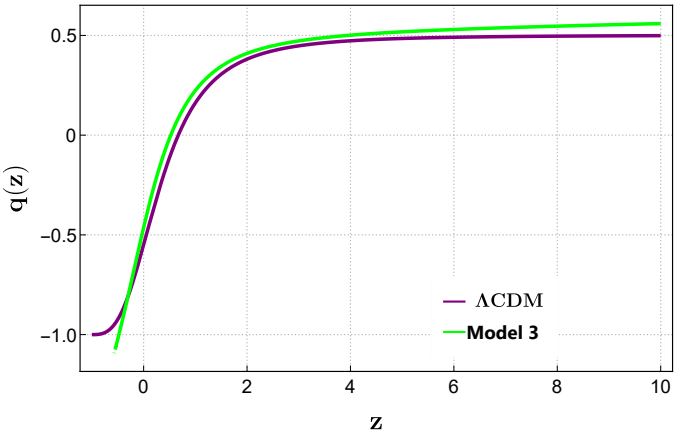


FIG. 19. Evolution of decceleration parameter with of Model 3 with respect to redshift.

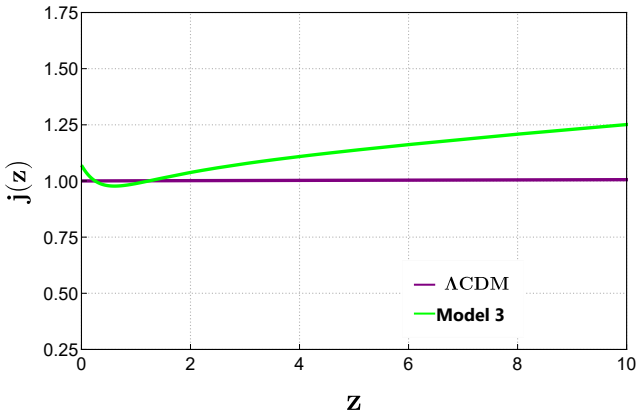


FIG. 20. Evolution of jerk parameter of Model 2 with respect to redshift.

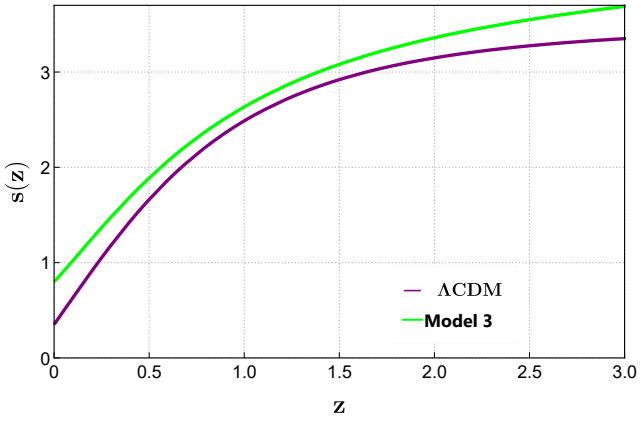


FIG. 21. Evolution of snap parameter with of Model 3 with respect to redshift.

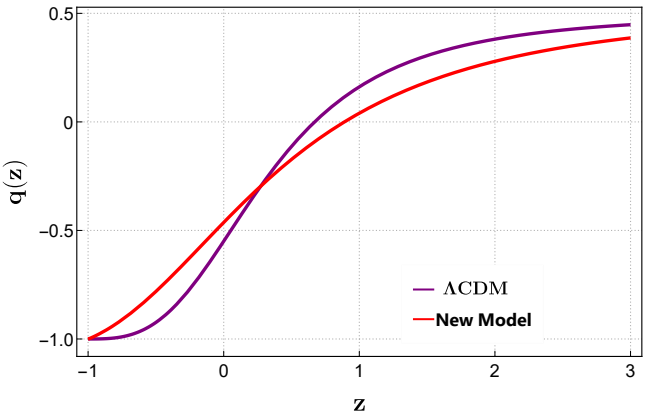


FIG. 22. Evolution of decceleration parameter with of Model 4 with respect to redshift.

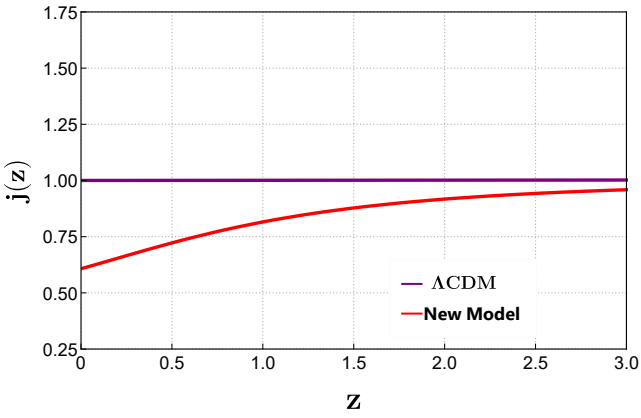


FIG. 23. Evolution of jerk parameter with of Model 4 with respect to redshift.

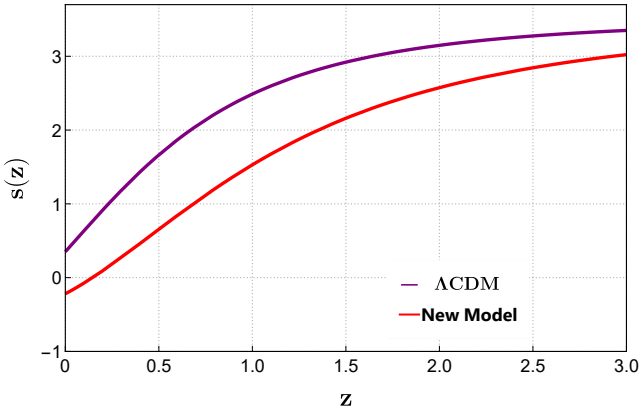


FIG. 24. Evolution of snap parameter with of Model 4 respect to redshift.

VII. DETAILED DESCRIPTION

A. Cosmological Parameters

1. Hubble parameter (H_0)

This Hubble parameter is the normalized rate of expansion, although this parameter is not constant. The current estimates of this parameter are "65-75". Model 3 has the "$H_0 < 65$ ", but on the other hand, Model 1 and Model 2 have this parameter in the Hubble range of "$70 < H_0 < 75$", and finally, in Model 4, "$H_0 > 70$ "(Table: I).

2. The matter density parameter ($\Omega_{\rm m0}$)

This parameter is the ratio of the actual (or observed) density ρ_m to the critical density ρ_c. The Model 1, Model 2 and Model 4 have a matter density in the range of "$0.22 < \Omega_{\rm m0} < 0.26$", but in Model 3, this parameter has the range "$\Omega_{\rm m0}$ ¡ 30" (Table: I).

3. The radiation density parameter ($\Omega_{\rm r0}$)

This parameter is the ratio of the actual (or observed) radiation density ρ_{rad} to the critical density ρ_c. This parameter for Model 1, Model 2, Model 3 and Model 4 lies in the range "$0.003 < \Omega_{\rm r0} < 0.008$" (Table: I).

B. Cosmographic Analysis

The cosmographic analysis provides a universal and effective way to compare the solutions of the theoretical models with cosmological observations. From the observational data, we obtain a set of cosmological parameters, which must be compared with the predicted values of the same parameters obtained from a given model. The result of the comparison allows us to conclude the acceptability of the considered model. Thus, for a complete comparison of all models with the observations and the ΛCDM model, we will consider an extended set of parameters constructed from the higher-order time derivatives of the scale factor. More exactly, we will concentrate on the comparative behaviour of the deceleration, jerk, and snap parameters of all models and ΛCDM models.

1. Deceleration parameter

While analyzing Model 1's trajectory, the behavior of the model's deceleration parameter is nearly comparable to the ΛCDM model in the redshift range of q ∈ [-0,2], but Model 1 endures a super acceleration in the future (Fig: 13). Model 2 appears to have the same behavior as ΛCDM in the q ∈ [-0.8,4], but Model 2 is slower since it achieves the value -0.83565 as $z \to 0$. (Fig: 16). In Model 3, this parameter appears to behave similarly to the ΛCDM model, in q ∈ [-0.5,6], before experiencing a super acceleration in the near future as "q = -1.25464" (Fig: 16). Model 4 behaves differently from the ΛCDM. but it aquire the same value as ΛCDM in near future 22).

2. The jerk parameter

The jerk parameter of Model 1 basically different from ΛCDM at high as well as low redshift. However, important Model 1 predicts the higher value of $j = 1.357595$ at $z = 0$ (Fig: 15), Meanwhile, on the other hand, Model 2 also shows different behaviour at both high and low redshift, and seems to coincide with ΛCDM at redshift value of $z = 2$ and $z = 0.342566$, and this model also predicts the higher value of $j = 1.134545$ (Fig: 18). Model 3 shows different behavior than the ΛCDM as at high redshift $z > 2$, it having a higher value than the ΛCDM of $j = 1.2657$, but it cuts the trajectory of ΛCDM, twice at $z = 1.5$ and $z = 0.3$, finally, at lower redshift, it attains the j value of 1.05417, which is a higher than ΛCDM (Fig: 21). Although the jerk of Model 4 is inconsistent with ΛCDM, since it shows different evolution at both high and low redshifts and predicts the lower value of $j = 0.608251$ (Fig: 24).

3. The snap parameter

This parameter of Model 1 and Model 2 is significantly systematic the difference with ΛCDM within the whole redshift range, but this parameter of Model 1 monotonically decreases in the redshift range of $z > 0.3$ and acquires a sudden increase in "s" value and predicts the higher value of $j = 1.608251$ then ΛCDM, but Model 2 predicts the same value of j as ΛCDM,(Fig: 15, 18). Model 3 trajectory shows a proper systematic difference with ΛCDM and predicts a higher value of "s" as 0.734546 21 Finally, the snap trajectory of Model 4 is notably non-identical with the ΛCDM, in the given redshift range and accommodate the "s" value of -0.222743 as $z \to 0$. which is lower than ΛCDM (Fig: 24).

VIII. INFORMATION CRITERIA

To discuss the viable model analysis, we need to know the study of information criteria (IC). The Akaike Information Criteria (AIC) [49] is merely used among all ICs. The AIC is an asymptotically unbiased estimator of Kullback-Leibler information as the AIC is an approximate minimization of the Kullback-Leibler information. The Gaussian estimator for the AIC can be written as [50–53] $\mathrm{AIC} = -2\ln(\mathcal{L}_{max}) + 2\kappa + \frac{2\kappa(\kappa+1)}{N-\kappa-1}$ where $\mathcal{L}_{max}$ is the maximum likelihood function, κ is

the number of parameters of the models, and N is the number of data points used in the data fit of the models. Since for the models, $N \gg 1$, so for this assumption, the above expression converts to the original AIC like $\mathrm{AIC} = -2\ln(\mathcal{L}_{max}) + 2\kappa$. If the set of models are given, the deviations of the IC values are reduced to $\triangle\mathrm{AIC} = \mathrm{AIC}_{model} - \mathrm{AIC}_{min} = \triangle\chi^2_{min} + 2\triangle\kappa$ In the study of data analysis, the more favorable range of $\triangle$AIC is $(0, 2)$. The low favorable range of $\triangle$AIC is $(4, 7)$, while $\triangle\mathrm{AIC} > 10$ provides less support model.

Model	χ^2_{min}	χ^2_{red}	AIC	ΔAIC
ΛCDM Model	1102.67	0.981	1106.67	0
Model 1	1103.21	0.961	1109.69	0.54
Model 2	1103.05	0.963	1107.05	0.38
Model 3	1103.85	0.965	1107.85	1.18
Model 4	1103.76	0.972	1109.76	3.09

TABLE II. Summary of the χ^2_{min}, χ^2_{red}, AIC and ΔAIC.

IX. DISCUSSIONS AND CONCLUSIONS

We have assumed the FRW model of the universe in the presence of radiation, dark matter and dark energy. Instead of considering the well-known parametrized dark energy equation of state, we have considered the analogous form of parametrized deceleration parameter for the dark energy component and found the Hubble parameter in terms of redshift with other model parameters. Here we have assumed Model 1 (Wetterich type), Model 2 (Barboza-Alcaniz type) and Model 3 (CPL type), which contains two unknown parameters. Also, we have introduced a new Model 4 for parametrized deceleration parameters, which also contains two unknown parameters. The model parameters have been constrained for $H(z)$ datasets, Pantheon datasets, and BAO datasets by MCMC method. Using the best-fit parameters, we have shown the nature of the deceleration parameter, jerk parameter and snap parameter. The viability of the models has been studied by the information criteria. We have compared all the models as well as compared with the ΛCDM model (which is the base model) to get which model is more viable than others. From Table: II, we observe that Models 1 - 4 are all viable models, but (i) Model 3 is more viable than Model 4, (ii) Model 1 is more viable than Model 3 and (iii) Model 2 is more viable than the Model 1 compared to the ΛCDM model.

Acknowledgement: TR is thankful to IIEST, Shibpur, India for providing Institute fellowship (SRF).

[1] P. Astier, R. Pain, Observational evidence of the accelerated expansion of the universe, Comptes Rendus Physique 13 (6-7) (2012) 521–538.
[2] S. Perlmutter, G. Aldering, G. Goldhaber, R. Knop, P. Nugent, P. Castro, S. Deustua, S. Fabbro, A. Goobar, D. Groom, et al., Measurements of Ω and λ from 42 high-redshift supernovae, Astrophys. J 517 (5) (1999).
[3] S. Weinberg, The cosmological constant problem, Reviews of modern physics 61 (1) (1989) 1.
[4] P. J. Steinhardt, L. Wang, I. Zlatev, Cosmological tracking solutions, Physical Review D 59 (12) (1999) 123504.
[5] E. Copeland, M. sami, and s. tsujikawa, Int. J. Mod. Phys. D 15 (2006) 1753.
[6] L. Amendola, S. Tsujikawa, Dark energy: theory and observations, Cambridge University Press, 2010.
[7] E. R. Harrison, Why the sky is dark at night, Physics Today 27 (2) (1974) 30–33.
[8] Y. L. Bolotin, V. Cherkaskiy, O. Lemets, D. Yerokhin, L. Zazunov, Cosmology in terms of the deceleration parameter. part i, arXiv preprint arXiv:1502.00811 (2015).
[9] Y. Gong, A. Wang, Reconstruction of the deceleration parameter and the equation of state of dark energy, Physical Review D 75 (4) (2007) 043520.
[10] J. Cunha, J. Lima, Transition redshift: new kinematic constraints from supernovae, Monthly Notices of the Royal Astronomical Society 390 (1) (2008) 210–217.
[11] V. Sahni, A. Shafieloo, A. A. Starobinsky, Two new diagnostics of dark energy, Physical Review D 78 (10) (2008) 103502.
[12] T. Holsclaw, U. Alam, B. Sansó, H. Lee, K. Heitmann, S. Habib, D. Higdon, Nonparametric reconstruction of the dark energy equation of state from diverse data sets, Physical Review D 84 (8) (2011) 083501.
[13] L. Xu, J. Lu, Cosmic constraints on deceleration parameter with sne ia and cmb, Modern Physics Letters A 24 (05) (2009) 369–376.
[14] R. Nair, S. Jhingan, D. Jain, Cosmokinetics: a joint analysis of standard candles, rulers and cosmic clocks, Journal of Cosmology and Astroparticle Physics 2012 (01) (2012) 018.
[15] B. Santos, J. C. Carvalho, J. S. Alcaniz, Current constraints on the epoch of cosmic acceleration, Astroparticle Physics 35 (1) (2011) 17–20.
[16] S. Capozziello, R. D'Agostino, O. Luongo, Thermodynamic parametrization of dark energy, Physics of the Dark Universe (2022) 101045.
[17] A. A. Mamon, S. Das, A parametric reconstruction of the deceleration parameter, The European Physical Journal C 77 (7) (2017) 1–9.
[18] Z. Zhang, M. Li, X.-D. Li, S. Wang, W.-S. Zhang, Generalized holographic dark energy and its observational constraints, Modern Physics Letters A 27 (20) (2012) 1250115.
[19] U. Debnath, Accretion of some classes of holographic de onto higher-dimensional schwarzschild black holes, Gravitation and Cosmology 26 (1) (2020) 75–81.
[20] A. G. Riess, L.-G. Strolger, J. Tonry, S. Casertano, H. C. Ferguson, B. Mobasher, P. Challis, A. V. Filippenko,

S. Jha, W. Li, et al., Type ia supernova discoveries at z¿ 1 from the hubble space telescope: Evidence for past deceleration and constraints on dark energy evolution, The Astrophysical Journal 607 (2) (2004) 665.

[21] C. Shapiro, M. S. Turner, What do we really know about cosmic acceleration?, The Astrophysical Journal 649 (2) (2006) 563.

[22] S. K. J. Pacif, R. Myrzakulov, S. Myrzakul, Reconstruction of cosmic history from a simple parametrization of h, International Journal of Geometric Methods in Modern Physics 14 (07) (2017) 1750111.

[23] W. Yu-Ting, X. Li-Xin, L. Jian-Bo, G. Yuan-Xing, Reconstructing dark energy potentials from parameterized deceleration parameters, Chinese Physics B 19 (1) (2010) 019801.

[24] A. Al Mamon, S. Das, A divergence-free parametrization of deceleration parameter for scalar field dark energy, International Journal of Modern Physics D 25 (03) (2016) 1650032.

[25] G. N. Gadbail, S. Mandal, P. K. Sahoo, Parametrization of deceleration parameter in f (q) gravity, Physics 4 (4) (2022) 1403–1412.

[26] T. Bandyopadhyay, U. Debnath, Fluid accretion upon higher-dimensional wormhole and black hole for parameterized deceleration parameter, International Journal of Geometric Methods in Modern Physics 19 (12) (2022) 2250182.

[27] R. Kundu, U. Debnath, A. Pradhan, Studying the optical depth behaviour of parametrized deceleration parameter in non-flat universe, International Journal of Geometric Methods in Modern Physics (2023).

[28] E. Barboza Jr, J. Alcaniz, Z.-H. Zhu, R. Silva, Generalized equation of state for dark energy, Physical Review D 80 (4) (2009) 043521.

[29] M. Chevallier, D. Polarski, Accelerating universes with scaling dark matter, International Journal of Modern Physics D 10 (02) (2001) 213–223.

[30] E. V. Linder, Exploring the expansion history of the universe, Physical Review Letters 90 (9) (2003) 091301.

[31] H. Jassal, J. Bagla, T. Padmanabhan, Wmap constraints on low redshift evolution of dark energy, Monthly Notices of the Royal Astronomical Society: Letters 356 (1) (2005) L11–L16.

[32] U. Seljak, A. Makarov, P. McDonald, S. F. Anderson, N. A. Bahcall, J. Brinkmann, S. Burles, R. Cen, M. Doi, J. E. Gunn, et al., Cosmological parameter analysis including sdss ly α forest and galaxy bias: constraints on the primordial spectrum of fluctuations, neutrino mass, and dark energy, Physical Review D 71 (10) (2005) 103515.

[33] D. Foreman-Mackey, D. W. Hogg, D. Lang, J. Goodman, emcee: the mcmc hammer, Publications of the Astronomical Society of the Pacific 125 (925) (2013) 306.

[34] W. Handley, M. Hobson, A. Lasenby, Polychord: nested sampling for cosmology, Monthly Notices of the Royal Astronomical Society: Letters 450 (1) (2015) L61–L65.

[35] A. Lewis, Getdist: a python package for analysing monte carlo samples, arXiv preprint arXiv:1910.13970 (2019).

[36] R. Jimenez, A. Loeb, Constraining cosmological parameters based on relative galaxy ages, The Astrophysical Journal 573 (1) (2002) 37.

[37] A. Bouali, I. Albarran, M. Bouhmadi-López, T. Ouali, Cosmological constraints of phantom dark energy models, Physics of the Dark Universe 26 (2019) 100391.

[38] Z. Zhai, Y. Wang, Robust and model-independent cosmological constraints from distance measurements, Journal of Cosmology and Astroparticle Physics 2019 (07) (2019) 005.

[39] D. M. Scolnic, D. Jones, A. Rest, Y. Pan, R. Chornock, R. Foley, M. Huber, R. Kessler, G. Narayan, A. Riess, et al., The complete light-curve sample of spectroscopically confirmed sne ia from pan-starrs1 and cosmological constraints from the combined pantheon sample, The Astrophysical Journal 859 (2) (2018) 101.

[40] N. B. Hogg, M. Martinelli, S. Nesseris, Constraints on the distance duality relation with standard sirens, Journal of Cosmology and Astroparticle Physics 2020 (12) (2020) 019.

[41] M. Martinelli, C. J. A. P. Martins, S. Nesseris, D. Sapone, I. Tutusaus, A. Avgoustidis, S. Camera, C. Carbone, S. Casas, S. Ilić, et al., Euclid: Forecast constraints on the cosmic distance duality relation with complementary external probes, Astronomy & Astrophysics 644 (2020) A80.

[42] C. Wetterich, Phenomenological parameterization of quintessence, Physics Letters B 594 (1-2) (2004) 17–22.

[43] Y. Gong, Model-independent analysis of dark energy: supernova fitting result, Classical and Quantum Gravity 22 (11) (2005) 2121.

[44] E. Barboza Jr, J. Alcaniz, A parametric model for dark energy, Physics Letters B 666 (5) (2008) 415–419.

[45] W. Yang, S. Pan, E. Di Valentino, R. C. Nunes, S. Vagnozzi, D. F. Mota, Tale of stable interacting dark energy, observational signatures, and the h0 tension, Journal of Cosmology and Astroparticle Physics 2018 (09) (2018) 019.

[46] C. Escamilla-Rivera, S. Capozziello, Unveiling cosmography from the dark energy equation of state, International Journal of Modern Physics D 28 (12) (2019) 1950154.

[47] S. Mandal, D. Wang, P. Sahoo, Cosmography in f (q) gravity, Physical Review D 102 (12) (2020) 124029.

[48] S. Pacif, S. Arora, P. Sahoo, Late-time acceleration with a scalar field source: Observational constraints and statefinder diagnostics, Physics of the Dark Universe 32 (2021) 100804.

[49] H. Akaike, A new look at the statistical model identification, IEEE transactions on automatic control 19 (6) (1974) 716–723.

[50] M. Li, X. Li, X. Zhang, Comparison of dark energy models: A perspective from the latest observational data, Science China Physics, Mechanics and Astronomy 53 (9) (2010) 1631–1645.

[51] K. Burnhan, D. R. Anderson, Model selection and multimodel inference, New York: Springer (2002).

[52] K. P. Burnham, D. R. Anderson, Multimodel inference: understanding aic and bic in model selection, Sociological methods & research 33 (2) (2004) 261–304.

[53] A. R. Liddle, Information criteria for astrophysical model selection, Monthly Notices of the Royal Astronomical Society: Letters 377 (1) (2007) L74–L78.

[54] F. Beutler, H.-J. Seo, A. J. Ross, P. McDonald, S. Saito, A. S. Bolton, J. R. Brownstein, C.-H. Chuang, A. J. Cuesta, D. J. Eisenstein, et al., The clustering of galaxies in the completed sdss-iii baryon oscillation spectroscopic survey: baryon acoustic oscillations in the fourier space, Monthly Notices of the Royal Astronomical Society 464 (3) (2017) 3409–3430.

[55] A. J. Ross, L. Samushia, C. Howlett, W. J. Percival, A. Burden, M. Manera, The clustering of the sdss dr7 main galaxy sample–i. a 4 per cent distance measure at z= 0.15, Monthly Notices of the Royal Astronomical Society 449 (1) (2015) 835–847.

[56] L. L. Honorez, B. A. Reid, O. Mena, L. Verde, R. Jimenez, Coupled dark matter-dark energy in light of near universe observations, Journal of Cosmology and Astroparticle Physics 2010 (09) (2010) 029.

[57] L. Anderson, É. Aubourg, S. Bailey, F. Beutler, V. Bhardwaj, M. Blanton, A. S. Bolton, J. Brinkmann, J. R. Brownstein, A. Burden, et al., The clustering of galaxies in the sdss-iii baryon oscillation spectroscopic survey: baryon acoustic oscillations in the data releases 10 and 11 galaxy samples, Monthly Notices of the Royal Astronomical Society 441 (1) (2014) 24–62.

[58] H.-J. Seo, S. Ho, M. White, A. J. Cuesta, A. J. Ross, S. Saito, B. Reid, N. Padmanabhan, W. J. Percival, R. De Putter, et al., Acoustic scale from the angular power spectra of sdss-iii dr8 photometric luminous galaxies, The Astrophysical Journal 761 (1) (2012) 13.

[59] L. Anderson, E. Aubourg, S. Bailey, D. Bizyaev, M. Blanton, A. S. Bolton, J. Brinkmann, J. R. Brownstein, A. Burden, A. J. Cuesta, et al., The clustering of galaxies in the sdss-iii baryon oscillation spectroscopic survey: baryon acoustic oscillations in the data release 9 spectroscopic galaxy sample, Monthly Notices of the Royal Astronomical Society 427 (4) (2012) 3435–3467.

[60] J. E. Bautista, M. Vargas-Magaña, K. S. Dawson, W. J. Percival, J. Brinkmann, J. Brownstein, B. Camacho, J. Comparat, H. Gil-Marín, E.-M. Mueller, et al., The sdss-iv extended baryon oscillation spectroscopic survey: baryon acoustic oscillations at redshift of 0.72 with the dr14 luminous red galaxy sample, The Astrophysical Journal 863 (1) (2018) 110.

[61] T. Abbott, F. Abdalla, A. Alarcon, S. Allam, F. Andrade-Oliveira, J. Annis, S. Avila, M. Banerji, N. Banik, K. Bechtol, et al., Dark energy survey year 1 results: Measurement of the baryon acoustic oscillation scale in the distribution of galaxies to redshift 1, Monthly Notices of the Royal Astronomical Society 483 (4) (2019) 4866–4883.

[62] R. Neveux, E. Burtin, A. de Mattia, A. Smith, A. J. Ross, J. Hou, J. Bautista, J. Brinkmann, C.-H. Chuang, K. S. Dawson, et al., The completed sdss-iv extended baryon oscillation spectroscopic survey: Bao and rsd measurements from the anisotropic power spectrum of the quasar sample between redshift 0.8 and 2.2, Monthly Notices of the Royal Astronomical Society 499 (1) (2020) 210–229.

[63] M. Ata, F. Baumgarten, J. Bautista, F. Beutler, D. Bizyaev, M. R. Blanton, J. A. Blazek, A. S. Bolton, J. Brinkmann, J. R. Brownstein, et al., The clustering of the sdss-iv extended baryon oscillation spectroscopic survey dr14 quasar sample: first measurement of baryon acoustic oscillations between redshift 0.8 and 2.2, Monthly Notices of the Royal Astronomical Society 473 (4) (2018) 4773–4794.

[64] V. de Sainte Agathe, C. Balland, H. D. M. Des Bourboux, M. Blomqvist, J. Guy, J. Rich, A. Font-Ribera, M. M. Pieri, J. E. Bautista, K. Dawson, et al., Baryon acoustic oscillations at z= 2.34 from the correlations of lyα absorption in eboss dr14, Astronomy & Astrophysics 629 (2019) A85.

[65] C. Blake, S. Brough, M. Colless, C. Contreras, W. Couch, S. Croom, D. Croton, T. M. Davis, M. J. Drinkwater, K. Forster, et al., The wigglez dark energy survey: Joint measurements of the expansion and growth history at z¡ 1, Monthly Notices of the Royal Astronomical Society 425 (1) (2012) 405–414.

X. APPENDIX

BAO name	**redshift z**	**Experiment**	**Measurement**	**Standarddeviation**	Ref.
6dFGS	0.106	r_s/D_V	0.336	0.015	[54]
SDSS DR7	0.15	$D_V(r_{s,fidd}/r_s)$	664	25.0	[55]
SDSS-DR7 + 2dFGRS	0.275	r_s/D_V	0.1390	0.0037	[56]
SDSS-DR11 LOWZ	0.32	$D_V(r_{d,fidd}/r_s)$	1264	25	[57]
SDSS-III DR8	0.54	D_A/r_s	9.212	0.41	[58]
SDSSIII/ DR9	0.57	D_V/r_s	13.67	0.22	[59]
SDSS-IV DR14	0.72	$D_V(r_{s,fidd}/r_s)$	2353	63	[60]
DES Year 1	0.81	D_A/r_s	10.75	0.43	[61]
DECals DR8	0.874	$D_A(r_{s,fidd}/r_s)$	1680	109	[60]
	0.72	$D_V(r_{s,fidd}/r_s)$	2353	63	
eBoss DR16 BAO+RSD	1.48	$D_H.r_s$	13.23	0.47	[62]
SDSS-IV/DR14	1.52	$D_V(r_{s,fidd}/r_s)$	3843	147.0	[**?**]
Boss Lya quasars DR9	2.3	$H.r_s$	34188	1188	[63]
BOSS DR14 Lya in LyBeta	2.34	$D_H.r_s$	8.86	0.29	[64]
WiggleZ	0.44		0.0870	0.0042	
	0.6	r_s/D_V	0.0672	0.0031	[65]
	0.73		0.0593	0.0020	

TABLE III. Summary of the Baryon Acoustic Oscillations measurements used in this paper.

Multimedia Tools and Applications
https://doi.org/10.1007/s11042-023-14353-w

Conv-CapsNet: capsule based network for COVID-19 detection through X-Ray scans

Pulkit Sharma[1] · **Rhythm Arya**[1] · **Richa Verma**[1] · **Bindu Verma**[1]

Received: 12 March 2022 / Revised: 9 June 2022 / Accepted: 2 January 2023
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

Abstract
Coronavirus, a virus that spread worldwide rapidly and was eventually declared a pandemic. The rapid spread made it essential to detect Coronavirus infected people to control the further spread. Recent studies show that radiological images such as X-Rays and CT scans provide essential information in detecting infection using deep learning models. This paper proposes a shallow architecture based on Capsule Networks with convolutional layers to detect COVID-19 infected persons. The proposed method combines the ability of the capsule network to understand spatial information with convolutional layers for efficient feature extraction. Due to the model's shallow architecture, it has $23M$ parameters to train and requires fewer training samples. The proposed system is fast and robust and correctly classifies the X-Ray images into three classes, i.e. COVID-19, No Findings, and Viral Pneumonia. Experimental results on the X-Ray dataset show that our model performs well despite having fewer samples for the training and achieved an average accuracy of 96.47% for multi-class and 97.69% for binary classification on 5-fold cross-validation. The proposed model would be useful to researchers and medical professionals for assistance and prognosis for COVID-19 infected patients.

Keywords COVID-19 detection · Capsule networks · Medical imaging · Chest X-ray classification

1 Introduction

In December of 2019, when Wuhan, a city in China, reported a cluster of fatal pneumonia cases, it marked the start of a series of cases rising worldwide and eventually converted into a global pandemic with a new variant SARS-CoV. If a person gets infected with COVID-19, it depletes the respiratory tract and affects the lungs. People generally experience high temperatures and body aches with cough and breathing problems. In many cases, people suffer from fatal pneumonia, which causes lung inflammation due to SARS-CoV. The after-effects of COVID-19 could be fatal and include degrading immunity, breathing problems,

✉ Bindu Verma
bindu.cvision@gmail.com

[1] Delhi Technological University, Delhi, India

heart problems, etc. COVID-19 [1] has now proven to be one of the most contagious viral pandemics in history. The fast spread of the virus forced most nations to shut their economy and work remotely. Medical operations and systems faced problems worldwide. The researchers have been on their toes to understand the scenario and the virus's behaviour. There have been extensive studies and experiments performed to devise radiological methods to detect the presence of the COVID-19 virus in the body. The most popular method used nowadays is RT-PCR through the mouth and nasal swabs and X-Ray scan analysis.

The second wave of the pandemic that hit India in March 2021 was devastating in many aspects, such as the unavailability of proper healthcare facilities, panic among the general public, lack of testing mechanisms, and less knowledge about the evolving virus. Even after more than two years of the global pandemic, people's physical and mental health has constantly declined. There is a need to have proper management for screening COVID-19 cases in the country and a need for rapid and promising results. The pandemic of SARS-CoV-2 new Coronavirus (COVID-19) has accompanied a torrent of disruption. One thing that every individual can follow apart from the proven public health practices is getting tested as soon as possible. This situation has led to a generation of quicker testing mechanisms to get faster and more promising results. This deadly disease in India brought multiple waves, with the fatal second wave leading to a high mortality rate. Many tests are done to find whether the COVID-19 virus is present inside the suspected person, including rapid mouth and throat swab tests, RT-PCR tests, Chest X-Rays, and various other radiological methods.

The X-Ray scan technology is cheaper, faster, and has less radiation than the CT scan procedure. However, some challenges are associated with them as well. It is not easy to diagnose pneumonia manually, and white spots on X-Ray images need to be examined to determine whether it is a coronavirus, tuberculosis, or bronchitis. Any misdiagnosis leads to the wrong treatment of the patients. Figure 1 shows the chest X-Ray image of a COVID-19 patient taken on the day 1,4,5 and 7 day.

Manual examination of the X-Ray images provides a correct diagnosis of COVID-19 at around $50 - 60\%$. Still, with the help of an artificial intelligence model, accurate diagnosis can be increased for both COVID-19 and other kinds of pneumonia. Another significant problem the researchers have encountered in this scenario is the lack of uniform, huge, and labelled datasets. For smaller datasets, Capsule Networks are known to perform better as compared to Convolutional Neural Networks [37, 44, 45]. Due to the increased research in computer vision and the recent advancement of computation systems, deep learning has been widely used in image classification and retrieval, object detection, medical imaging, signal processing, and video classifications in recent years [7, 10–13]. Deep learning-based models are also widely used in various classification problems [21, 54, 55] using AlexNet, VGGNet, GoogleNet, ResNet, and Inception Net, which have achieved good recognition accuracy. Hinton et al. [45] proposed a new class of convolutional neural networks, Capsule Networks (CapsNet), that have shown an outstanding performance in many image processing and natural language processing problem [3, 31, 38, 56, 58, 60]. The advantage of using capsule network over other CNN model is that it is a shallow and lighter network and requires less number of parameters compared to the other CNN models. Moreover, a capsule network gives remarkable performance even when training samples are less and without any data augmentation. A Capsule network can capture the spatial and temporal information from the image, improving the accuracy of the traditional CNN network, where spatial information is lost at the pooling layer. According to Sabour [45] capsule network has view in-variance property and view equivalence in the internal structure of the network. So given all the facts, it is believed that the capsule network is an excellent choice for our study. Thus,

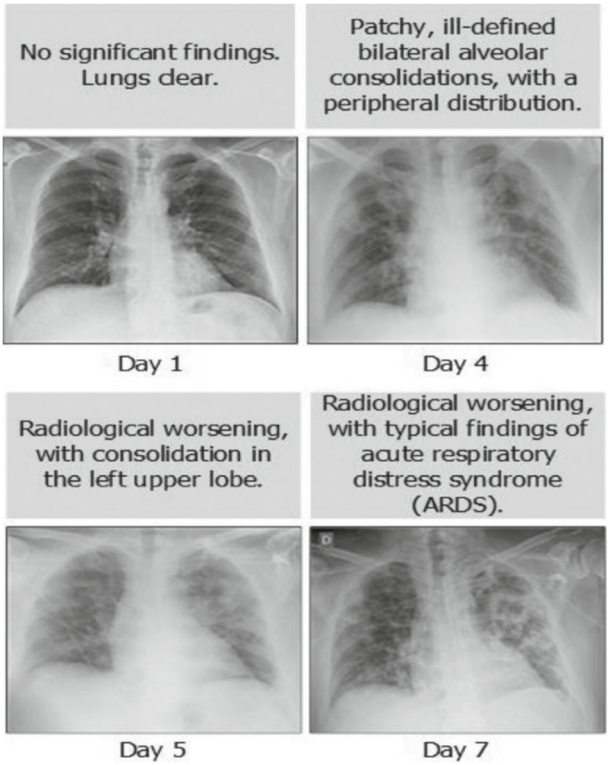


Fig. 1 Chest X-Ray image of COVID-19 Patient [17]

we propose a modified capsule network that includes additional convolutional layers before the primary capsule layers. A chest X-Ray image is passed to the convolutional layers during the training process. They extract discriminating features from the images and pass the feature vector into primary capsule layers. A total of 32 capsules are generated, and the vectors are further passed to the X-Ray capsule layer. This layer contains $16D$ vectors and is used to classify the chest X-Ray images into three classes (COVID-19, Normal, and Pneumonia) using M_l margin loss.Overall, this paper presents a novel framework for COVID-19 detection using chest X-Ray and has several key contributions:

- This paper proposes a new capsule network architecture, Conv-CapsNet (Convolutional Capsule Network), with additional convolution layers followed by primary and X-Ray capsule layers. The advantage of adding convolutional layers is to provide the most discriminating feature map to the primary caps layer.
- Conv-CapsNet can capture the spatial and temporal information from the image and achieve improved accuracy over the traditional deep learning models in which spatial information is lost at the pooling layer.
- Proposed Conv-CapsNet can achieve good accuracy even if fewer data samples are available and outperform the literature without any data augmentation or new data.
- The proposed framework can be deployed as one of the useful techniques in the medical field for classification purposes and prognosis. It should help radiologists diagnose COVID-19 as a precautionary measure.

The rest of the paper is organized as follows: Related work is discussed in Section 2. In Section 3, the proposed work is discussed in detail with Conv-CapsNet architecture and algorithm. We have done an ablation study with experimental results and analysis in Section 4. Section 5 discusses the conclusion of the study with the future directions.

2 Related work

When the pandemic began, researchers and medical professionals quickly tried to combat the virus by developing various detection methods such as using RT-PCR, CT Scans, X-Rays, and symptomatic diagnosis. In Artificial Intelligence, various imaging analyzing systems and recommendation systems were tested with limited data samples available. Initially, when there was a panic situation around the world, other than RT-PCR and rapid antigen tests, the suspected individuals underwent X-Ray scanning. The scans were analyzed using radio-logical methods. The locations of lung opacity were analyzed, and a threshold was kept to determine whether the patient is COVID-19 positive or not [18]. Various telehealth services, including mobile and social media-based services, were also being provided for the patients with different network models [53]. Similarly, to assist COVID-19 quarantine patients, a wearable device that helps to monitor heart rate, oxygen level, and temperatures proved helpful. Oxygen therapies and ventilators were being provided for critical cases [26, 27]. In the field of Artificial Intelligence (AI), many machine learning models such as Linear Regression, Support Vector Machine, and Multi-Layer Perceptron were being developed using various features such as sex, age, and clinical values [41, 61] to predict COVID-19 disease. Rahman et al. [43] developed a neural network to detect whether a person was wearing a mask in public or not with an accuracy of around 98%. Likewise, in deep learning, based on a few X-Ray scans and CT scans available, neural network-based models were being used to extract the features and classify the patient whether they are infected or not [9, 52].

Neural network architectures were being used to study X-Ray scans for medical purposes such as the detection of pneumonia. This methodology was further used to study COVID-19 infected chest X-Rays for prediction and classification of COVID-19 [22, 25, 59]. Ismael et al. [28] provided a novel CNN model wherein deep features were extracted from ResNet50 and SVM with linear kernel function on a dataset containing 180 COVID-19 images and 200 healthy images. They obtained an accuracy of 94.7%. Apostolopoulos et al. [6] used a convolutional neural network to classify the chest X-Ray images into a binary class COVID-19 and no findings. Similarly, Narin et al. [35] analyses a various deep learning models such as ResNet50, ResNet101, ResNet152, InceptionV3 and Inception-ResNetV2 for automatic detection of COVID-19. Wang et al. [57] proposed a new deep neural network called a COVID-Net, which is an open-source tool and also provided an open-source dataset COVIDx for researchers. Sethy et al. [47] extracted the in-depth features using ResNet50 and then used a support vector machine to classify the CXR images with around 95% accuracy. In addition, Law et al. [32] implemented transfer learning using VGG-16 on the COVID-19 patient data set and found that VGG-16 gives superior performance metrics as compared to the other ResNet models. In recent studies, author Sharma et al. [48] proposed a COVDC-Net that is a combination of MobileNetV2 and VGG16 to extract the features. Then both model scores were fused to get the final prediction for multi-class classification. Similarly, CVDNet model [36] was proposed based on a residual neural

network with two parallel levels for local and global features, CoroDet model [23] was proposed, which is a 22 layer deep CNN with both X-Rays and CT scans. Rahman et al. [42] in their study utilized various image enhancement techniques such as HE, CLAHE, and BCET to understand their effect on COVID-19 detection. They also proposed a novel U-Net model with gamma correlation-based enhancement performing the best. Mousavi et al. [33] presented a CNN-LSTM-based network testing on six different databases and separated COVID-19 infected scans from healthy ones and other lung diseases with more than 90% accuracy. Distinctive filters could also be utilized to study different kinds of pneumonia. Karthik et al. [29] proposed a CNN architecture that could learn unique filter patterns for different kinds of pneumonia. It achieved an F1 score of 97.2%.

Similarly, in other studies, Rahimzadeh et al. [40] developed an integrated CNN-based model for classification using Xception and ResNet50V2. Traditional methods are also used to detect the COVID-19 [5] using classifiers such as SVM, CNN, and RF to detect the COVID-19 infection. Author Muhammad et al. [34] used data mining approaches such as decision tree, support vector machine, naive Bayes, logistic regression, random forest, and K-nearest neighbour algorithms were applied to predict the minimum and the maximum number of days needed to recover from a virus. When there are fewer samples to train the model, data can be generated by an adversarial network [30] used a generative adversarial network (GAN) to generate the more COVID-19 X-Ray images. In the paper [30] despite having fewer training samples, they have achieved good accuracy using ResNet, AlexNet, and VGGNet. Ucar et al. [51] proposed a novel model that used the X-Ray images to classify confirmed and negative cases of COVID-19 based on squeeze-net.

Transfer Learning techniques have been helpful as well in this research area. Avola et al. [8] presented a comparative analysis of X-Ray images with 12 Image pre-trained models where MobileNet performed the best. Chandra et al. [15] used two-phase classification using a majority vote-based classifier ensemble with different supervised algorithms. Dimeglio et al. [20] by employing a well-known DenseNet-121 deep learning architecture, classify the chest X-Ray image into COVID-19 positive and negative classes. Sheykhivand et al. [49] developed a model for pneumonia classification in which GANs architecture is used with LSTM networks and transfer learning. Chakraborty et al. [14] developed a deep learning model to classify COVID, pneumonia, and healthy from the total dataset. Dalvi et al. [19] employed a well-known technique of deep feature extraction along with ResNet50 and Logistic Regression, and an advanced model was constructed to predict COVID with reasonable accuracy. Islam et al. [24] developed a technique combining CNN and LSTM models to detect whether a person is infected or not using X-Ray images. Author Saha et al. [46] proposed an EMCNet model where in-depth and high-level features are extracted using a convolutional neural network from chest X-Ray images. Then machine learning classifiers such as random forest, support vector machine, ADA boost, and decision tree are used to detect the COVID-19. Author Rakhami et al. [4] used a combination of convolutional neural networks such as VGG19, DenseNet121, and Inception V3 to extract the feature then RNN was used to detect the COVID-19. They also visualize the specific reason for the image responsible for the detection using Gradient-weighted Class Activation Mapping (Grad-CAM).

In the recent studies, we see that chest X-Ray images are very few, and most of the authors either used data augmentation technique or GAN to generate the new images so that overfitting and underfitting problems of the deep learning models can be solved. But above methods take much time to train the model as the number of trainable parameters is more and not suitable for real-time deployment. Also, most of the methods extract features

from deep learning models and detection using machine learning algorithms. These models are not feasible for real-time implementation. Thus, this paper proposes a modified capsule network combination of convolutional filters, primary capsule layers, and digits caps layers that work perfectly if the training sample is less. With fewer datasets available to research in this area, Capsule Networks perform better in accuracy than other Convolution-based architectures. Moreover, a combination of only a few layers has a shallow structure and is suitable for real-time deployment in clinical trials.

3 Proposed work

3.1 Capsule networks

Capsule Networks and the concept of dynamic routing between capsules were given by Hinton and his fellow researchers in 2017 in his research paper [45]. He proposed this network by providing a considerable performance for the MNIST digit dataset. These networks showed improvement in performance over Convolutional Neural Networks consisting of a small number of trainable parameters [37, 44, 45]. The main reason behind this could be that CNN has pooling layers used for down-sampling the maps of images. In this way, a CNN is not able to capture small details. Capsule Networks were developed to understand the spatial features and minute details, which were one of the drawbacks of using the pooling layer in the CNN. These networks consist of capsules, which are different from neurons as they output vectors that could differentiate images in different directions.

3.1.1 Architecture of capsule network

The Capsule Network is used to classify handwritten digits of the MNIST dataset. There are six layers in a typical CapsNet, which are used for encoding and decoding purposes. The first layer is a convolutional layer followed by a primary layer consisting of one more convolutional layer with a squashing function. Further, it consists of a digit or label capsule where dynamic routing occurs. The last three layers are fully connected (FC) layers used in image reconstruction (Fig. 2).

Unlike convolutional layer that used different activation function to activate the neurons, primary caps layer used a non-linear squashing function in (1) to convert the neurons into the vector form.

$$v_j = \frac{||S_j||^2}{1 + ||S_j||^2} \frac{S_j}{||S_j||} \tag{1}$$

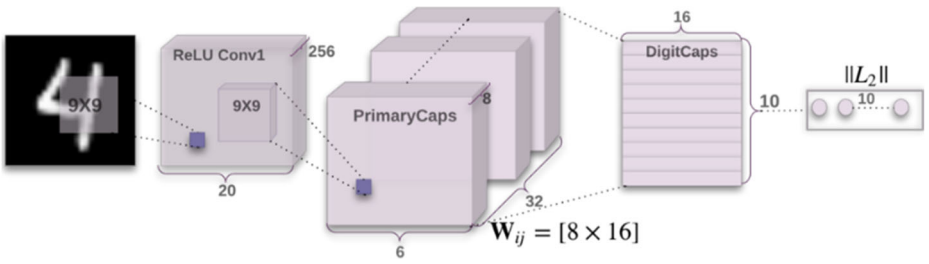


Fig. 2 Original capsule network architecture [45]

where v_j is a computed output vector of capsule j and S_j is total input to capsule. The total input to a capsule S_j can be calculated in (3) by the weighted sum over all the prediction vector $\hat{P}_{j|i}$ from the lower layer capsules. The prediction vector in below layer can be computed using (2) as multiplication of output p_i and weight matrix W_{ij}.

$$\hat{P}_{j|i} = W_{ij} p_i \tag{2}$$

$$S_j = \sum_i k_{ij} \hat{P}_{j|i} \tag{3}$$

where, k_{ij} is a coupling coefficient that is calculated using (4) using the dynamic routing algorithm [45]. Dynamic routing is a process between the capsules to decide how to move forward in the network. This happens with routing by agreement between higher and lower capsules. The capsules at high level receive input from the lower level capsule by its agreement iteratively. It is recommended to use only three iterations because higher iterations would overfit the model.

$$k_{ij} = \frac{exp(n_{ij})}{\sum_k exp(n_{ik})} \tag{4}$$

where, n_{ij} denotes the log probability calculated using softmax function. The sum of coupling coefficient k_{ij} between capsule i and top layer should be 1. To calculate the k_{ij}, n_{ij} must be calculated using (5).

$$n_{ij} \leftarrow n_{ij} + \hat{P}_{j|i} \cdot v_j \tag{5}$$

Initially, the value of n_{ij} is 0, and using this, we can get k_{ij} and p_i, which is the output of the previous layer of the capsules. Using these three values, we can calculate the next level capsule input S_j.

For margin loss, the loss is calculated in (6) for every class to determine whether a particular object is present in the class or not.

$$M_l = C_k \max(0, u^+ - ||v_k||^2) + \lambda(1 - C_k) \max(0, ||v_k|| - u^-)^2 \tag{6}$$

where, $C_k = 1$ iff a digit of class k is present and $u^+ = 0.9$ and $u^- = 0.1$ are the hyper parameter denotes the down weighting of the loss. The total loss is the sum of the losses of all the digit capsules.

3.2 Proposed convolutional capsule network

The inability of CNNs to distinguish different directions and spatial arrangement of the entire or a portion of the image is one of the critical issues that guided the development of CapsNet. Because of down sampling, CNN loses some picture properties. As a result, vast amounts of training data are required for loss compensation. While capsule network uses capsules which is a vector with spatial information. This means that Capsule Networks will perform better on smaller datasets without augmenting the dataset. Capsule networks are shallow network systems that mimic how a human brain learns by modelling and simulating hierarchical relationships in processing a picture.

We were inspired by the research [45] that introduces a new training mechanism, called dynamic routing between capsules, for iteratively training the capsule network architecture, which gives a greatly enhanced performance over the conventional CNNs. Modification of capsule network is proposed in which we have added 4 convolutional layer followed by capsule network layers. We name the modified capsule network a Convolutional Capsule Network (Conv-CapsNet). Figure 3 shows the proposed network architecture. More Convolutional Layers to the capsule network were added to provide more effective information in

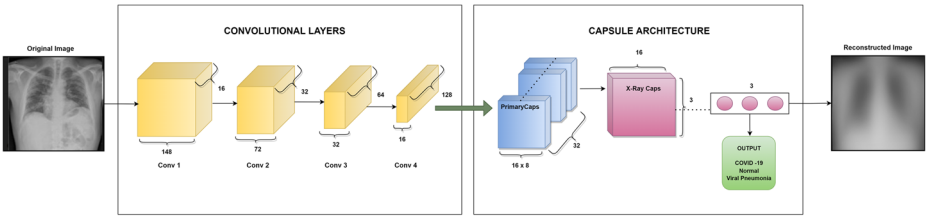


Fig. 3 Proposed convolutional capsule network architecture

terms of features of the primary capsule. After experimentation with the number of layers with kernel and filter sizes, four Convolutional Layers were the optimal number of layers to perform feature extraction. We resized these images to 150 × 150 for the best possible performance after experimenting with different sizes of images.

1: TRAINING PHASE
2: INPUT: Labelled training set $\overline{I} = I^1, I^2, I^3$
3: CONVOLUTION LAYERS:
4: Feature Maps $\overline{I} = I^1, I^2, I^3 \rightarrow \overline{F} = F^1, F^2, F^3$
5: CAPSULE LAYERS: $\leftarrow \overline{F}$
6: Calculate activation v_j for l + 1 layer
7: $p \leftarrow$ activation for l layer
8: Prediction Vector $\hat{P}_j = W.p$, W = weighted matrix
9: Output Vector $S = \Sigma k.\hat{P}_j$, k = coupling coefficient
10: Dynamic Routing for k in lower level capsules
11: Activation $v_j = squash(P)$
12: Calculate margin loss M_l for class k with threshold 0.5
13: DECODER:
14: Calculate reconstruction loss R_k
15: Mask activity vectors a except a_k, k = correct class
16: Reconstruct with decoder FC layers
17: OUTPUT:
18: Reconstructed Image
19: Vector $\overline{A}$ with the probability of class

Algorithm 1 Proposed Conv-CapsNet.

Initially, an image is passed to the series of convolution layers. The advantage of adding a more convolutional layer is to provide a most discriminating feature map to the primary capsule layer. A 150 × 150 size image is passed into the convolutional layer to classify the X-Ray image. The detailed input dimension of each layer is shown in Table 1. First convolutional layer trains 16 kernel of size 5 × 5 with a stride of 1. Then the max-pooling operation is performed with a stride of 2. The same kernel size is used in the next two consecutive layers, and kernel numbers are 32 and 64, respectively. Then, the fourth convolutional layer 128 kernel is trained of size 9 × 9 with the stride of 1. ReLU activation function is used in all the layers.

Table 1 Proposed Conv-CapsNet layers

Layer Type	Output Shape
Conv 2D	[-1, 16, 148, 148]
ReLU	[-1, 16, 148, 148]
MaxPool 2D	[-1, 16, 74, 74]
Conv 2D	[-1, 32, 72, 72]
ReLU	[-1, 32, 72, 72]
MaxPool 2D	[-1, 32, 36, 36]
Conv 2D	[-1, 64, 32, 32]
ReLU	[-1, 64, 32, 32]
MaxPool 2D	[-1, 64, 17, 17]
Conv 2D	[-1, 128, 16, 16]
ReLU	[-1, 128, 16, 16]
Conv 2D	[-1, 256, 5, 5]]
Primary Capsule	[-1, 800, 8]
X-Ray Capsule	[-1, 3, 16]
Linear	[-1, 512]
ReLU	[-1, 512]
Linear	[-1, 1024]
ReLU	[-1, 1024]
Linear	[-1, 22500]
Sigmoid	[-1, 22500]

Input to the Primary Capsule Layer is a refined feature of size $16 \times 16 \times 128$ and output is 32 capsules each of size $6 \times 6 \times 8$. The primary capsule layer first applies convolutional operation with 9×9 kernel, then squash the output to get a capsule. The output of the capsule layer is $6 \times 6 \times 8$ capsules where 6×6 grid shared their weights and orientation. The final layer is an X-Ray caps layer which contains $16D$ vectors per class and is used to classify the X-Ray image into three classes COVID-19, Normal and Viral Pneumonia using M_l margin loss. The X-Ray caps layer is also used to reconstruct the image using three fully connected layers.

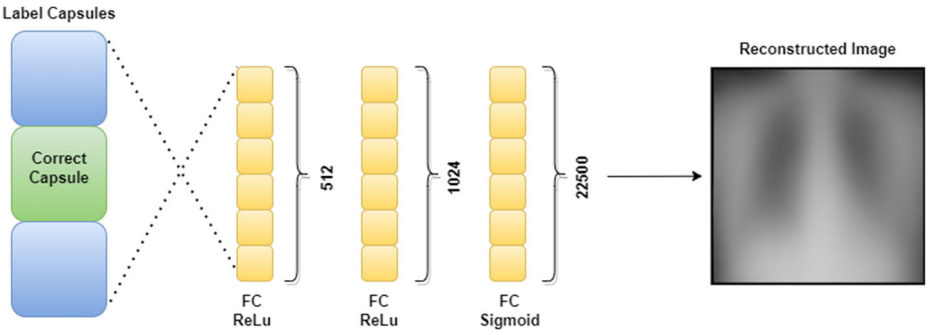


Fig. 4 Reconstructing a decoding structure from the DigitCaps layer

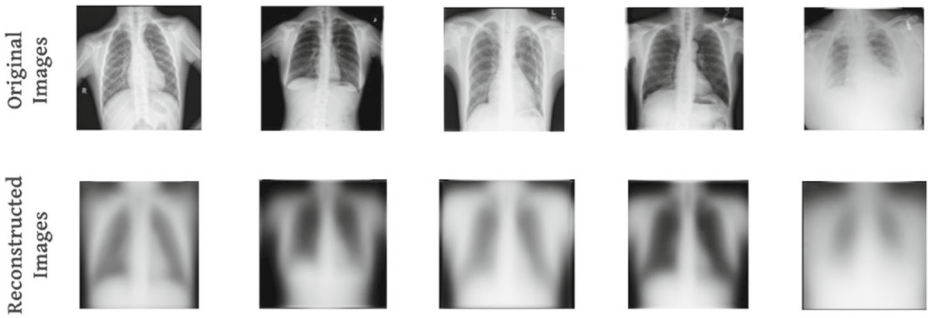


Fig. 5 Figure shows the original image and reconstructed image

3.3 Image reconstruction

During the training process, activity vectors are masked out, which are used to reconstruct the input image. The output of the X-Ray caps layer is fed into the three fully connected layers that model the pixel intensities as shown in Fig. 4. The sum of square differences between the reconstructed image and the actual input image. Low error represents that the reconstructed image is similar to the input image. The reconstruction loss is scaled down to 0.005 so that it does not dominate the margin loss during training. As shown in Fig. 5 we can say that reconstruction from the $16D$ X-Ray caps layer is robust while keeping only essential information. Some of the reconstructed images are a blur due to the change in the training data or CapsNet's inability to handle the noise in the data.

4 Experimental results

4.1 Dataset and setup

The dataset of **COVID-19 Radiography** [16, 42] which has been developed by the researchers at University of Qatar with the help of medical professionals at Dhaka is publicly available to the researchers.[1] It is a huge database of chest X-Ray scans for COVID-19 based pneumonia, chest X-Ray scans of normal people as well as people suffering from viral pneumonia as shown in Fig. 6. The metadata of the dataset has also been provided separately in CSV files containing the URLs of the source of the images. The dataset includes images of each class - COVID-19 (3616), Normal (10192) and Viral Pneumonia (1345) as shown in Table 2.

The Conv-CapsNet is implemented in Python using the PyTorch library. Experiments are performed on 64-bit windows 10, Intel Core i7 processor with 16GB RAM, 8GB NVIDIA GETFORCE GTX graphic card. Table 3 shows the hyper parameters of the proposed model.

4.2 Experimental discussion

We have not deployed the model in real-time for clinical use, but it can be deployed for the clinical trial. In the literature, Certain technologies and tools have been developed recently,

[1] https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database.

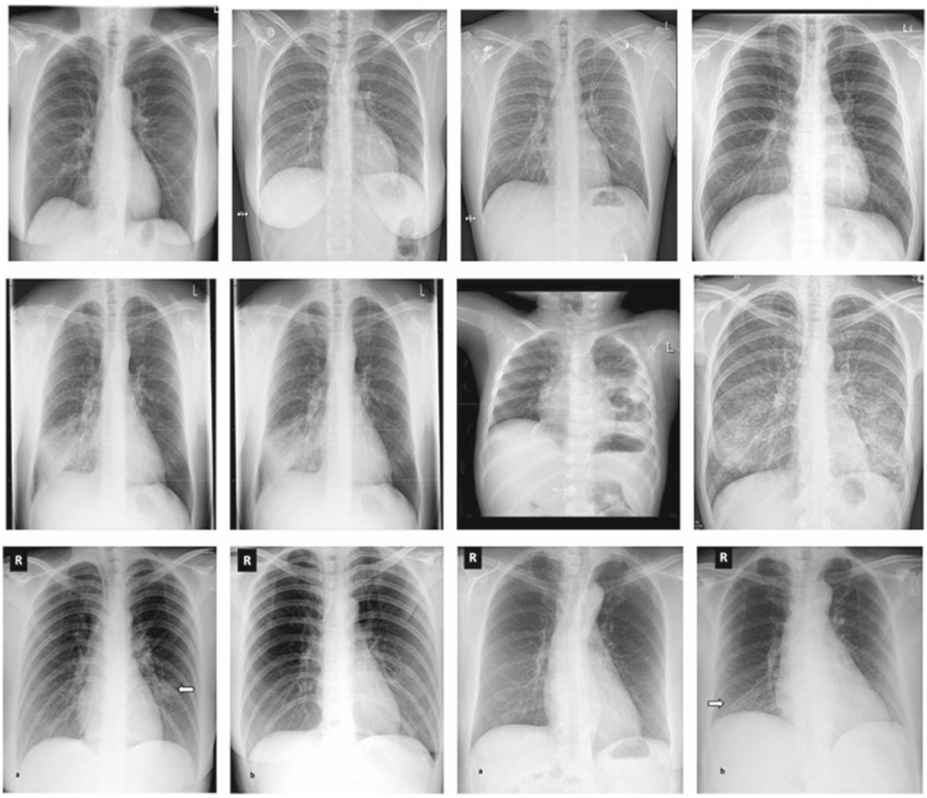


Fig. 6 Four scans of COVID-19 patients, viral induced pneumonia and normal scans in the subsequent rows

such as XraySetu, a platform developed in India for Chest X-Ray interpretation over WhatsApp. Doctors and Patients who have used this tool were happy that the tool helped in the early diagnosis of patients, and this tool has the potential to reach rural and underdeveloped areas. There are limitations to these techniques as well, doctors can not completely rely on these for medical use, and professional intervention is necessary. The good accuracy of our proposed model and shallow architecture made it feasible for real-time deployment. Once the model is trained on the Chest X-ray data, it can be saved on the cloud and can be used for the diagnosis.

Table 2 Data set split

	Normal	Viral Pneumonia	Normal	Total
Train	6523	861	2314	9698
Validation	1631	215	579	2425
Test	2038	269	723	3030

Table 3 Hyper parameters for proposed model

Epochs	30
Batch Size	64
Learning Rate	0.005
Routing	3
Optimizer	Adam
Loss	Categorical Cross Entropy

4.2.1 Image pre-processing

The images are modified and enhanced in such a way so that the data is more suitable to the model and improves performance. The initial image size in the dataset was 300 × 300 pixels. Next, we resize the images into 150 × 150 pixels. Furthermore, we converted the RGB image to a Grayscale image by reducing the channels from 3 to a single channel. We then analyzed their pixel intensities and applied the Contrast Limited Adaptive Histogram Equalization (CLAHE) technique. The complete procedure is shown in Fig. 7.

4.2.2 Evaluation metrics

We evaluated our proposed model regarding precision and recall and performed 5-fold cross-validation to get the average accuracy. We randomly split the dataset into 80% of the training samples and 20% of the testing samples and then trained the model in 5 fold and recorded the average accuracy of the proposed model. We have also compared the literature based on the average accuracy achieved on the X-Ray dataset. The used parameters to compare the performance are defined as follows:

- True Positive (TP) shows the number of correctly identified COVID-19.
- False Negative (FN) shows the number of incorrectly identified Non-Findings.
- True Negative (TN) shows the number of correctly identified No-findings.
- False Positive (FP) shows the number of incorrectly identified COVID-19.

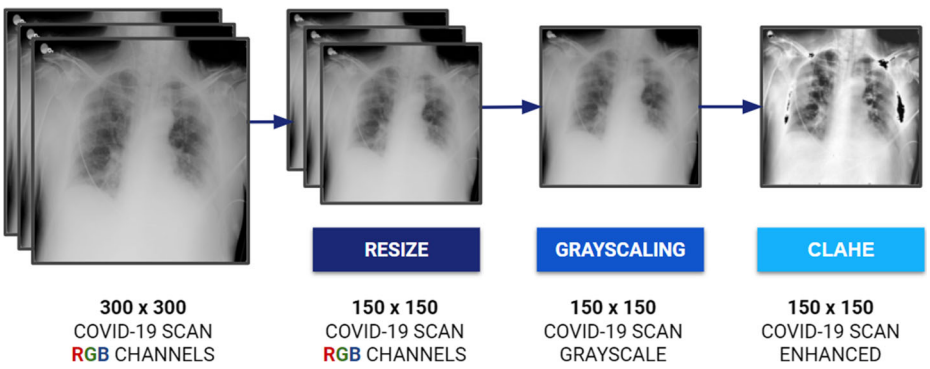


Fig. 7 Enhancement route followed

Sensitivity refers to the rate of positive cases who actually have the condition and are detected true and as a result the test yielded a positive result.

$$Sen = \frac{TP}{(TP + FN)} \times 100 \tag{7}$$

Specificity refers to the rate of negative cases who actually did not have the condition and are detected false and as a result the test yielded a negative result .

$$Spe = \frac{TN}{(TN + FP)} \times 100 \tag{8}$$

Precision is an estimate of appropriateness in extraction of information.

$$Pre = \frac{TP}{(TP + FP)} \tag{9}$$

F1 score It can be defined as the harmonic mean of the recall and precision.

$$F1 = \frac{2TP}{(2TP + FP + FN)} \tag{10}$$

Accuracy can be calculated using the following equation.

$$Acc = \frac{TP + FN}{(TP + FP + TN + FN)} \times 100 \tag{11}$$

AUC-ROC curve is a measure explaining how much model can separate different classes. Higher the AUC, better is the performance of the model.

4.3 Results of the proposed model

The proposed Conv-CapsNet is evaluated for the multi-class and the binary class. In multi-class, three classes are used, which are COVID-19, Normal and Viral Pneumonia, and in binary class, only two classes are used, which are COVID-19 and Non-COVID-19 as shown in Fig. 8. We evaluated our model using 5-fold cross-validation. Training and validation

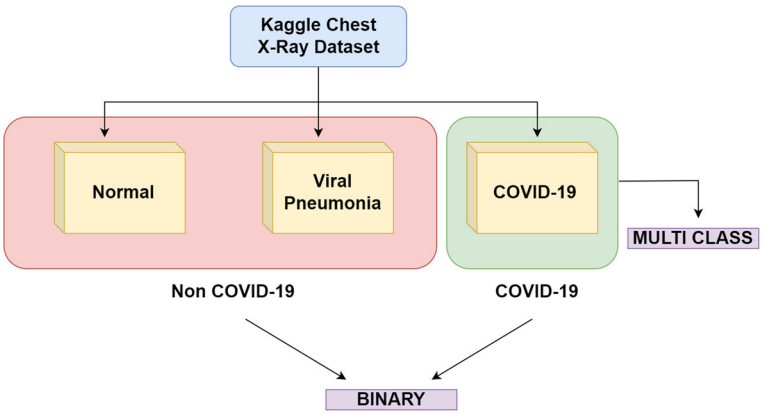


Fig. 8 Dataset classes

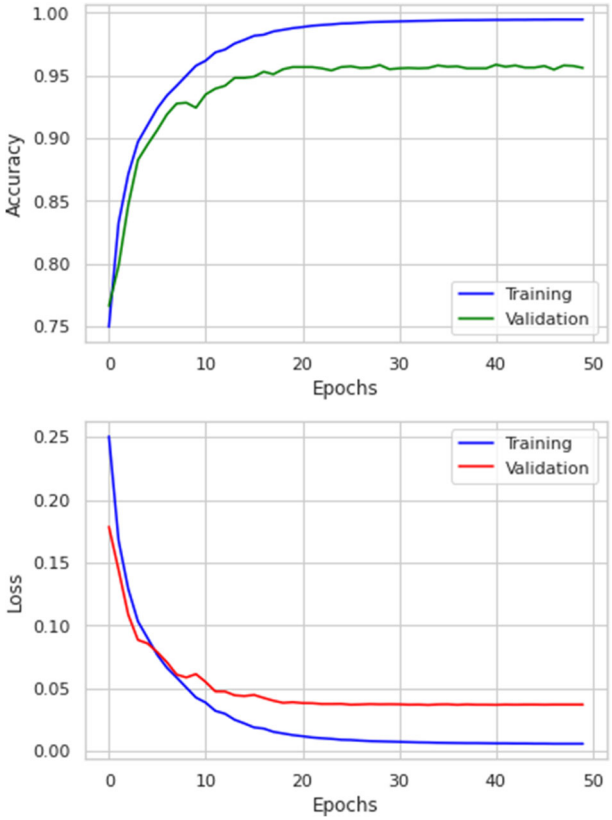


Fig. 9 Performance diagram of Conv-CapsNet on X-Ray dataset using 80 – 20% split for multi-class classification. The graph shows training accuracy vs validation accuracy and training loss vs validation loss

results of multi-class classification are shown in Fig. 9. Class wise average, precision, recall and F1-scores are shown in Table 4. F1 scores for all the classes are more than 94%, which means the proposed model is balanced and perform well for all the classes. PR curve was plotted for all three classes as shown in Fig. 10. The model provides better accuracy for the Normal class, with almost similar performance for COVID-19 and Viral Pneumonia classes. The proposed model achieved 96.47% average accuracy. The confusion matrix of the proposed model is shown in Fig. 11. It is observed that around 4% of scans which were

Table 4 Precision, recall and F1 score of all three classes for multi-class classification

Class	Pre (%)	Rec (%)	F1 (%)
COVID-19	94.65	93.08	93.86
Viral Pneumonia	95.16	95.12	95.15
Normal	97.22	97.79	97.56
Accuracy			96.47
Macro Avg.	96.59	95.87	96.22
Weighted Avg.	96.47	96.47	96.47

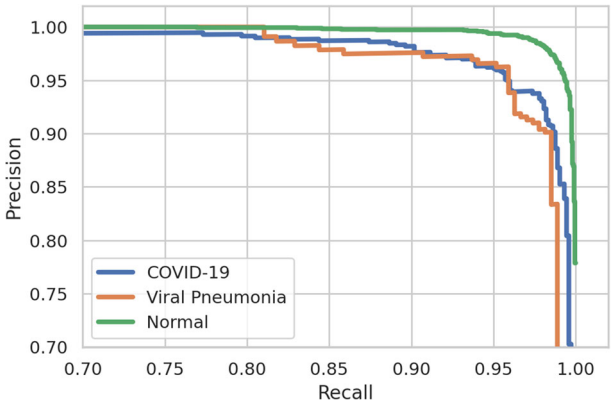


Fig. 10 Precision recall curve for COVID-19, Viral Pneumonia and Normal classes

COVID-19 infected were predicted as Normal, which is a significant improvement from the original capsule network where the percent of scans was 21%. Figure 12 shows the ROC-AUC curve for all the three classes as well as their macro average. The model performs well with an AUC of 0.99 for all classes. It means it distinguishes well between the classes.

We performed a comparative analysis of the proposed model, Conv-CapsNet, with the original capsule network. The performance of the proposed Conv-CapsNet is analyzed based on the classification accuracy of each class and comparing the confusion matrices for more insights. As we can see in Fig. 13, Conv-CapsNet performed better than the original capsule network in terms of class-wise accuracy. Moreover, the performance of the COVID-19 class has significantly improved in our proposed model (94.95% compared to the original capsule, which is very low (78.01%). The main focus of this study is to find the maximum number of positive cases so that further spread of the virus can be stopped. In our proposed Conv-CapsNet, with evidence, it has performed better, and it achieved more than 90% accuracy.

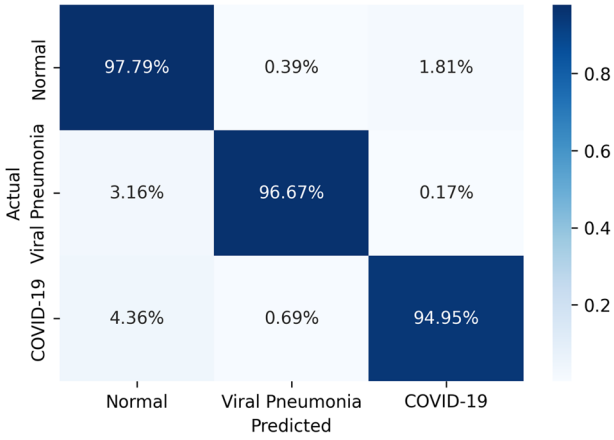


Fig. 11 Figure shows the confusion matrix of the proposed convolutional capsule network for multi-class classification

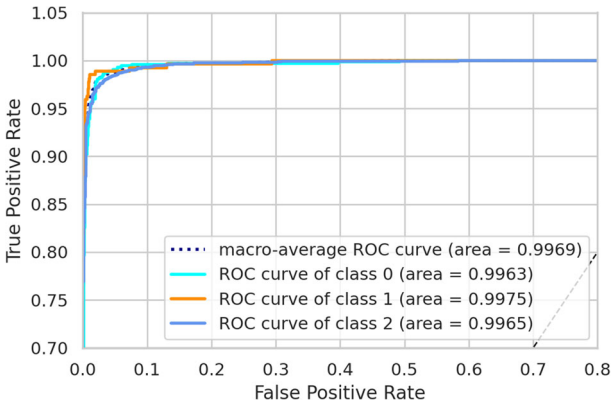


Fig. 12 ROC curves for Multi-Class Classification with COVID-19 (0), Viral Pneumonia (1) and Normal (2) classes

Different parameters used to compare both the networks is shown in Table 5.

On Comparing the Confusion Matrices of both the models in Figs. 11 and 14, there has been a significant reduction in the false negatives in the case of COVID-19 class in proposed Conv-CapsNet.

For binary class classification using proposed Conv-CapsNet, training and validation accuracy and loss curves are shown in Fig. 15. Comparative class wise accuracy of original capsule network and Conv-CapsNet is shown in Fig. 16. Here, accuracy for COVID-19 class significantly improved from 78.7% to 94.6%. Confusion matrices for both the models are shown in Figs. 17 and 18. Here, incorrectly predicted Non COVID-19 cases reduced from 21.3% to 2.67%. In case of binary-class classification, we obtained more than 94% accuracy in both the classes and achieved average accuracy of 97.69% as shown in Table 6. Precision, Recall and F1 measures are also shown in the table and PR curve is plotted in Fig. 19 which shows that both the classes have good performance and Non COVID-19 class

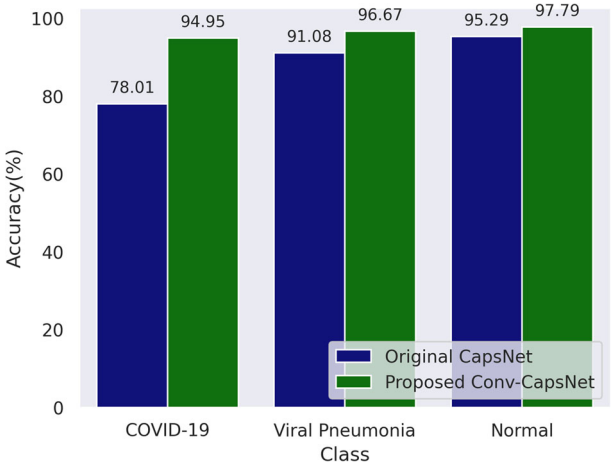


Fig. 13 Class wise accuracy of original capsule network and proposed Conv-CapsNet for multiclass classification

Table 5 Different parameters of Conv-CapsNet and original capsule network

Parameter	Original capsule network	Conv-CapsNet
Image Size	28 × 28	150 × 150
Learning Rate	0.001	0.0005
Batch Size	64	64
Epochs	20	50
Loss Weights	0.392	5.392
Routings	3	3
Optimizer	Adam	Adam
F1-Score	90.24%	96.47%
COVID-19 Class Accuracy	78.01%	93.08%

performs slightly better than COVID-19 class. Figure 20 shows the ROC-AUC curve for the binary class with AUC score of 0.99 for both the classes signifying that the proposed model is a good fit.

4.4 Comparative analysis with the literature

We performed a comparative analysis with the literature for binary and multi-class classification. Many authors worked on binary classification, where detection was mainly focused on COVID-19 and No-Findings classes. Whereas in multi-class classification, the Pneumonia class was also included as the 3rd class. Comparison with literature for the binary class is shown in Table 7. In binary classification, the majority of the authors obtained more than 95% accuracy, and our proposed approach obtained an excellent accuracy of 97.69%, which is comparable with the works [6, 35, 47]. These papers have utilized a deeper convolutional neural network with complex structures with more trainable parameters and performed data

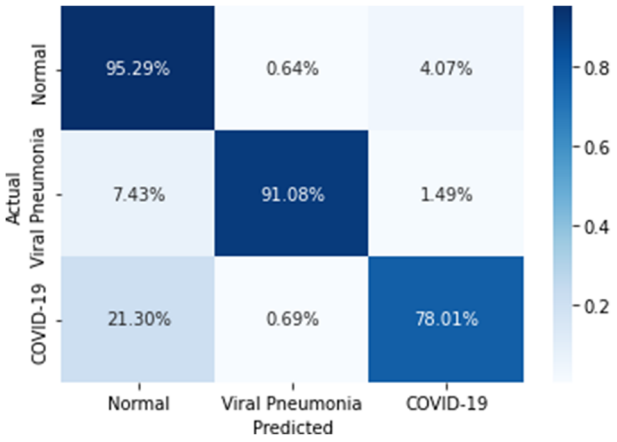


Fig. 14 Figure shows the confusion matrix of the original capsule network for multi-class classification

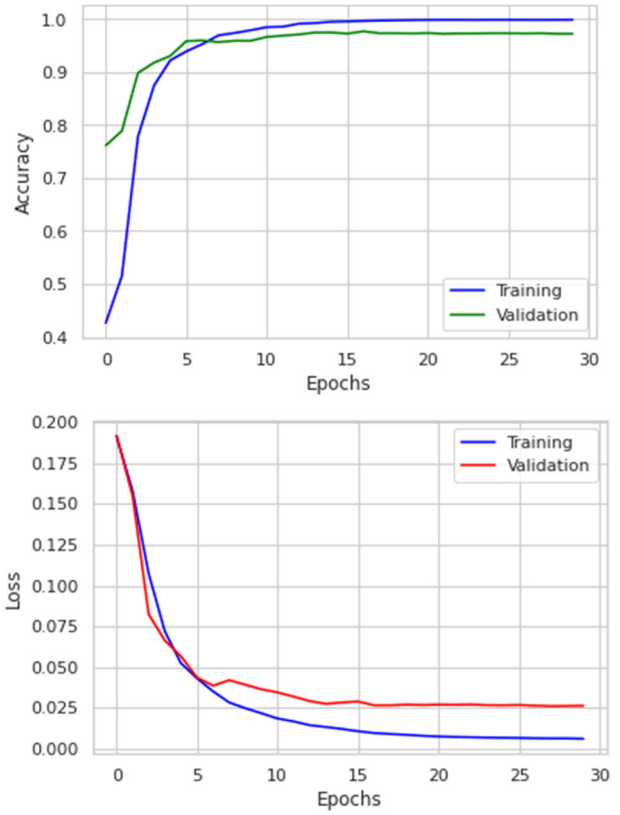


Fig. 15 Performance diagram of Conv-CapsNet on X-Ray dataset using $80-20\%$ split for binary class. Graph shows training accuracy vs validation accuracy and training loss vs validation loss

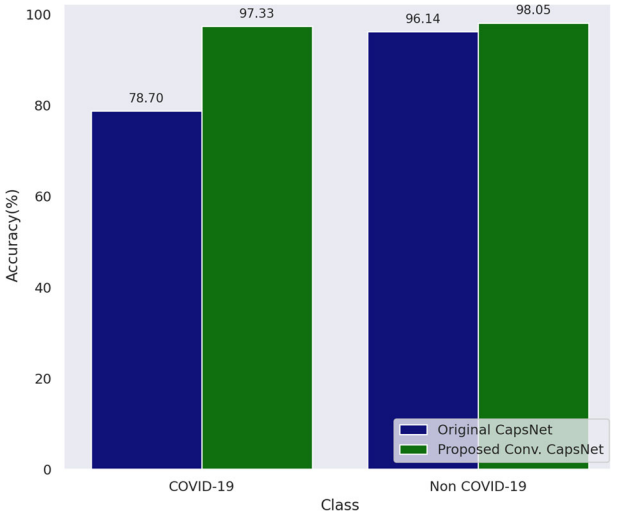


Fig. 16 Class-wise accuracy of proposed Conv-CapsNet for binary class

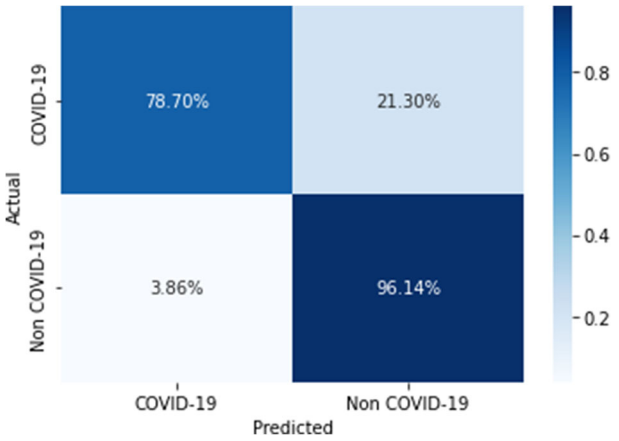


Fig. 17 Figure shows the confusion matrix of original capsule network for binary classification

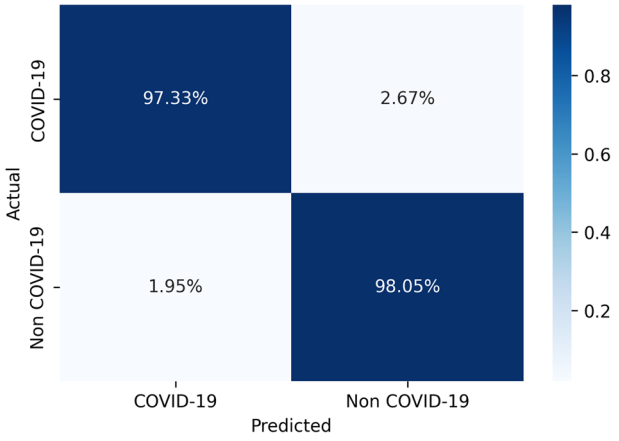


Fig. 18 Figure shows the confusion matrix of proposed convolutional capsule network for binary classification

Table 6 Precision, recall and F1-score of the two classes for binary classification

Class	Pre (%)	Rec (%)	F1 (%)
COVID-19	95.38	94.33	94.85
Non COVID-19	98.23	98.57	98.40
Accuracy			97.69
Macro Avg.	96.81	96.45	96.63
Weighted Avg.	97.69	97.69	97.69

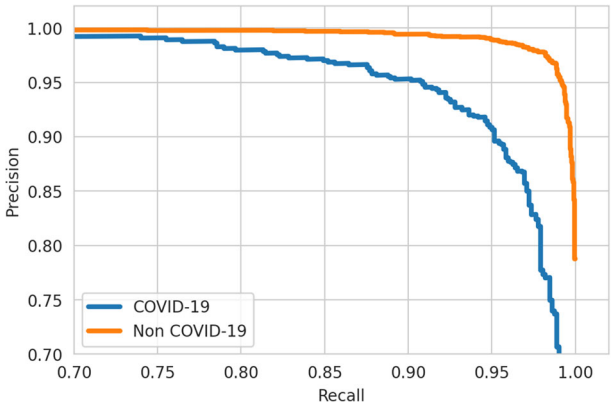


Fig. 19 Precision recall curve for COVID-19 and non COVID-19 classes

augmentation techniques to obtain more data samples. Our proposed model Conv-CapsNet has a shallow structure with fewer parameters and better accuracy, and it performed on fewer training samples. This makes it fast and robust, which is suitable for real-time applications.

Further, we performed a comparative analysis with literature for multi-class classification as shown in Table 8. The results clearly show that the proposed model achieved a good accuracy for the Radiography dataset. We performed a comparative analysis of various other models given in recent times with the proposed approach, and it gives an accuracy of 96.47% for the same dataset. The advantage of the proposed model, Conv-CapsNet, over other approaches is that it is a shallow network and takes less time for training and the size of the trained model is compact. The number of parameters required to train the model was around 23 million, which is less than the traditional CNN models for obtaining the same accuracy. The model provided accuracy comparable to other state-of-the-art architectures with less trainable parameters and without any data augmentation techniques.

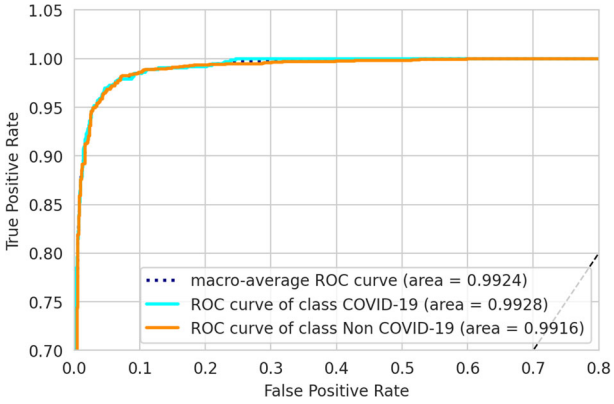


Fig. 20 ROC curves for Binary classes and their average with AUC scores

Table 7 Comparative analysis with state-of-the-art-methods on binary class classification

Paper	Model	Datasets	Accuracy
Saha et al. [46]	EMCNet (CNN + Ensemble of ML Classifiers)	4600 COVID-19, 2300-No Findings	98.91%
Ouchicha et al. [36]	CVDNet (Deep CNN)	219 COVID-19, 2686-No Findings	97.20%
Hussain et al. [23]	CoroDet	500 COVID 19, 800-No Findings	91.2%
Apostolopoulos et al. [6]	VGG-19	224 COVID-19, 504-No Findings	98.75%
Narin et al. [35]	ResNet50	50 COVID-19, 50 No-Findings	98.0%
Sethy et al. [47]	CNN + SVM	25 COVID-19, 25 No-Findings	95.3%
Alqudah et al. [5]	CNN	1525 COVID-19, 3050 Non-COVID-19	95.2
Dimeglio et al. [20]	DenseNet-121	15000 COVID-19 and Non-COVID-19	99%
Chakraborty et al. [14]	CNN	3797 COVID-19 and No-Findings	97.11%
Toraman et al. [50]	Convolutional CapsNet	1050COVID-19, 1050No-Findings	97.24%
Proposed Model	Conv-CapsNet	3616 COVID-19, 11537 Non COVID-19	97.69%

Table 8 Comparative analysis with state-of-the-art-methods on multi-class classification

Paper	Model	Datasets	Accuracy
Apostolopoulos et al. [6]	VGG19	224 COVID-19, 504-No Findings,700 Pneumonia	93.48%
Afshar et al. [2]	COVID-CAPS	COVID-19, Bacterial, Normal, Non-COVID-19 Viral	95.7
Ouchicha et al. [36]	CVDNet (Dee- CNN)	219 COVID-19, 51341-Normal, 1345 Pneumonia,	96.69%
Hussain et al. [23]	CoroDet	500 COVID-19, 800 Normal, 800 Pneumonia-Bacteria	94.2%)
Rahimzadeh et al. [40]	XCeptionNet, ResNetV2	180 COVID-19	80%
Loey et al. [30]	ResNet, AlexNet, GoogleNet	69 COVID-19	85.2%
Ucar et al. [39]	CNN	76COVID-19, 1583 Non-COVID-19, 4290 Pneumonia	98.3%
Dalvi et al. [19]	ResNet50 and Logistic Regression	1200 COVID-19, 1200 Non-COVID-19, 1200 Pneumonia	98.2%
Islam et al. [24]	CNN+ LSTM	1525 COVID-19, 1525 Non-COVID-19, 1525 Pneumonia	98.9%
Toraman et al. [50]	Convolutional CapsNet	1050COVID-19, 1050 No-Findings, 1050 Pneumonia	84.2%
Sharma et al. [48]	COVDC-Net	1784 COVID-19, 1755 Healthy, 1345 Pneumonia	96.48%
Wang et al. [57]	COVID-Net	53 COVID-19, 8066 Healthy, 5526 Pneumonia	93.3%
Law et al. [32]	VGG-16	1200 COVID-19, 1341 Healthy, 1345 Pneumonia	94%
Chandra et al. [15]	Ensemble Classifier	2346 COVID-19, 2346 Normal, 2346 Pneumonia	93.41%
Proposed Model	Conv-CapsNet	3616 COVID-19, 10192 Normal, 1345 Pneumonia	96.47%

5 Conclusion and future work

The study concludes that Capsule Networks have been effective in classifying the X-Ray scans compared to the Convolutional Neural Networks for smaller datasets. Our proposed Conv-CapsNet can achieve a 96.47% accuracy for diagnosing COVID-19 with Normal and Viral Pneumonia patients and 97.49% for binary class classifications. The disadvantage with the original Capsule Network over the proposed system is that it under-performed for the COVID-19 class, conveying that the model overfits the majority class. The proposed system is not advised for medical prescription but could help medical professionals and researchers for prognosis and research purposes.

Much further research has been done regarding Capsule Networks in healthcare as it is a relatively new model and has been tested on fewer datasets. Moreover, they suffer a disadvantage in handling complex datasets. With better feature map extraction using models such as VGG, RNN or EfficientNet, Capsule-based networks' performance could be tested. Various transfer and ensemble-based learning techniques are also being deployed for detection using X-Rays. CT scans are now available for testing. Existing models could be tested for more robust and accurate predictions for both scans.

Data Availability The data that support the findings of this study are available in at Kaggle at https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database, reference number [16, 42].

Declarations

Conflict of Interests We declare that we have no conflict of interest.

References

1. About worldometer (2020) COVID-19 data - Worldometer (9th Jul 2020), https://www.worldometers.info/coronavirus/
2. Afshar P, Heidarian S, Naderkhani F, Oikonomou A, Plataniotis KN, Mohammadi A (2020) Covid-caps: a capsule network-based framework for identification of covid-19 cases from x-ray images. Pattern Recogn Lett 138:638–643
3. Afshar P, Mohammadi A, Plataniotis KN (2018) Brain tumor type classification via capsule networks. In: 2018 25th IEEE international conference on image processing (ICIP). IEEE, pp 3129–3133
4. Al-Rakhami MS, Islam MM, Islam MZ, Asraf A, Sodhro AH, Ding W (2021) Diagnosis of covid-19 from x-rays using combined cnn-rnn architecture with transfer learning. MedRxiv, pp 2020–08
5. Alqudah AM, Qazan S, Alqudah A (2020) Automated systems for detection of covid-19 using chest x-ray images and lightweight convolutional neural networks
6. Apostolopoulos ID, Mpesiana TA (2020) Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks. Phys Eng Sci Med 43(2):635–640
7. Asraf A, Islam M, Haque M, Islam M (2020) Deep learning applications to combat novel coronavirus (covid-19) pandemic. SN Comput Sci 1:363
8. Avola D, Bacciu A, Cinque L, Fagioli A, Marini MR, Taiello R (2022) Study on transfer learning capabilities for pneumonia classification in chest-x-rays images. Comput Methods Programs Biomed 221:106833
9. Bassi PR, Attux R (2021) A deep convolutional neural network for covid-19 detection using chest x-rays. Res Biomed Eng:1–10
10. Bhatti U, Huang M, Wu D, Zhang Y, Mehmood A, Han H (2018) Recommendation system using feature extraction and pattern recognition in clinical care systems. Enterprise Inf Syst 13:1–23
11. Bhatti UA, Yu Z, Chanussot J, Zeeshan Z, Yuan L, Luo W, Nawaz SA, Bhatti MA, Ain QU, Mehmood A (2022) Local similarity-based spatial–spectral fusion hyperspectral image classification with deep cnn and gabor filtering. IEEE Trans Geosci Remote Sens 60:1–15
12. Bhatti U, Yuan L, Yu Z, Li J, Nawaz SA, Mehmood A, Zhang K (2020) Hybrid watermarking algorithm using clifford algebra with arnold scrambling and chaotic encryption. IEEE Access, pp 1–1

13. Bhatti U, Zeeshan Z, Nizamani MM, Bazai S, Yu Z, Yuan L (2021) Assessing the change of ambient air quality patterns in jiangsu province of China pre-to post-covid-19. Chemosphere 288:132569
14. Chakraborty S, Paul S, Hasan K (2022) A transfer learning-based approach with deep cnn for covid-19- and pneumonia-affected chest x-ray image classification. SN Comput Sci 3(1):1–10
15. Chandra TB, Verma K, Singh B, Jain D, Netam S (2020) Coronavirus disease (covid-19) detection in chest x-ray images using majority voting based classifier ensemble. Expert Syst Appl 165:113909
16. Chowdhury MEH, Rahman T, Khandakar A, Mazhar R, Kadir MA, Mahbub ZB, Islam KR, Khan MS, Iqbal A, Emadi NA, Reaz MBI, Islam MT (2020) Can ai help in screening viral and covid-19 pneumonia? IEEE Access 8:132665–132676
17. Covid-19 pneumonia - evolution over a week (2020) Case study, radiopaedia.org. https://doi.org/10.53347/rID-75264
18. Cozzi D, Albanesi M, Cavigli E, Moroni C, Bindi A, Luvarà S, Lucarini S, Busoni S, Mazzoni LN, Miele V (2020) Chest x-ray in new coronavirus disease 2019 (covid-19) infection: findings and correlation with clinical outcome. Radiol Med (Torino) 125:730–737
19. Dalvi J, Bohra A (2021) Covid-19 detection through deep feature extraction. arXiv:2111.10762
20. Dimeglio N, Romano S, Vesseron A, Pelegrin V, Ouchani S (2021) Covid-detect: a deep learning based approach to accelerate covid-19 detection. In: International conference on model and data engineering. Springer, pp 166–178
21. Gupta I, Garg N, Aggarwal A, Nepalia N, Verma B (2018) Real-time driver's drowsiness monitoring based on dynamically varying threshold. In: 2018 Eleventh international conference on contemporary computing (IC3). IEEE, pp 1–6
22. Haque IRI, Neubert J (2020) Deep learning approaches to biomedical image segmentation. Inf Med Unlocked 18:100297
23. Hussain E, Hasan M, Rahman MA, Lee I, Tamanna T, Parvez MZ (2021) Corodet: a deep learning based classification for covid-19 detection using chest x-ray images. Chaos Solitons, Fractals 142:110495
24. Islam MZ, Islam MM, Asraf A (2020) A combined deep cnn-lstm network for the detection of novel coronavirus (covid-19) using x-ray images. Inf Med Unlocked 20:100412
25. Islam M, Karray F, Alhajj R, Zeng J (2021) A review on deep learning techniques for the diagnosis of novel coronavirus (covid-19). IEEE Access 9:30551–30572
26. Islam M, Mahmud S, Jibril M, Islam M, Nooruddin S, Ayon S (2020) Wearable technology to assist the patients infected with novel coronavirus (covid-19). SN Comput Sci 1:320
27. Islam M, Ullah S, Mahmud S, Raju SMTU (2020) Breathing aid devices to support novel coronavirus (covid-19)infected patients. SN Comput Sci, vol 1
28. Ismael AM, Şengür A (2020) Deep learning approaches for COVID-19 detection based on chest x-ray images. Expert Syst Appl 164:114054
29. Karthik R, Menaka R, Hariharan M (2020) Learning distinctive filters for COVID-19 detection from chest x-ray using shuffled residual CNN. Appl Soft Comput 99:106744
30. Kim HW, Capaccione K, Li G, Luk L, Widemon RS, Rahman O, Beylergil V, Mitchell R, D'Souza BM, Leb JS et al (2020) The role of initial chest x-ray in triaging patients with suspected covid-19 during the pandemic. Emerg Radiol 27(6):617–621
31. LaLonde R, Bagci U (2018) Capsules for object segmentation. arXiv:1804.04241
32. Law BK, Lin LP (2021) Development of a deep learning model to classify x-ray of covid-19, normal and pneumonia-affected patients. In: 2021 IEEE international conference on signal and image processing applications (ICSIPA). IEEE, pp 1–6
33. Mousavi Z, Shahini N, Sheykhivand S, Mojtahedi S, Arshadi A (2022) Covid-19 detection using chest x-ray images based on a developed deep neural network. SLAS Technol 27(1):63–75. https://www.sciencedirect.com/science/article/pii/S247263032100011X
34. Muhammad LJ, Islam MM, Usman SS, Ayon SI (2020) Predictive data mining models for novel coronavirus (covid-19) infected patients' recovery. SN Comput Sci 1(4):206. https://doi.org/10.1007/s42979-020-00216-w
35. Narin A, Kaya C, Pamuk Z (2021) Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks. Pattern Anal Appl:1–14
36. Ouchicha C, Ammor O, Meknassi M (2020) Cvdnet: a novel deep learning architecture for detection of coronavirus (covid-19) from chest x-ray images. Chaos Solitons Fractals 140:110245
37. Patrick MK, Adekoya AF, Mighty AA, Edward BY (2022) Capsule networks–a survey. J King Saud Univ-comput Inf Sci 34(1):1295–1310
38. Qiao K, Zhang C, Wang L, Yan B, Chen J, Zeng L, Tong L (2018) Accurate reconstruction of image stimuli from human fmri based on the decoding model with capsule network architecture. arXiv:1801.00602

39. Rahimzadeh M, Attar A (2020) A new modified deep convolutional neural network for detecting covid-19 from x-ray images. p 2004
40. Rahimzadeh M, Attar A (2020) A modified deep convolutional neural network for detecting covid-19 and pneumonia from chest x-ray images based on the concatenation of xception and resnet50v2. Inf Med Unlocked 19:100360
41. Rahman MM, Islam MM, Manik MMH, Islam MR, Al-Rakhami MS (2021) Machine learning approaches for tackling novel coronavirus (covid-19) pandemic. SN Comput Sci 2(5):384. https://doi.org/10.1007/s42979-021-00774-7
42. Rahman T, Khandakar A, Qiblawey Y, Tahir A, Kiranyaz S, Kashem SBA, Islam MT, Maadeed SA, Zughaier SM, Khan MS, Chowdhury MEH (2020) Exploring the effect of image enhancement techniques on covid-19 detection using chest x-rays images. arXiv:2012.02238
43. Rahman MM, Manik MMH, Islam MM, Mahmud S, Kim JH (2020) An automated system to limit covid-19 using facial mask detection in smart city network. In: 2020 IEEE international IOT, electronics and mechatronics conference (IEMTRONICS). IEEE, pp 1–5
44. Rajasegaran J, Jayasundara V, Jayasekara S, Jayasekara H, Seneviratne S, Rodrigo R (2019) Deepcaps: going deeper with capsule networks. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 10725–10733
45. Sabour S, Frosst N, Hinton GE (2017) Dynamic routing between capsules. arXiv:1710.09829
46. Saha P, Sadi M, Islam M (2020) Emcnet: automated covid-19 diagnosis from x-ray images using convolutional neural network and ensemble of machine learning classifiers. Inf Med Unlocked 22:100505
47. Sethy PK, Behera SK (2020) Detection of coronavirus disease (covid-19) based on deep features
48. Sharma A, Singh K, Koundal D (2022) A novel fusion based convolutional neural network approach for classification of covid-19 from chest x-ray images. Biomed Signal Process Control, p 103778
49. Sheykhivand S, Mousavi Z, Mojtahedi S, Rezaii TY, Farzamnia A, Meshgini S, Saad I (2021) Developing an efficient deep neural network for automatic detection of covid-19 using chest x-ray images. Alexandria Eng J 60(3):2885–2903
50. Toraman S, Alakus TB, Turkoglu I (2020) Convolutional capsnet: a novel artificial neural network approach to detect covid-19 disease from x-ray images using capsule networks. Chaos Solitons Fractals 140:110122
51. Ucar F, Korkmaz D (2020) Covidiagnosis-net: deep bayes-squeezenet based diagnosis of the coronavirus disease 2019 (covid-19) from x-ray images. Med Hypotheses 140:109761
52. Uddin A, Talukder B, Monirujjaman Khan M, Zaguia A (2021) Study on convolutional neural network to detect covid-19 from chest x-rays. Math Probl Eng, vol 2021
53. Ullah S, Islam M, Mahmud S, Nooruddin S, Raju SMTU, Haque M (2021) Scalable telehealth services to combat novel coronavirus (covid-19) pandemic. SN Comput Sci, vol 2
54. Verma B, Choudhary A (2018) A framework for driver emotion recognition using deep learning and grassmann manifolds. In: 2018 21st International conference on intelligent transportation systems (ITSC). IEEE, pp 1421–1426
55. Verma B, Choudhary A (2018) Deep learning based real-time driver emotion monitoring. In: 2018 IEEE international conference on vehicular electronics and safety (ICVES). IEEE, pp 1–6
56. Vesperini F, Gabrielli L, Principi E, Squartini S (2019) Polyphonic sound event detection by using capsule neural networks. IEEE J Sel Top Signal Process 13(2):310–322
57. Wang L, Lin ZQ, Wong A (2020) Covid-net: a tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images. Sci Reports 10(1):1–12
58. Xi E, Bing S, Jin Y (2017) Capsule network performance on complex data. arXiv:1712.03480
59. Yamashita R, Nishio M, Do RKG, Togashi K (2018) Convolutional neural networks: an overview and application in radiology. Insights Imaging 9(4):611–629
60. Zhao W, Ye J, Yang M, Lei Z, Zhang S, Zhao Z (2018) Investigating capsule networks with dynamic routing for text classification. arXiv:1804.00538
61. Zoabi Y, Deri-Rozov S, Shomron N (2021) Machine learning-based prediction of covid-19 diagnosis based on symptoms. npj Digital Med 4(1):1–5

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.

EGU23-39
https://doi.org/10.5194/egusphere-egu23-39
EGU General Assembly 2023
© Author(s) 2023. This work is distributed under
the Creative Commons Attribution 4.0 License.

Deep Convolutional Architectures for Uncertainty Quantification and Forecast in Inundation Problems

Azzeddine Soulaïmani[1], Azzedine Abdedou[1], Yash Kumar[2], and Pratyush Bhatt[2]
[1]Ecole de technologie supérieure, Mechanical Engineering, Canada
[2]Delhi Technological University, Department of mechanical engineering, India

Most science and engineering problems are modeled by time-dependent and parametrized nonlinear partial differential equations. Their resolution with traditional computational methods may be too expensive, especially in the context of predictions with uncertainty quantification or optimization, to allow for rapid predictions. In this talk, we will overview data-driven methods aimed at representing high-fidelity computational models by means of reduced-dimension surrogate ones. Different approaches will be presented for the uncertainty quantification for reliable predictions and forecasts in inundation problems.

Particularly, a non-intrusive reduced-order model based on convolutional autoencoders is proposed as a data-driven tool to build an efficient nonlinear reduced-order model for stochastic spatiotemporal large-scale physical problems. The method uses two-level autoencoders to reduce the spatial and temporal dimensions from a set of high-fidelity snapshots collected from an in-house high-fidelity numerical solver of the shallow-water equations. The encoded latent vectors, generated from two compression levels, are then mapped to the input parameters using a regression-based multilayer perceptron. The accuracy of the proposed approach is compared to the linear reduced-order technique-based artificial neural network (POD-ANN) on benchmark tests (the Burgers and Stoker's solutions) and a hypothetical dam-break flow problem over a complex bathymetry river. The numerical results show that the proposed nonlinear framework presents strong predictive abilities to accurately approximate the statistical moments of the outputs for complex stochastic large-scale and time-dependent problems, with low computational cost during the predictive online stage.

The caveat that remains is the long-term temporal extrapolation for problems marked by sharp gradients and discontinuities. Our study explores forecasting convolutional architectures (LSTM, TCN, and CNN) to obtain accurate solutions for time-steps distant from the training domain, on advection-dominated test cases. A simple convolutional architecture is then proposed and shown to provide accurate results for the forecasts. To evaluate the epistemic uncertainties in the solutions, the methodology of deep ensembles is adopted.

REFERENCES

- Bhatt, Y. Kumar and A. Soulaïmani. Deep Convolutional Architectures for Extrapolative Forecast

in Time-dependent Flow Problems, DOI: 10.48550/arXiv.2209.09651.

- Abdedou and A. Soulaïmani. Reduced-order modeling for stochastic large-scale and time-dependent problems using deep spatial and temporal convolutional autoencoders. arXiv:2208.03190[physics.flu-dyn].
- Jacquier, A. Abdedou, V. Delmas and A. Soulaimani. Non-intrusive reduced-order modeling using uncertainty-aware Deep Neural Networks and Proper Orthogonal Decomposition: Application to flood modeling. Journal of Computational Physics. Volume 424, 1 January 2021, 109854.
- Abdedou and A. Soulaïmani. A non-intrusive reduced-order modeling for uncertainty propagation of time-dependent problems using a B-splines Bézier elements-based method and proper orthogonal decomposition: Application to dam-break flows. Computers & Mathematics with Applications. Volume 102, 15 November 2021, Pages 187-205.
- Chaudhry and A. Soulaimani. A Comparative Study of Machine Learning Methods for Computational Modeling of the Selective Laser Melting Additive Manufacturing Process. Appl. Sci. 2022, 12(5), 2324; https://doi.org/10.3390/app12052324.
- Delmas and A. Soulaimani. Parallel high-order resolution of the Shallow-water equations on real large-scale meshes with complex bathymetries. Journal of Computational Physics. Volume 471, 15 December 2022, 111629

Journal of Physics: Conference Series

PAPER • OPEN ACCESS

Dependence of Matrix Elements on Hydrostatic Pressure and Temperature

To cite this article: Mona *et al* 2023 *J. Phys.: Conf. Ser.* **2426** 012058

View the article online for updates and enhancements.

You may also like

- Charge transport dynamics and the effects on electrical tree degradation under DC voltages in thermally aged silicone rubber
Yunxiao Zhang, Yuanxiang Zhou, Chao Wu et al.
- Effect of load on the friction-wear behavior of magnetron sputtered DLC film at high temperature
Sun Ze and Kong Dejun
- Traffic flow of connected and automated vehicles at lane drop on two-lane highway: An optimization-based control algorithm versus a heuristic rules-based algorithm
Huaqing Liu, , Rui Jiang et al.

ECS 244th Electrochemical Society Meeting

October 8 – 12, 2023 • Gothenburg, Sweden

50 symposia in electrochemistry & solid state science

Abstract submission deadline:
April 7, 2023

Read the call for papers & **submit your abstract!**

This content was downloaded from IP address 14.139.251.98 on 01/03/2023 at 08:56

Dependence of Matrix Elements on Hydrostatic Pressure and Temperature

Mona[1], Aakash[2], Priyanka and Rinku Sharma

Department of Applied Physics, Delhi Technological University, Delhi, 110042, India

Email- monadevi_2k20mscphy16@dtu.ac.in, akash_2k20mscphy01@dtu.ac.in

Abstract - The matrix elements of a 2 – D electron gas restricted in a GaAs cylindrical quantum wire including Rashba Spin Orbit Interaction, under simultaneous effect of hydrostatic pressure and temperature is investigated. The strong dependency of hydrostatic pressure and absolute temperature, on energy band gap and effective mass of charge carriers of the quantum wire, have been found. Because of this, the linear and non-linear properties of the quantum wire also get affected. We also observed that under the influence of electric and mamagnetic field, the matrix elements get affected strongly with the change in Rashba spin orbit interaction and Magnetic field.

Keywords – 2-D = Two Dimensional, GaAs = Gallium Arsenide, SOI = Spin Orbit Interaction.

1. Introduction

Recently, the low dimensional semiconductor structures are attracting much attention e.g., quantum dot, wire, well etc. because of their utilization in the area of optoelectronics [1]–[4]. Remarkable physical properties have been shown by the quantum wires of all these structures, in observing materiality in future automations [5]. Due to such discrete characteristics, scientists have been paid much attention to the nanostructures, as they provide a broad area for research to be aware of them [6]. Quantum wire which is a one-dimensional structure having the motion of electrons restricted in two directions, quantum effects get more obvious in comparison to the two-dimensional quantum structures [7].

Great interest has been shown in the learnings of spin-dependent phenomena of quantum restrained structures due to their essential part in upcoming electronic devices with low power consumption, high speed and a high degree of performance [1], [5], [8]–[10]. These devices are based on the spin degree of freedom which is helpfull for information processing including electron charge. The work of these proposed devices is manipulating the electron spin over spin orbit interaction. Rashba spin orbit interaction has convenient benefits as it rely upon the electronic conditions of the quantum heterostructures and alternating the gate voltage exert influence on the strength of Rashba SOI, for managing the final spin orientation of electrons [11], [12] In this field, Datta-Das spin field effect transistor is one out of the innovative gadget idea. Many theoretical and experimental inspections have been done on the influence of Rashba SOI on the optical affections of nanostructures. However, from the experimental and theoretical viewpoints, the real existence of Rashba term in asymmetric quantum wires is still debatable. From the experimental standpoint, the trouble to discover the Rashba spin orbit coupling constant α, turn out the debate [13]–[16].

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

Optical properties possessed by quantum nanostructures, can be freely manipulated by apparently confined parameters that are hydrostatic pressure and temperature [17]–[20]. The optical matrix elements of the quantum wire structures depend on the direction of crystallographic wire and an optical anisotropy can appear there which will be in a perpendicular plane to the wire [21]. Recently, many researchers have investigated that in case of low-dimensional structures, the optical and electronic properties get influenced by hydrostatic pressure and temperature with varying confinement potential [22]–[24]. The optical properties of exciton with hydrostatic pressure was demonstrated by Duque et al., Ungan et al. look over the effects of hydrostatic pressure in quantum well on the intersubband transitions [25]. Rezaei et al. reported that the linear and nonlinear properties get influenced extensively in case of quantum dot with the application of electric and magnetic fields [26]. In disc shaped quantum dot, the consequence of hydrostatic pressure and temperature was investigated by Liang and Xie [27]. Mughnetsyan et al. have demonstrated the integrated outcome of hydrostatic pressure and temperature in case of quantum ring. In low-dimensional structures, the impact of so many parameters on the optical properties have been explored by a number of researchers, however, according to all the information till now, no one discussed the upshot of pressure, temperature and external magnetic field with SOI on the optical properties in a parabolic confinement of quantum wire. For quantum confined structures, the knowledge of Rashba SOI on the optical reciprocation is interesting and is going to be advantageous for many future applications namely quantum information systems and opt- spintronics devices [28], [29]. Therefore, to learn about the optical properties under the effect of hydrostatic pressure and temperature with Rashba SOI in case of quantum wire is going to be fascinating.

In this conversation, we have used the density matrix formalism to compute the optical matrix elements for an electron applying electric and magnetic field both at the same time, under the domination of hydrostatic pressure and temperature for a cylindrical quantum wire. From the results, we observed that matrix elements vary linearly with the pressure and inversely with temperature. We also detected that the peaks of absorption coefficient also behave in the same manner as the matrix elements, with the variation of hydrostatic pressure and temperature. In section III, the theory part is given, the result and conclusion of the conversation is mentioned in section IV and V.

2. **Theoretical framework**

Let us assume, a 2-D GaAs quantum wire, with parabolic confined potential, given by:

$$V(x,y) = \frac{1}{2}m^*[P,T]\omega_0^2(x^2+y^2) \qquad (1)$$

where $m^*[P,T]$ is the effective mass of electron under the impact of hydrostatic pressure and temperature. The Hamiltonian for a single electron in terms of mass can be given as

$$H_0(x,y)=\frac{1}{2m^*[P,T]}(p+eA)^2+\frac{1}{2}m^*[P,T]\omega_0^2(x^2+y^2) \qquad (2)$$

When we consider the spin, two more terms will add up to the Hamiltonian in equation (2), where the 1^{st} term is the Zeeman term and the 2nd term is the spin orbit interaction term as:

$$H_{spin} = \frac{1}{2}g\mu_B B\sigma + H_{soi} \qquad (3)$$

Where g represents the effective Lande factor for semiconductor. H_{soi} represents the spin orbit interaction term which consists of two terms as H_{soi}^R which is for Rashba SOI and H_{soi}^D which is for

Dresselhaus spin orbit interaction. Rashba term dominates over the Dresselhaus term that is why, here we will consider only the Rashba term, which is:

$$H_{soi}^{R} = \frac{\alpha}{\hbar}[\sigma * (p + eA)]_z \qquad (4)$$

Where α is the Rashba coupling coefficient which can be restrained by applying the varying gate voltage in z-direction. Thus, the total Hamiltonian will be the combination of both the terms as H_0 and H_{soi}^{R} in the effective mass approximation and will be given as

$$H = \frac{p^2}{2m^*[P,T]} + \frac{e}{m^*[P,T]}A.p + \frac{e^2A^2}{2m^*[P,T]} + \frac{1}{2}m^*[P,T]\omega_0^2(x^2+y^2) + \frac{1}{2}g\mu_B B\omega + \frac{\alpha}{\hbar}(\sigma * p)_z + \frac{e\alpha}{\hbar}(\sigma * A)_Z \quad (5)$$

Now, the Hamiltonian gets transformed into the polar co-ordinates, so the eigen energy comes out as:

$$E_{n\sigma} = \hbar\omega\left(n + \frac{1}{2}\right) - \frac{e^2E^2}{2m^*(P,T)\omega^2} + \frac{\omega_0^2\hbar^2k_y^2}{\omega^2 2m*(P,T)} - \frac{e^2EB\hbar k_Y}{m*(P,T)^2\omega^2} + \frac{1}{2}g\mu_B\sigma B \quad (6)$$

Energy scale correlating with the confining potential and Rashba SOI are

Ep= $\hbar\omega_0$ and Δ_{so}=$\frac{m(P,T)\alpha^2}{2\hbar^2}$ respectively. Expanding φ(x), concerning

$\psi_{n\sigma}$(x) as φ(x) = $\Sigma_{n\sigma}a_{n\sigma}\psi_{n\sigma}$(x)

And using the condition of orthogonality, we have calculated the matrix elements $\left(M_{if} = |\langle\Phi_i|x,y|\Phi_j\rangle|\right)$ of a quantum wire by using the density matrix formalism.

3. **Result**

To calculate the eigenfunctions and eigenvalues for studying the purpose of hydrostatic pressure and temperature with SOI, we have done the numerical calculations in this project. To look over the optical properties, we have used the calculated results. For the calculation of optical matrix elements of the GaAs quantum wire, we have taken the successive range of parameters as: magnetic field, $B = 1T - 2.5T$ and Rashba coupling coefficient, $\alpha = 0.51nmeV - 0.57nmeV$. Recently, J. R. Bindal et al., experimentally reported the higher values of Rashba SOI [30]. For the computation, we have taken the material parameters of GaAs, for that an electron has the effective mass $0.03m_0$, where m_0 represents the rest mass of the electron.

In Fig.1, we showed the variation of matrix elements with hydrostatic pressure and the alteration of effective mass as a function of hydrostatic pressure for $P = 0kbar - 100kbar$. It is detected that the optical matrix elements go on increasing on increasing the pressure. From eq. (6), we can observe the dependence of energy band gap on hydrostatic pressure. This linear dependency of $E_{n\sigma}$ on hydrostatic pressure bring about the inflection in the graph among matrix elements and pressure as shown in Fig.1. With the increase in pressure, energy band gap also increases linearly and due to this reason, the effective mass of the electron also found to be increasing.

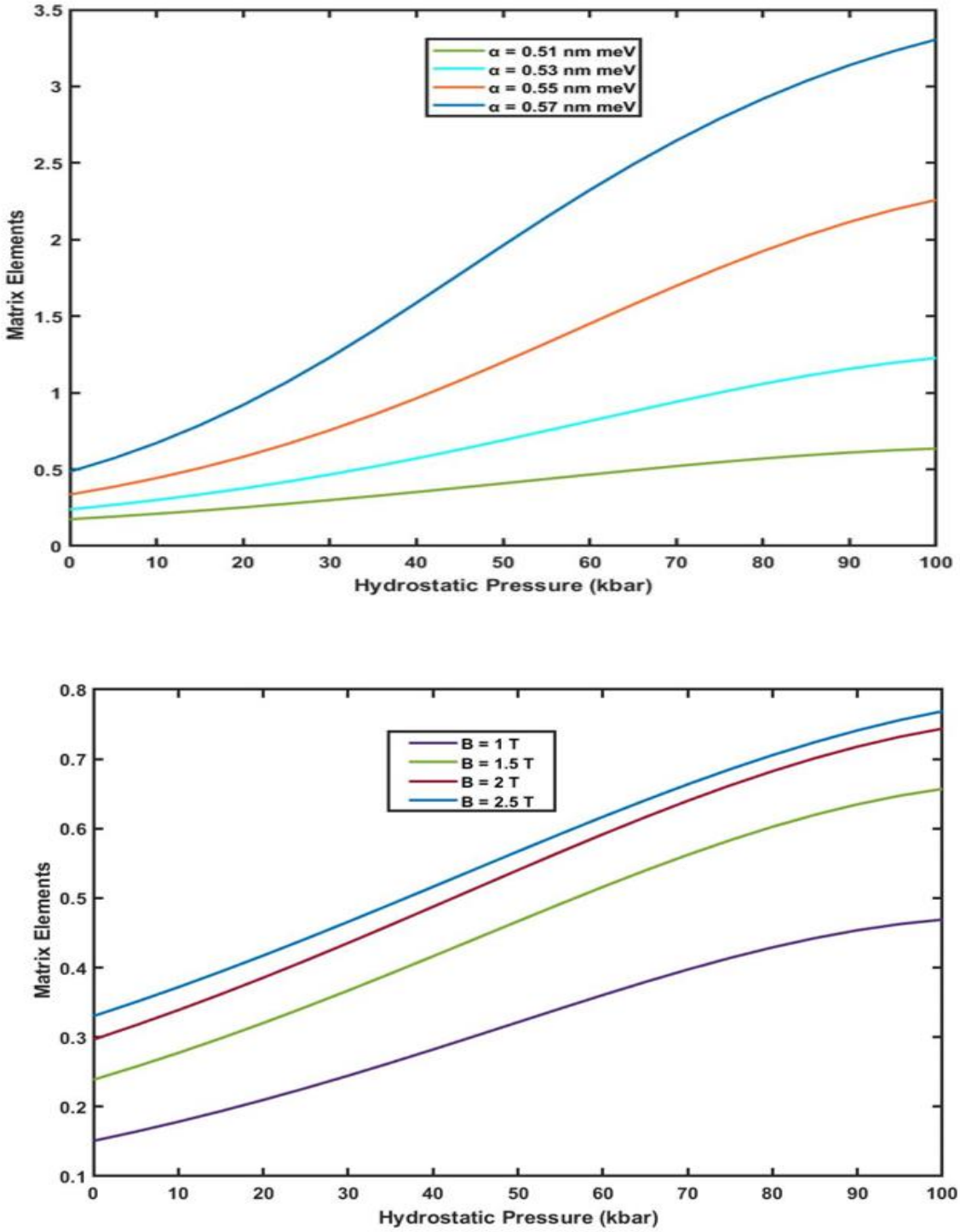


Fig.1. The matrix element variation with hydrostatic pressure at 'zero' external electric field (E) for: (a) B = 1T, ∝ = 0.51nm eV to 0.57 nm eV (b) ∝ = 0.51 nm eV, B = 1 to 2.5 T.

In Fig.1(a), we have taken a constant value of magnetic field, $B = 1T$ and varied the value of Rashba coupling coefficient, $\alpha = 0.51$nm eV-0.57nm eV and observed the linear dependency of matrix elements on the hydrostatic pressure. Then in Fig.1(b), we have varied the value of magnetic field, B=1T-2.5T and taken the constant value of the Rshba coupling coefficient, $\alpha = 0.51nm\ eV$.

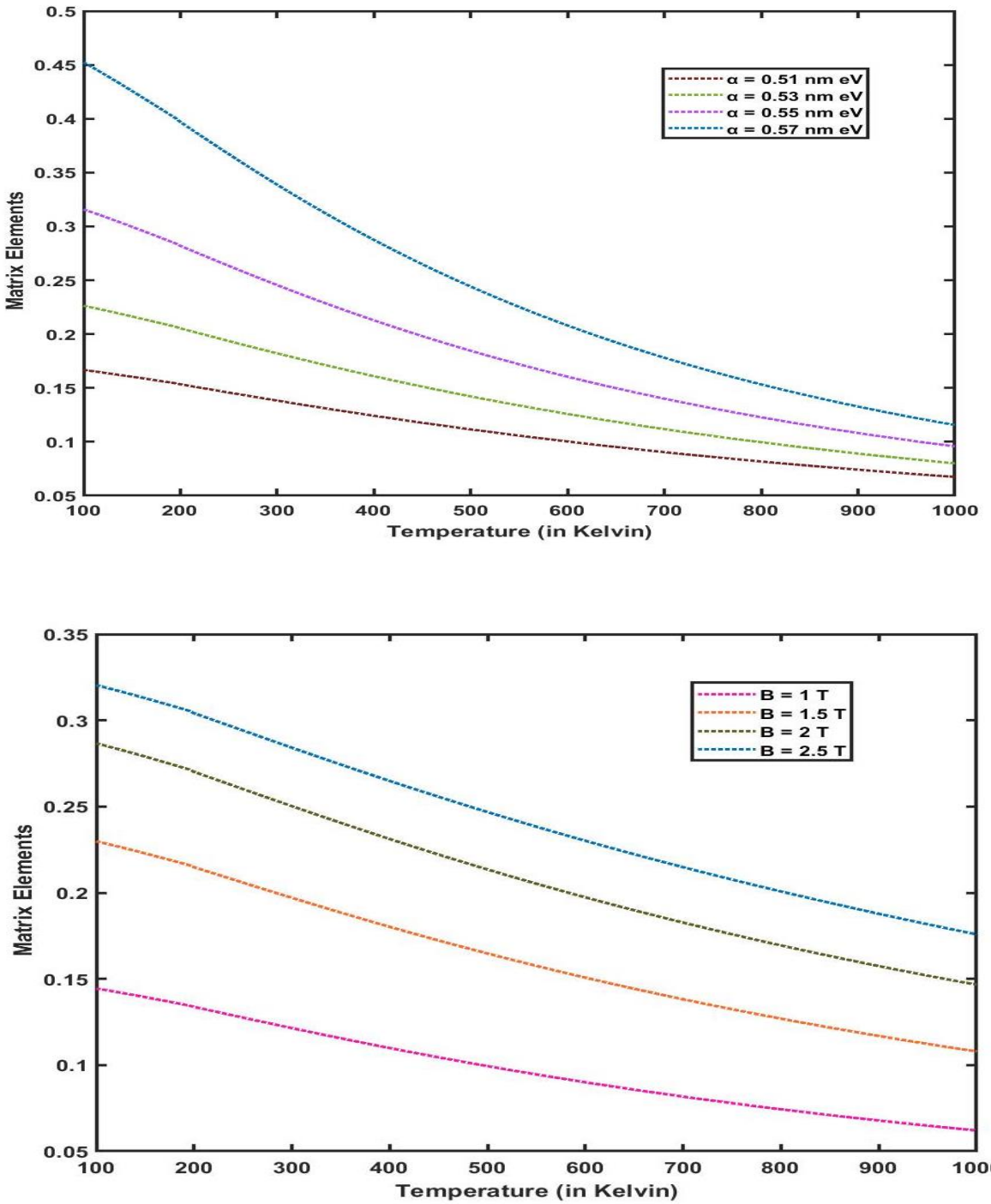


Fig. 2. The matrix element variation with temperature at zero external electric field (E) for: (a) B = 0T, $\propto$ = 0.51 to 0.57 nm eV (b) $\propto$ = 0.51 nm eV, B = 1 to 2.5 T.

In Fig.2, we showed the fluctuation of matrix elements with temperature and the alteration of effective mass as a function of temperature for $T = 100K - 1000K$. We observed that with increasing the temperature, the matrix elements found to be decreasing, it is also observed that on increasing the temperature, the band gap will decrease and because of this, the effective mass of the electron diminish.

In Fig.2(a), we have taken a constant value of magnetic field, $B = 1T$ and varied the value of Rashba coupling coefficient, $\alpha = 0.51$nm eV-0.57nm eV and observed the inverse dependence of matrix elements on the temperature. On increasing the Rashba SOI factor, the transition energy is also found to increase. Then in Fig.2(b), we have varied the value of magnetic field, B=1T-2.5T and taken the constant value of the Rshba coupling coefficient, $\alpha = 0.51nm\ eV$.

4. Conclusion

We presented the detailed inspection of the outcome of hydrostatic pressure and temperature with SOI on the linear and non-linear optical traits, in case of a quantum wire. We have shown the results throughout the effective mass approximation using density matrix formalism. Before that, we have also look over the impact of electric and magnetic field with Rashba spin orbit interaction in a quantum wire. We have also calculated the energy dispersion for a quantum wire.

We have found that the matrix elements depend linearly on pressure and inversely on temperature. Also, we observed that the energy band gap also varies linearly with pressure and inversely with temperature. Blue shifting in the peaks of absorption coefficient is detected with the increase in pressure and if the temperature is increased, a red shift in the peaks of absorption coefficient is observed. The transition energy is also found to be in linear dependence with the Rashba SOI.

To sum up, the linear and non-linear properties of quantum wires are greatly influenced by the hydrostatic pressure and temperature, as well as the magnetic field with Rashba SOI. Hence, hydrostatic pressure and temperature changed the energy eigenstates and eigenvalues, as done by all the other parameters. By adjusting and regulating these parameters, we can tune and detune the THz laser in the low-dimensional structures so these characteristics would be extremely beneficial in the creation of tunable optoelectronic devices.

5. References

[1] Antil, S., Kumar, M., Lahon, S., Dahiya, S., Ohlan, A., Punia, R., & Maan, A. S. (2019). Influence of hydrostatic pressure and spin orbit interaction on optical properties in quantum wire. *Physica B: Condensed Matter*, *552*, 202-208.

[2] Baghramyan, H. M., Barseghyan, M. G., Kirakosyan, A. A., Restrepo, R. L., Mora-Ramos, M. E., & Duque, C. A. (2014). Donor impurity-related linear and nonlinear optical absorption coefficients in GaAs/Ga1− xAlxAs concentric double quantum rings: Effects of geometry, hydrostatic pressure, and aluminum concentration. *Journal of luminescence*, *145*, 676-683.

[3] Bandyopadhyay, S. (2012). *Physics of nanostructured solid-state devices*. Springer Science & Business Media.

[4] Dakhlaoui, H. (2015). Linear and nonlinear optical absorption coefficients and refractive index changes in GaN/AlxGa (1− x) N double quantum wells operating at 1.55 μ m. *Journal of Applied Physics*, *117*(13), 135705.

[5] Csavina, J., Field, J., Félix, O., Corral-Avitia, A. Y., Sáez, A. E., & Betterton, E. A. (2014). Effect of wind speed and relative humidity on atmospheric dust concentrations in semi-arid climates. *Science of the Total Environment*, *487*, 82-90.

[6] Dahiya, S., Lahon, S., & Sharma, R. (2020). Effects of temperature and hydrostatic pressure on the optical rectification associated with the excitonic system in a semi-parabolic quantum dot. *Physica E: Low-dimensional Systems and Nanostructures*, *118*, 113918.

[7] Gou, B. P., & Kong, X. J. (2005). Electron energy levels for a finite rectangular quantum wire in a transverse magnetic field. *Journal of applied physics*, *98*(5), 053710.

[8] Žutić, I., Fabian, J., & Sarma, S. D. (2004). Spintronics: Fundamentals and applications. *Reviews of modern physics*, *76*(2), 323.

[9] Schliemann, J., Egues, J. C., & Loss, D. (2003). Nonballistic spin-field-effect transistor. *Physical review letters*, *90*(14), 146801.

[10] Wolf, S. A., Awschalom, D. D., Buhrman, R. A., Daughton, J. M., von Molnár, V. S., Roukes, M. L., ... & Treger, D. M. (2001). Spintronics: a spin-based electronics vision for the future. *science*, *294*(5546), 1488-1495.

[11] Antil, S., Kumar, M., & Jha, P. K. Spin effect on energy levels of quantum wire in presence of electric and magnetic field. In *Spin* (Vol. 2250, p. 2734).

[12] Dresselhaus, G. (1955). Spin-orbit coupling effects in zinc blende structures. *Physical Review*, *100*(2), 580.

[13] Dayem, A. H., & Martin, R. J. (1962). Quantum interaction of microwave radiation with tunneling between superconductors. *Physical Review Letters*, *8*(6), 246.

[14] Därr, A., Kotthaus, J. P., Ando, T., & Fumi, F. G. (1976). Proceedings of the 13th International Conference on the Physics of Semiconductors.

[15] Rowe, A. C. H., Nehls, J., Stradling, R. A., & Ferguson, R. S. (2001). Origin of beat patterns in the quantum magnetoresistance of gated InAs/GaSb and InAs/AlSb quantum wells. *Physical Review B*, *63*(20), 201307.

[16] J. Ohkawa, F., & Uemura, Y. (1974). Quantized surface states of a narrow-gap semiconductor. *Journal of the Physical Society of Japan*, *37*(5), 1325-1333.

[17] Lu, L., Xie, W., & Shu, Z. (2011). Combined effects of hydrostatic pressure and temperature on nonlinear properties of an exciton in a spherical quantum dot under the applied electric field. *Physica B: Condensed Matter*, *406*(19), 3735-3740.

[18] Zeng, Z., Garoufalis, C. S., Baskoutas, S., & Bester, G. (2013). Electronic and optical properties of ZnO quantum dots under hydrostatic pressure. *Physical Review B*, *87*(12), 125302.

[19] Ungan, F., Restrepo, R. L., Mora-Ramos, M. E., Morales, A. L., & Duque, C. A. (2014). Intersubband optical absorption coefficients and refractive index changes in a graded quantum well under intense laser field: effects of hydrostatic pressure, temperature and electric field. *Physica B: Condensed Matter*, *434*, 26-31.

[20] Ozturk, E., & Sokmen, I. (2013). Nonlinear intersubband absorption and refractive index changes in square and graded quantum well-modulated by temperature and hydrostatic pressure. *Journal of luminescence*, *134*, 42-48.

[21] Yamaguchi, A. A., & Usui, A. (1996). Theoretical gain of [hhk]-oriented quantum wire lasers. *Journal of applied physics*, *79*(6), 3340-3342.

[22] Schönhöbel, A. M., Girón-Sedas, J. A., & Porras-Montenegro, N. (2014). Quasistationary states in single and double GaAs – (Ga, Al) As quantum wells: Applied electric field and hydrostatic pressure effects. *Physica B: Condensed Matter*, *442*, 74-80.

[23] Rezaei, G., & Kish, S. S. (2013). Linear and nonlinear optical properties of a hydrogenic impurity confined in a two-dimensional quantum dot: effects of hydrostatic pressure, external electric and magnetic fields. *Superlattices and Microstructures*, *53*, 99-112.

[24] Manaselyan, A. K., Barseghyan, M. G., Kirakosyan, A. A., Laroze, D., & Duque, C. A. (2014). Effects of applied lateral electric field and hydrostatic pressure on the intraband optical transitions in a GaAs/Ga1− xAlxAs quantum ring. *Physica E: Low-dimensional Systems and Nanostructures*, *60*, 95-99.

[25] Ungan, F., Restrepo, R. L., Mora-Ramos, M. E., Morales, A. L., & Duque, C. A. (2014). Intersubband optical absorption coefficients and refractive index changes in a graded quantum well under intense laser field: effects of hydrostatic pressure, temperature and electric field. *Physica B: Condensed Matter*, *434*, 26-31.

[26] Rezaei, G., & Kish, S. S. (2013). Linear and nonlinear optical properties of a hydrogenic impurity confined in a two-dimensional quantum dot: effects of hydrostatic pressure, external electric and magnetic fields. *Superlattices and Microstructures*, *53*, 99-112.

[27] Liang, S., & Xie, W. (2011). Effects of the hydrostatic pressure and temperature on optical properties of a hydrogenic impurity in the disc-shaped quantum dot. *Physica B: Condensed Matter*, *406*(11), 2224-2230.

[28] Jungwirth, T., Wunderlich, J., & Olejnik, K. Nature Mater. 11, 382 (2012).

[29] Dai, J. H., Lee, J. H., Lin, Y. L., & Lee, S. C. (2008). In (Ga) As quantum rings for terahertz detectors. *Japanese journal of applied physics*, *47*(4S), 2924.

[30] Bindel, J. R., Pezzotta, M., Ulrich, J., Liebmann, M., Sherman, E. Y., & Morgenstern, M. (2016). Probing variations of the Rashba spin–orbit coupling at the nanometre scale. *Nature Physics*, *12*(10), 920-925.

energies

MDPI

Article

Design and Evaluation of a Water-Based, Semitransparent Photovoltaic Thermal Trombe Wall

Sheel Bhadra [1,†], Niloy Sen [2,†], Akshay K K [3], Harmeet Singh [4] and Paul G. O'Brien [4,*]

[1] Department of Mechanical Engineering, Delhi Technological University, Delhi 110042, India
[2] Department of Mechanical Engineering, Jadavpur University, Kolkata 700032, India
[3] Department of Mechanical Engineering, IIT Madras, Chennai 600036, India
[4] Department of Mechanical Engineering, Lassonde School of Engineering, York University, Toronto, ON M3J 2S5, Canada
* Correspondence: paul.obrien@lassonde.yorku.ca
† These two authors contributed equally to this paper.

Abstract: Trombe walls are a passive solar technology that can contribute to the reduction of building heating loads. However, during warmer weather conditions, Trombe walls may cause overheating. In this work, we investigate the feasibility of using Trombe walls to perform multiple functions during warm weather conditions including (1) heating and storing water for building applications, (2) providing occupants with visibility to the outdoors, and (3) generating electric power. Experiments are performed on a small-scale prototype comprising a clear water storage container with a transparent window and a tinted acrylic sheet that is immersed in the water. Photovoltaic cells are placed on the bottom half of the front face of the water storage container. Results show that water at the top of the clear container can be heated to temperatures as high as 45 °C when subjected to solar-simulated radiation for five hours. Numerical simulations predict that similar temperatures can be reached if the Trombe wall is scaled to full size. Furthermore, the cooler water at the bottom of the water storage container acts as a heatsink that reduces the extent to which the temperature of the PV cells is elevated. Results show the temperature and open circuit voltage of the PV cells are about 50 °C and 0.66 V, respectively, when water is present. However, when the water is absent from the container, the temperature of the PV cells increases up to 90 °C and their open circuit voltage drops to 0.60 V. The results show that water-based, semitransparent photovoltaic thermal Trombe walls have the potential to operate as multifunctional building envelopes that simultaneously provide for daylighting, heated water and electric power, and further research in this area is warranted.

Keywords: building energy; Trombe wall; thermal energy storage; building envelopes; building integrated photovoltaics

check for updates

Citation: Bhadra, S.; Sen, N.; K, A.K.; Singh, H.; O'Brien, P.G. Design and Evaluation of a Water-Based, Semitransparent Photovoltaic Thermal Trombe Wall. *Energies* **2023**, *16*, 1618. https://doi.org/10.3390/en16041618

Academic Editor: Angelo Zarrella

Received: 8 January 2023
Revised: 29 January 2023
Accepted: 2 February 2023
Published: 6 February 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Energy use in the building sector reached 132 EJ in 2021, which accounted for 30% of global energy consumption [1]. A significant portion of this energy demand is from the residential sector used for space and water heating purposes. Hence, there is an urgent requirement to shift to renewables to meet building energy demands. One promising source of clean energy is the sun, because it is widely distributed and can provide up to ~1000 W/m^2 per day [2].

As urban energy density continues to increase, it is expected that harnessing solar energy on building rooftops and facades will make a significant contribution toward reducing demands on energy infrastructures [3–6]. In particular, Trombe walls have been used for decades as a passive solar heating method because they can be highly efficient, relatively simple to construct and cost-effective [7–9]. Trombe walls typically consist of a large thermal wall, a clear glazed cover and an air channel between the thermal wall and the glazed cover [10]. Sunlight that passes through the glazed cover is absorbed by the

thermal wall. The absorbed sunlight generates heat, which can be stored in the thermal wall, which is made of materials with high thermal capacity such as brick, concrete, sand or stone. The generated heat can also be transferred to the air within the channel between the thermal wall and the outer glazing. Air vents are often located at the top and bottom of the thermal wall. When air within the channel is heated, it can be passively transported through the vent at the top of the thermal storage wall to the interior of the building, and colder air is drawn through the vent at the bottom of the thermal wall. Furthermore, solar thermal energy generated in Trombe walls during peak sunshine hours can be stored and used to heat the building after sunset.

Under favorable conditions Trombe walls can reduce the energy consumed in a building by 30% [8] and there has been much research and development toward improving the performance of Trombe walls. Mokni et al. found that a Trombe wall in Saudi Arabia could satisfy 80% of the heating needs of a room on a sunny day, or about 40% of the heating needs under conditions of high-wind or scattered clouds [11]. Yang et al. performed numerical simulations to show the addition of the phase change materials to Trombe walls results in a decrease in annual energy consumption of 18% [12]. Liu et al. showed that by using two air channels in a Trombe wall, the indoor air temperature could be increased by 5.6 °C as compared to the case when a traditional Trombe wall was used [13].

Although Trombe walls can reduce building energy loads, the design of effective Trombe walls is a challenging task that depends on climatic and seasonal factors [14]. During night or cloudy periods, the temperature of the thermal storage wall may drop below that of the indoor temperature, creating an inverse thermosiphon phenomenon whereby the airflow through the Trombe wall vents is reversed [15]. Moreover, if the thermal storage wall is too thick, then solar thermal energy stored in this wall may not be adequately transferred to the indoors. Another challenge, which we focus on in this research, is that during hot weather conditions, Trombe walls can cause overheating in well-insulated buildings [16,17]. Researchers have recommended using ventilation and solar shading to reduce overheating caused by Trombe walls [18]. Jaber and Ajib [19] recommended using roller shutters to prevent solar radiation from entering the building and insulation curtains between the glass and masonry wall layer to avoid heat transfer to the building. Building overhangs can also be designed to shade Trombe walls in hot climates during summer months [20]. However, these solutions for preventing overheating from Trombe walls do not fully utilize the available solar energy.

Recently, Mohamad et al. [21] designed a water-based Trombe wall wherein the water could be removed from the Trombe wall and used as a domestic water supply during hot weather conditions to prevent overheating. Singh et al. proposed semitransparent, water-based Trombe walls wherein a tinted acrylic sheet was inserted in a water-based storage wall to absorb solar radiation [22]. In this configuration, the Trombe wall could be designed to have the appearance of a tinted window, providing architects with aesthetic options for integrating Trombe walls into building facades. In this work, we investigate the potential of placing PV cells on the bottom portion of the front surface of the water storage wall within a semitransparent, water-based Trombe wall. In this configuration, the relatively cooler water at the bottom of the water storage wall acts as a heat sink to prevent the PV cells from overheating, which is important for achieving high efficiencies. Thus, the Trombe wall investigated in this work is designed to enhance the utility of the wall during hot weather conditions by producing both hot water and electric power.

Crystalline silicon photovoltaic (c-SiPV) modules dominate the market due to their low cost and high efficiency (~20%). However, one disadvantage of c-SiPV cells is their negative temperature coefficient of ~ 0.45%/°C, which implies that their efficiency decreases as their temperature increases [23–27]. The resulting decrease in efficiency can be significant, as the temperature of PV cells may be elevated by 20 °C or more. Thus, by reducing the temperature of PV cells during operation, their output can be increased. For example, Rahman et al. [26] experimentally investigated a PV module fitted with finned tubes for cooling and an efficiency improvement of up to 15.72% was observed.

Research on PV cells integrated in Trombe walls (referred to as PVTWs) has been reported in the literature. Hu et al. studied the performance of different PVTW configurations. They considered a Trombe wall module with blinds and PV cells (PVBTW), a Trombe wall configuration wherein the PV cells were fixed to the exterior glazing (PVGTW) and a third configuration where the PV cells were attached to the thermal storage wall (PVMTW). Considering both heat gains and electric power production, the PVBTW module performed 14.5% better than the PVGTW configuration and 14.1% better than the PVMTW configuration [28]. Koyunbaba et al. [29] presented a building integrated photovoltaic (BIPV) Trombe wall and their experimental results showed the daily average electrical and thermal efficiency of the BIPV Trombe wall system reached 4.52% and 27.2%, respectively. Jie et al. [30] studied a Trombe wall consisting of a PV glass panel with an area of 2.66 m × 0.84 m, a 0.18 m wide vented air channel, and a thermal storage wall. Experimental results showed the Trombe wall could increase the indoor temperature by 7.7 °C as compared to a reference room. Moreover, Ahmed et al. integrated a water-cooling system at the rear side of the PV cells within a PV-Trombe wall, such that heated water could be supplied to the building. In this Trombe wall configuration, PV cells were located at the inside surface of the glazing and a serpentine-shaped copper pipe with the dimensions of 15 mm diameter and 11.34 m length was placed between the PV cells and the air channel. Thermal and electrical efficiencies of 79.9% and 10.7%, respectively, were achieved [31].

To the best of the authors knowledge, research has yet to be reported on water-based PVTWs, wherein PV cells are integrated with a thermal energy storage medium comprising water. Herein, we study the benefits of placing PV cells on the lower portion of a water storage wall within a Trombe wall. Due to stratification, heated water rises to the top of the thermal storage wall while relatively cooler water resides at the bottom of the thermal storage medium. We investigate the potential benefits of placing PV cells on the bottom portion of the thermal storage wall where the temperature is relatively low. Thus, we report herein on a feasibility study about a novel semitransparent, water-based Trombe wall with integrated PV cells that can make full use of the solar irradiance to provide multiple functions: (1) it can function as a tinted semitransparent window; (2) it can generate heated water; (3) it can produce electric power.

2. Materials and Methods

2.1. Experimental Setup

A small Trombe wall prototype was built to experimentally investigate the ability to store solar thermal energy in water storage walls comprising a tinted acrylic sheet. The Trombe wall prototype is framed with wood having a cross section of 3.8 cm × 3.8 cm and its overall dimensions are 58 cm × 36 cm × 68 cm. The wooden frame is insulated with FOAMULAR extruded polystyrene boards and sealed using silicone sealant.

A 3 mm thick clear acrylic sheet with a height and width of 57.5 cm and 42.5 cm, respectively, is tightly fitted and sealed at the front face of the Trombe wall prototype. Similarly, a 3 mm thick acrylic sheet having dimensions of 30 cm × 30 cm is integrated at the back side, such that one can see through the transparent prototype. That is, the presence of the acrylic sheets at the front and back of the Trombe wall prototype allows for light to pass through the structure. An acrylic container with a width, height, and depth of 30 cm, 30 cm, and 20 cm, respectively, is placed within the Trombe wall prototype and is used to store the water. More details about this Trombe wall prototype are available in previously reported work [22]. In this work, monocrystalline silicon PV cells (Maxeon Gen II from Sunpower, with an efficiency rating of ~ 22%) with an area of 125 mm × 125 mm are located at the bottom half of the front surface of the water container as shown in Figure 1a,b. An insulated door with a small opening is used as the side of the prototype to allow users to change the thermal energy storage medium and to adjust the position of the PV cells and thermocouples.

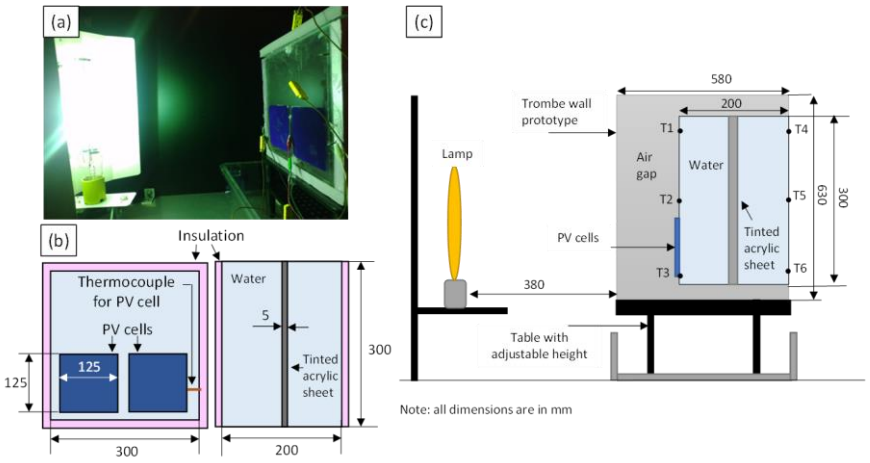


Figure 1. (**a**) A picture of the lamp and water container with insulation at its sides and bottom and PV cells on the bottom half of its front surface. (**b**) A schematic diagram showing the front (left) and cross section from the side view (right) of the water container. (**c**) A schematic diagram showing the side view of the experimental setup.

Solar radiation is simulated using a 1000 W Hortilux Blue Metal Halide bulb mounted vertically with a white reflector behind it. The front face of the Trombe wall prototype is 38 cm away from the metal halide bulb, as shown in Figure 1c. Experiments were carried out with and without a 5 mm thick tinted acrylic sheet inserted into the middle of the acrylic container (shown in Figure 1b). By placing the tinted acrylic sheet in the middle of the container (rather than placing it against the inner surface of the front or rear side of the container) its contact area with the water is maximized, which facilitates the transfer of heat generated in the acrylic container to the surrounding water. The tinted acrylic sheet absorbs solar-simulated light, which increases the amount of heat generated in the water. The reflectance (R) and transmittance (T) of the tinted acrylic sheet, and of the clear acrylic sheet used at the front and back of the Trombe wall prototype, were measured over a wavelength range from 220 nm to 1400 nm using a UV-Vis spectrophotometer (Shimadzu 2600i). The reflectance and transmittance of the tinted and clear acrylic sheets was measured using an FTIR spectrometer (VERTEX 70) over the wavelength range from 1400 nm to 25 µm. The absorptance (A) of these acrylic sheets was calculated using the reflectance and transmittance measurements and the equation A + R + T = 1. The transmittance and absorbance of the clear and tinted acrylic sheets are provided in Figure S1 in the Supplementary Materials.

The temperature of the water within the acrylic container was measured over the duration of the experiments using six type K thermocouples (labeled T1 through T6) at the locations shown in Figure 1. Thermocouples T1, T2, and T3 are installed 25 mm from the front surface. T2 is placed at the center of the container, whereas T1 and T3 are 5 mm from the top and bottom of the container, respectively. T4, T5, and T6 follow a similar configuration at the rear side of the container. Two PV cells are attached to the front face of the acrylic container, located at its bottom half as shown in Figure 1b, using silicon heat dissipation pads (Thermalright, Extreme Odyssey II). A thermocouple is placed between the PV cell and the acrylic container to monitor the temperature of the PV cell. This thermocouple is placed between the heat dissipation pad and the acrylic container near the center of the rear side of the PV cell. A data acquisition system (LabJack T7 pro) is used to record the temperature measurements and open circuit voltage of the PV cells. The open circuit voltage of silicon-based PV cells decreases substantially with increasing temperature, whereas the temperature dependence of the short-circuit current is comparatively small [23]. For example, Cuce et al. showed that the open circuit voltage decreases from about 19.5 V to 15 V as the temperature of a c-SiPV cell is increased from 15 °C to 60 °C, while the effects of this temperature increase on the short circuit current were negligible [32]. Thus,

in this work, the temperature dependence of the open circuit voltage is used to evaluate the feasibility of the proposed water-based Trombe wall.

2.2. Experimental Procedure

Each experiment consisted of a charging period followed by a discharging period. The solar-simulated light source is turned on at the beginning of the charging period, which lasts for 5 h for all experiments. In this work, 5 h was selected for the duration of the experiments because this is a good representation of the number of peak sun hours received in hot weather climates over the course of one day [33]. After 5 h, the metal halide lamp is turned off and the discharging period begins. During the discharging period, which lasts for 24 h, the heat stored within the water is dissipated to the surroundings. The total duration of each experiment, including the charging and discharging periods, is 29 h.

Experiments were conducted for the twelve cases shown in Table 1. For Cases 1 to 6, PV cells are not used. For Cases 1 and 2, the experimental setup consists of only the water container with and without the acrylic sheet, respectively. Cases 3 and 4 are similar to Cases 1 and 2, respectively, but the top, bottom and sides of the water container are insulated. Cases 5 and 6 are similar to Cases 3 and 4, but the water container is housed within the Trombe wall prototype shown in Figure 1.

Table 1. Experimental cases investigated.

Case	Water	Insulation	Trombe Wall	Acrylic Sheet (Tinted)	PV Cell	Abbreviation
1	✓					W
2	✓			✓		W-A
3	✓	✓				W-I
4	✓	✓		✓		W-I-A
5	✓	✓	✓			W-I-T
6	✓	✓	✓	✓		W-I-T-A
7	✓	✓			✓	W-I-P
8	✓	✓		✓	✓	W-I-A-P
9	✓	✓	✓		✓	W-I-T-P
10	✓	✓	✓	✓	✓	W-I-T-A-P
11		✓			✓	I-P
12		✓	✓		✓	I-T-P

For cases 6 to 12, PV cells are placed at the bottom front face of the water container as shown in Figure 1. For Cases 7 and 8, the water container is not housed within the Trombe wall prototype. For Case 8, the tinted acrylic sheet is within the water container, whereas for Case 7, the tinted acrylic sheet is not present. Cases 9 and 10 are similar to Cases 7 and 8, but with the water container housed within the Trombe wall prototype. For Cases 11 and 12, the water is removed from its container to provide a reference for investigating the extent to which the water is able to prevent the temperature of the PV cells from increasing.

2.3. Numerical Modeling Methods

To numerically model the temperature of the water storage medium over the duration of the charging phase, a three-dimensional model was first created in SOLIDWORKS software. The Reynolds number was determined as ~2500 and the heat transfer and turbulent flow for the cases described in Table 1 were modeled using ANSYS Fluent. All simulations were performed considering the standard k-epsilon turbulence model. The "SIMPLE" algorithm was employed to couple the pressure and velocity and an ambient temperature of 21 °C was used to initialize the model. For all simulations, the time step was 5 s and a quadrilateral mesh of 5 mm was used to discretize the governing equations

into a series of algebraic equations. Independence of the results from the time step and mesh size were checked (when decreasing either the time step or the mesh size to half of the values stated above the results changed by less than 0.01%). The convergence criteria of the CFD simulation were based on the residuals of the governing equations to be less than 1.0×10^{-5}. The acrylic container (300 mm × 200 mm × 300 mm) housing the water and tinted acrylic sheet is shown in Figure 2a. The mesh used to numerically model the thermal properties of this water storage medium is shown in Figure 2b. The lab-scale Trombe wall prototype is shown in Figure 2c, and the mesh used to numerically model the prototype is shown in Figure 2d. The thickness of the container is 5 mm on all sides except for the top, where it is 2.5 mm.

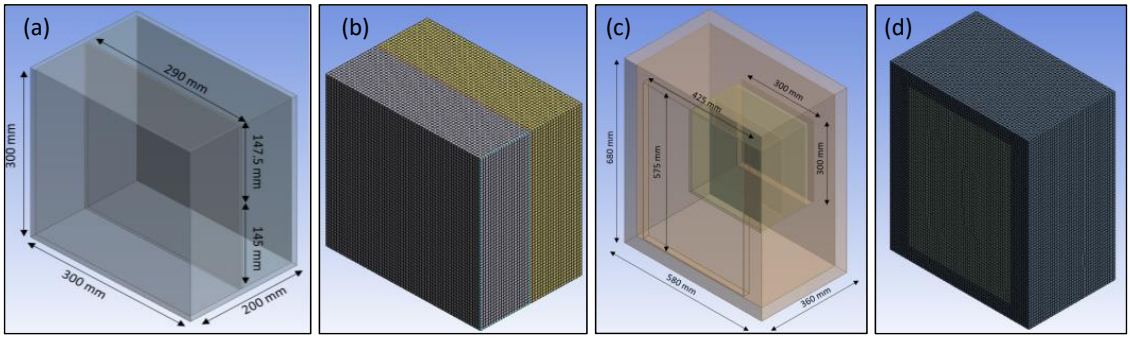


Figure 2. (**a**) A diagram of the water storage medium with the tinted acrylic sheet at its midplane. (**b**) The mesh used to perform the simulations for the water storage medium with the acrylic sheet at its midplane. (**c**) A diagram of the water storage medium within the small-scale Trombe wall prototype. (**d**) The mesh used to perform the simulations for the Trombe wall prototype.

To simulate the heat generated when incident light is absorbed a volumetric heat generation term is imparted in the cell zone. The incident light intensity measured 84 m W/cm^2 using a power meter (Thorlabs, PM 100D). To simplify the simulations for this feasibility study, it is assumed that 70% of the incident light intensity is absorbed by the acrylic sheet, 20% is absorbed by the water, and 10% is transmitted through the structure. As shown in Figure S1, the absorbance spectrum of the tinted acrylic sheet decreases from over 90% at 300 nm to less than 50% at 1600 nm, and for wavelengths above 2400 nm the absorbance is about 90%. Thus, assuming an average absorption of 70% for the tinted acrylic sheet is reasonable for the feasibility study conducted in this work. The volumetric heat generation term is determined by dividing the intensity of light absorbed by each component by its volume. Furthermore, a convective heat transfer coefficient of 10 W/m^2·K is assumed at the top and sides of the structure for all cases, while the bottom of the structure is assumed to be perfectly insulated for all cases.

3. Results

3.1. Experimental Results

The temperature profiles measured at the front of the water container (using T1, T2 and T3 in Figure 1) for Cases 1 to 6 over the 29 h duration of the experiments are shown in Figure 3. The temperature measurements taken near the rear side of the water container were similar to those measured at the front and are shown in Figure S2 in the supplementary material. The results for Case 1 are shown in Figure 3a. The initial temperature of the water at all locations was around 21 °C and the temperature started rising at t = 0 when the lamp was turned on and kept increasing over the next 5 h (over the charging period). As expected, the temperature of the water at the top of the container increased faster than the water at the middle and bottom due to thermal buoyancy forces caused by the density differences associated with water at different temperatures. The temperature at the top, middle and bottom of the water container increased to around 33 °C, 31 °C and 30 °C, respectively, and the average temperature gain for the water was around 10 °C over the charging period.

At the end of the charging period, the lamp was turned off and the discharging period began. The water temperature decreased over the discharging phase as heat was lost to the surroundings and the temperature of the water at the end of the discharging period was about 21.5 °C.

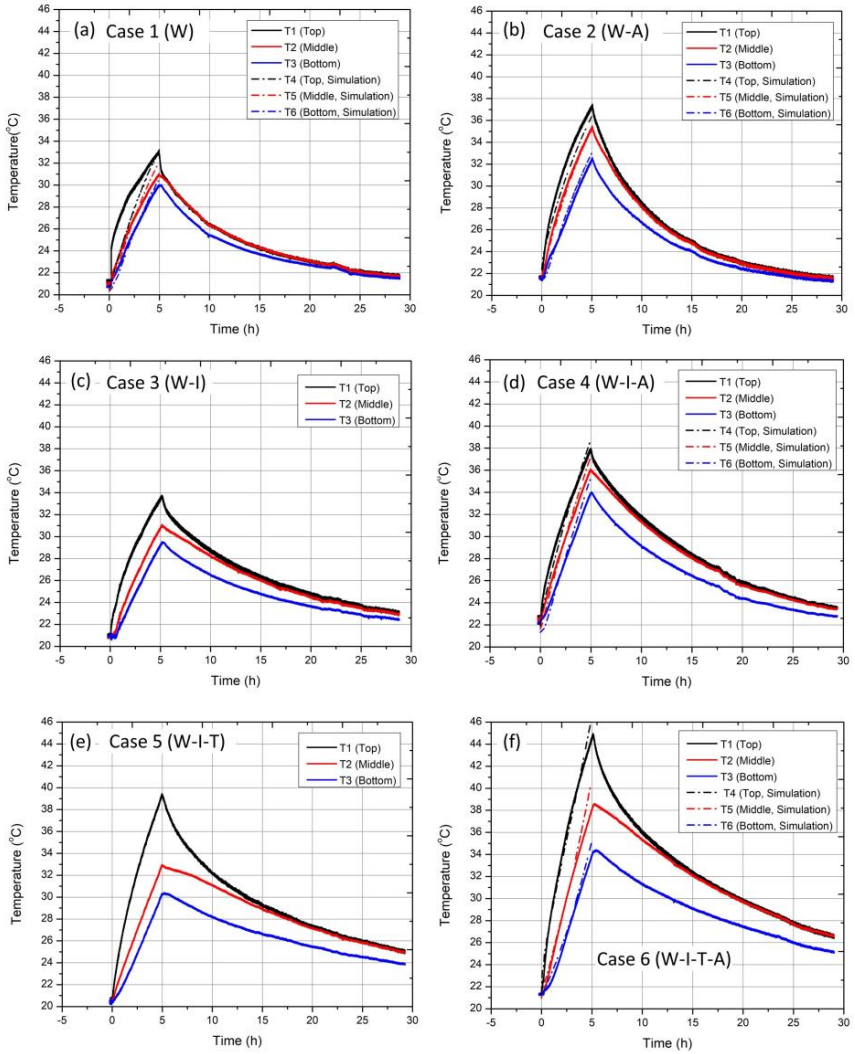


Figure 3. Temperature profiles during the charging and discharging periods for (**a**) just the container with water, (**b**) the water container with the tinted acrylic sheet inserted at its midplane, (**c**) the container with water when its sides are insulated, (**d**) the insulated water container when the tinted acrylic sheet is inserted at the midplane of the water, (**e**) the Trombe Wall, (**f**) Trombe wall acrylic sheet.

The results for Case 2 are shown in Figure 3b. The initial temperature of the water was about 21.5 °C and the temperature at the top, middle and bottom of the water container increased to approximately, 37 °C, 35 °C and 32.5 °C, respectively, by the end of the charging phase. Notably, the temperature of the water at the end of the charging phase is greater for Case 2 than for Case 1, which is attributed to the heat generated when a portion of the incident light is absorbed by the tinted acrylic sheet.

The temperature profiles for Case 3 are shown in Figure 3c. Compared to Case 1, the addition of the insulation at the sides of the water container results in a small increase in the water temperature throughout the container. In comparing the results from Case 3 to those of Case 2, it can be noted that inserting the tinted glass sheet at the center of the water container results in a larger temperature increase than adding insulation at the sides of the container. However, the ability to retain the stored thermal energy is better for Case 3 than for Case 2. For example, during the first five hours of the discharging phase (at the 10 h

mark of the experiment) the temperature of the water at the top of the acrylic container decreased by 9 °C for Case 2 and by 5 °C for Case 3.

The temperature profile measurements for Case 4 are shown in Figure 3d. With the addition of both the insulation and the tinted acrylic sheet the water temperature at the end of the charging phase is increased to 38 °C. Figure 3e, which plots the results for Case 5, shows the water temperature can be significantly increased by placing the water container in the Trombe wall prototype. When the tinted acrylic sheet is inserted into the water container within the Trombe wall prototype (Case 6, shown in Figure 3f), the temperature of the water at the top of the container is further increased to 45 °C. Moreover, the results from the numerical simulations over the charging phase while the water is subjected to solar simulated light are shown as the dashed lines in Figure 3b,d,f. The results from the experiment and simulations are in good agreement and are within 2 °C by the end of the charging phase for all locations (including the temperature at the top, middle and bottom of the water storage medium).

The maximum increase in temperature, which occurs at the top of the water container at the end of the charging phase is shown for Cases 1 through 10 in Figure 4. The lowest increase in temperature is ΔTmax = 11.7 °C for Case 1, whereas the highest increase of ΔTmax = 23.6 °C occurs for Case 6.

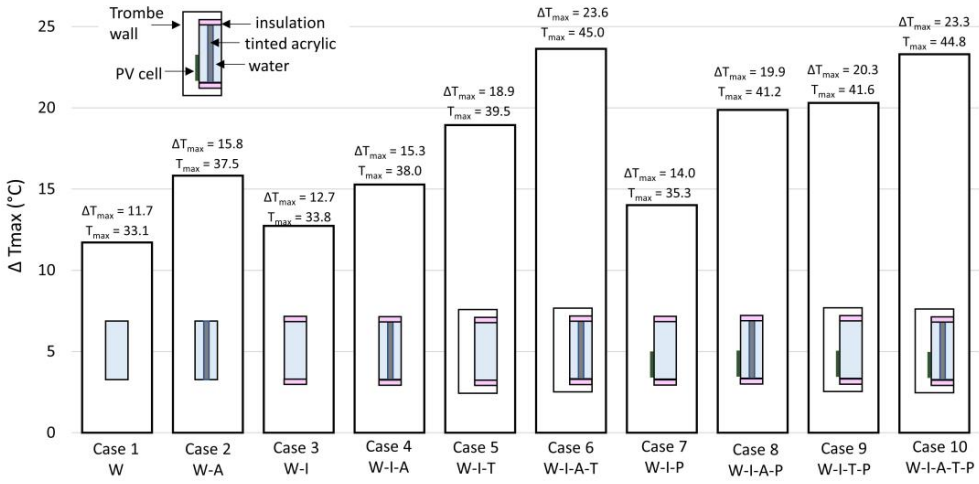


Figure 4. The temperature increase (ΔTmax) is measured at the top of the acrylic container for Cases 1 to 10. The maximum temperature, Tmax, measured at the top of the water container is also provided above the bar graph for each case.

The addition of the PV cells at the front bottom face of the acrylic container generally caused an increase in ΔTmax. For example, in comparing Cases 3 (W-I) and 7 (W-I-P), the addition of the PV cells resulted in an increase from ΔTmax = 12.7 °C to ΔTmax = 14.0 °C. In comparing Cases 4 (W-I-A) and 8 (W-I-A-P), the addition of the PV cells resulted in an increase from ΔTmax = 15.3 °C to ΔTmax = 19.9 °C. However, when comparing Cases 6 and 10, a slight decrease from ΔTmax = 23.6 °C to ΔTmax = 23.3 °C occurs with the addition of the PV cells when the Trombe wall prototype and tinted acrylic sheet are present in the experimental setup. Based on these results the heat generated when incident light is absorbed by the PV cells can cause the temperature of the water to increase during the charging phase if the amount of light being converted to heat in the water is low or if the insulation is relatively poor. However, for the case when the Trombe wall prototype and tinted acrylic sheet are present the temperature of the water is already elevated, and the addition of the PV cells does not have a significant effect on the water temperature.

The temperature profile for the water and PV cell for Case 7 are shown in Figure 5a,b, respectively. The open circuit voltage, Voc, of the PV cell is also shown in Figure 5b. In this work, the temperature and open circuit voltage of the PV cells were measured only during the charging phase. At the beginning of the charging phase, the output voltage is about

0.715 V, and this value decreases to 0.68 V by the end of the charging phase. This decrease in Voc is attributed to the increase in the PV cell temperature. Before the solar-simulated light is turned on, the temperature of the PV cell is about 21 °C. Once the light is turned on, the PV cell temperature is quickly increased to about 42 °C after the first 15 min of the charging phase. The PV cell temperature continues to rise gradually to a maximum temperature of 47 °C at the end of the charging phase.

The water and PV cell temperature profiles for Case 8 are shown in Figure 5c,d, respectively. Similarly to Case 7, for Case 8, the temperature of the PV cell increases from 21 °C before the charging phase begins to about 47 °C when the charging phase is completed. Likewise, the value of Voc decreases from 0.715 V to about 0.68 V. Figure 5e shows the temperature profiles in the water-based thermal storage medium for Case 9 and Figure 5f shows the temperature and open circuit voltage of the PV cell over the duration of the charging phase. The temperature of the PV cell increases from 21 °C at the onset of the charging phase to 47 °C at the end of the charging phase. At the beginning of the charging phase the open circuit voltage is 0.69 V, which is less than that for Cases 7 and 8 because the incident light intensity for Case 9 is reduced due to absorption and reflection by the acrylic window at the front face of the Trombe wall prototype. By the end of the charging phase, the open circuit voltage is reduced to 0.66 V due to the temperature rise of the PV cell.

The temperature profiles of the water within the Trombe wall prototype for Case 10 are shown in Figure 5g and the corresponding measurements for the temperature and Voc of the PV cell are shown in Figure 5h. The temperature of the PV cell increases from 21 °C to 49 °C during the charging phase and Voc decreases from 0.685 V to 0.66 V over the same time period. The temperature and Voc of the PV cell during the charging phase for Cases 11 and 12 are shown in Figure 5i,j, respectively. For case 11, Voc decreases from 0.68 V to about 0.64 V as the temperature of the PV cell increases from 21 °C to slightly over 70 °C over the five-hour duration that the PV cell is subjected to solar-simulated light. For Case 12, the temperature of the PV cell increases from 21 °C to about 89 °C, and Voc decreases from 0.685 to less than 0.6 V.

The ability for the cooler water at the bottom of the water storage medium to act as a heatsink that reduces the extent to which the temperature of the PV cell rises is clear from the data listed in Table 2. The temperature at the bottom of the water storage medium near its front side (T3) at the end of the charging phase ranges from 30.2 °C to 34.8 °C for Cases 7 through 10. The corresponding temperature of the PV cell ranges from 48.0 to 49.7 °C for Cases 7 through 10. However, for Cases 11 and 12, when the water is not present in the water storage medium, the temperature of the PV cell greatly increases to 71.4 °C and 90 °C, respectively. Furthermore, Voc decreases from 0.68 V to 0.64 V for Case 11 and from 0.658 V to less than 0.6 V for Case 12 while being subjected to the incident solar radiation over five hours. Notably, the drop in the output voltages observed for Cases 11 and 12 are comparable to those reported in the literature for the corresponding increase in temperature [34,35].

Table 2. Temperatures and open circuit voltages for Cases 7–12.

CASE	T3 at the End of the Charging Phase (°C)	ΔT3 at the End of the Charging Phase (°C)	T of the PV Cell at the End of the Charging Phase (°C)	ΔT for the PV Cell at the End of the Charging Phase (°C)	Voc at the End of the Charging Phase (V)	ΔVoc at the end of the Charging Phase (V)
7 (W-I-P)	30.2	9.3	48.0	27.1	0.680	0.034
8 (W-I-A-P)	30.6	9.5	48.2	27.4	0.679	0.036
9 (W-I-T-P)	32.5	11.1	48.7	27.6	0.670	0.028
10 (W-I-T-A-P)	34.8	13.3	49.7	28.5	0.668	0.024
11 (I-P)	—	—	71.4	50.5	0.642	0.038
12 (I-T-P)	—	—	88.8	68.2	0.600	0.068

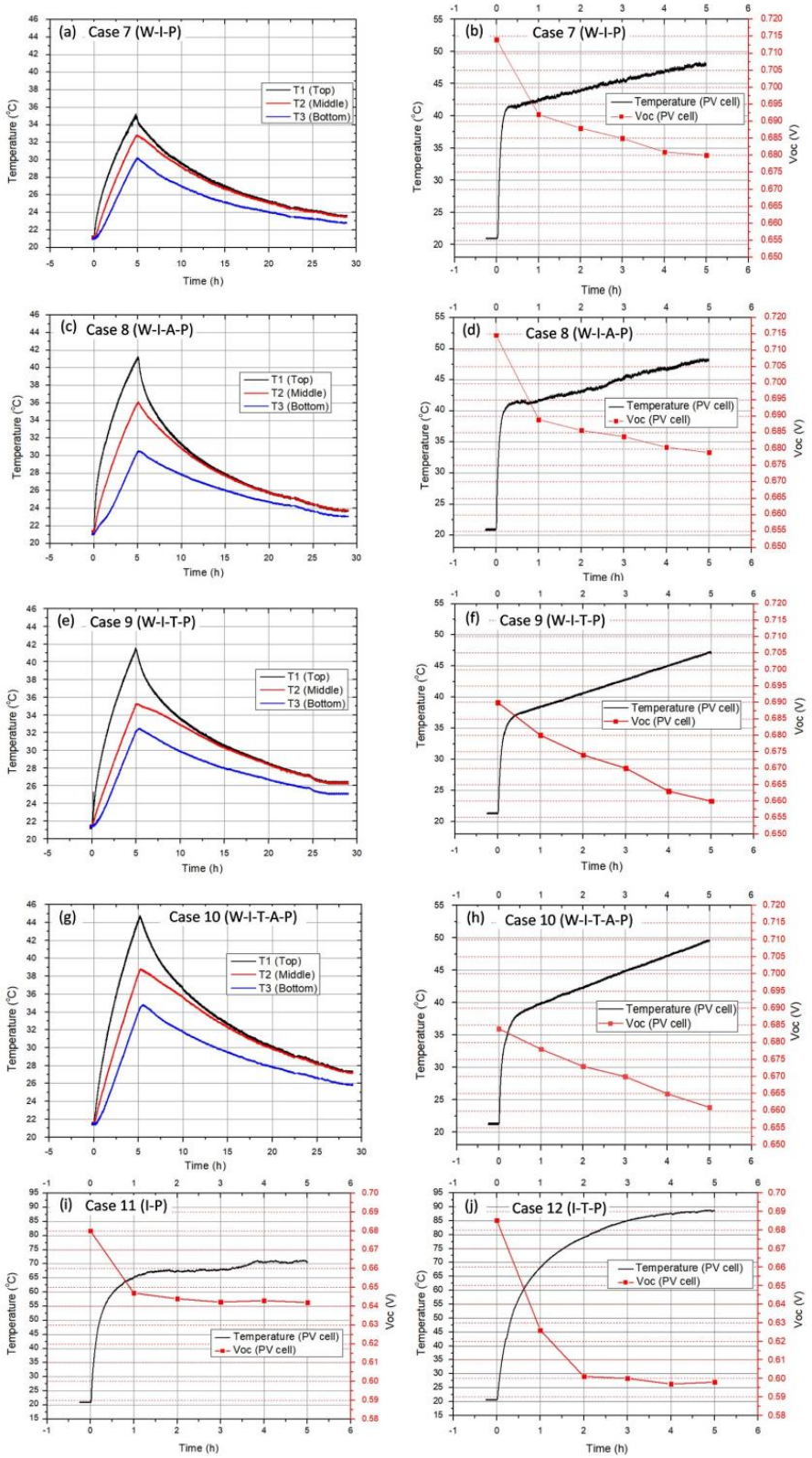


Figure 5. Temperature profiles of the water for (**a**) Case 7, (**c**) Case 8, (**e**) Case 9, and (**g**) Case 10. Temperature and open circuit voltage profiles for the PV cell for (**b**) Case 7, (**d**) Case 8, (**f**) Case 9, (**h**) Case 10, (**i**) Case 11, and (**j**) Case 12.

3.2. Results from the Numerical Simulations

Schematic diagrams of a full-sized Trombe wall are shown in Figure 6a,b. Similar to the small-scale Trombe wall prototype, the clear water storage container has a depth of

20 mm, and the tinted acrylic sheet resides at its midplane. For the full-scale Trombe wall configuration, there is a 100 mm thick air channel between the water storage medium and the indoor wall of the building. Furthermore, as shown in Figure 6a, it is assumed that PV cells are placed on the bottom half of the front face of the water storage medium. There are 6 rows and 17 columns of PV cells (102 PV cells in total), each with a facial area of 125 mm × 125 mm, such that almost half of the bottom surface of the Trombe wall is covered with PV cells. The temperature at the top, middle and bottom of the full scale Trombe wall is plotted in Figure 6c. In simulating these results, it was assumed that the heat generation terms for the bottom half of the structure were reduced by 50% due to the presence of the PV cells. Heat generated due to absorption of sunlight in the PV cells was not considered. All other methods used in the simulation are similar to those described in Section 2.3. As shown by the black line in Figure 6c, under this conservative assumption, the temperature at the top of the water storage medium in the Trombe wall reaches 44 °C when the wall is subjected to sunlight for five hours. Notably, the simulated temperature profiles shown in Figure 6 are comparable to the experimentally measured temperature profiles shown in Figure 5. Nevertheless, it should be noted that the results in Figure 6 are based on rough assumptions for the values of the heat generated when light is absorbed in the Trombe wall and for the heat loss at the boundaries of the structure. Thus, the initial modelling results presented in this feasibility study show the Trombe wall proposed in this work has the potential to effectively provide heated water that is useful for building applications. Subsequent work should include the modelling and fabrication of a large-scale physical prototype to fully validate the concept.

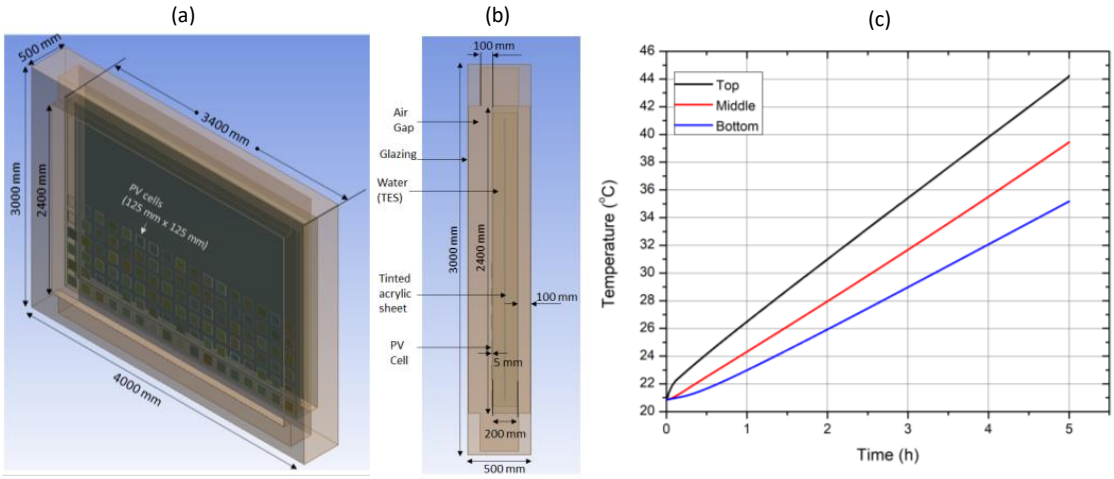


Figure 6. (**a**) Diagram of a 3 m tall Trombe wall; (**b**) cross-sectional view of a 3 m tall Trombe wall; (**c**) comparison of the simulated temperature profiles at the top, middle and bottom of the water storage medium within the 3 m Trombe wall with the small scale prototype.

4. Discussion

The results presented in this work demonstrate that Trombe walls can be used to simultaneously provide multiple functions. The semitransparent, water-based Trombe wall can generate and store heated water, allow for some degree of indoor lighting, and produce electric power. For example, the temperature of the water at the top of the acrylic container reached almost 45 °C, while PV cells were attached to the bottom half of its front surface for Case 10, which is sufficient to provide for most hot water applications in buildings [36]. As one example, it has been reported that dishwashers operate efficiently at temperatures as low as 38 °C, and the temperature of water used for showering is typically in the range of 37–40 °C. Moreover, the recommended water temperature in newer washing machines for high efficiency is approximately 49 °C. Furthermore, it has been suggested that the safest temperature for a hot water tank is 49 °C. Considering the temperature of the inlet water for hot water tanks in buildings during warm weather conditions is commonly ~18 °C,

the Trombe wall presented in this work could provide preheated water to the hot water tank to minimize the amount fuel it must consume to meet water heating demands [37]. The heated water can also be stored within the Trombe wall and used hours after sunlight is no longer available. For Case 10, the temperature of the heated water is about 37 °C at the ten-hour mark of the experiment (five hours after the light is switched off), which is still useful for most applications in buildings. The stored water could also be used as preheated water that is sent to a boiler. Furthermore, for Case 10, the temperature of the PV cell increased up to about 47 °C, which is not uncommon for PV cells operating in the field [38,39].

The performance of the Trombe wall presented in this work would be enhanced by operating it as a dynamic building envelope [40,41]. The PV Cells and tinted acrylic sheet could be integrated as light-weight panels that can be removed to satisfy varying demand profiles for electric power, heat and indoor lighting. For example, during summer, a tinted acrylic sheet that absorbs more solar irradiance could be used in the Trombe wall to prevent overheating, whereas a lightly tinted acrylic sheet could be used in winter to allow more solar heat and light to enter the building. It should also be noted that there is a tradeoff between the level of daylight provided to the interior of the building through the Trombe wall and maximizing the temperature the water can be heated to. In this regard, the optical properties of the tinted acrylic sheet used in this work could also be optimized. The visible light spectrum extends from 380 nm to 700 nm; however, as shown in Figure S1, the transmitted light spectrum in the range from 800 nm to 1800 nm is fairly high (about 35% on average). It would be preferrable if the tinted acrylic sheet absorbed all the incident solar radiation over this spectral region to heat the water in the storage medium to the greatest extent possible. It should also be noted that coated acrylics can be used in place of the tinted acrylic to provide for a selection of available brightness and colors [42].

Moreover, the performance of the water-based, semitransparent photovoltaic thermal Trombe wall presented herein could be further improved by using semitransparent PV cells [43–48] and by using optical coatings to increase the transparency of the Trombe wall window [49–52]. Semitransparent PV cells transmit light below the bandgap of the PV cells, which would allow more light to be absorbed by the acrylic sheet in the water storage medium. Semitransparent PV cells may also reduce the amount of absorbed and reflected light, which could reduce the temperature in the air channel of the Trombe wall and the PV cell. The results presented herein demonstrate the feasibility of the water-based, semitransparent photovoltaic thermal Trombe wall design. However, further research is required to validate the concept using a full-scale prototype.

There are some practical considerations to take into account when scaling the water-based Trombe wall proposed in this work. The strength of the transparent material used to contain the water would have to be designed to safely withstand the mechanical loads imposed by the water's weight. Notably, polycarbonate sheets have been fabricated for use in load-bearing structures [53]. Furthermore, supplemental heating may be needed to prevent the water from freezing if the building is unoccupied for long periods in cold climates and its temperature is allowed to drop. Moreover, to keep the stored water clear, it should not be left stagnant and should be filtered prior to entering the storage medium [54,55].

5. Conclusions

In this work, we experimentally investigated the feasibility of using Trombe walls to perform multiple functions during warm weather conditions including (1) heating and storing water for building applications, (2) providing occupants with visibility to the outdoors, and (3) generating electric power. Experiments were carried out on a small-scale prototype comprising a clear water storage container that is inside an insulated enclosure with transparent windows at its front and rear sides. A tinted acrylic sheet was immersed in the water within the clear container and was illuminated with solar-simulated light. Photovoltaic cells were placed at the lower half of the front face of the water storage

container. The cooler water at the bottom of the water storage container acted as a heat sink that reduced the extent to which the temperature of the PV cells increased.

The presence of the PV cells on the lower half of the front face of the water container did not have a significant effect on the water temperature. When the PV cells were absent, the temperature of the water at the top of the storage container was elevated to about 46 °C after being subjected to incident solar radiation for five hours. On the other hand, the temperature of the water at the top of the storage container was elevated to almost 45 °C when subjected to solar-simulated radiation over the same timeframe. This result suggests that electric power from the PV cells and hot water can be generated simultaneously in the water-based Trombe wall configuration presented in this work. The prototype also exhibited the ability to store thermal energy, as the temperature at the top of the water container was 37 °C five hours after the solar-simulated light was switched off. This is an important result because it suggests the water-based Trombe wall will be able to satisfy a portion of the demand for heated water in buildings after the sun has set. The results also show the benefits of thermal stratification within the water tank. While the temperature at the top of the water storage medium was about 45 °C, the temperature of the water near the bottom of the storage container did not surpass 35 °C. The relatively cooler water at the bottom of the container functions as an effective heat sink that limits the extent to which the temperature of the PV cells is elevated.

To demonstrate the ability and importance of the water to function as an effective heat sink, the temperature and open circuit voltage of the PV cells was measured with and without water present in the storage container while the Trombe wall prototype was subjected to incident solar-simulated radiation for five hours. When the water was present, the open circuit voltage was about 0.685 V just after the light was turned on. Over the five hour duration that the light was on, the temperature of the PV cell increased to about 50 °C, and the open circuit voltage of the PV cell decreased to 0.66 V. However, for the experiment wherein the light was removed from the container, the temperature of the PV cell increased to almost 90 °C after being subjected to light from the solar simulator for five hours and the open circuit voltage of the PV cell decreased from 0.685 V to less than 0.60 V over this time period.

Further research is required to investigate the feasibility and potential of water-based, semitransparent photovoltaic thermal Trombe walls at full-scale and under actual weather conditions. Nevertheless, the results attained thus far in this work bode well for the potential of this Trombe wall design and warrant its further development.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/en16041618/s1. Figure S1: Transmittance spectrum for the clear acrylic sample (top left), transmittance spectrum for the tinted acrylic sample (top right), absorptance spectrum for the clear acrylic sample (bottom left) and the absorptance spectrum for the tinted acrylic sample (bottom right). Figure S2: Temperature profiles at the front and back of the water storage medium (measured using thermocouples T1 through T6 in Figure 1) during the charging and discharging periods for the experiments carried out for (a) Case 1, (b) Case 2, (c) Case 3, (d) Case 4, (e) Case 5, (f) Case 6, (g) Case 7, (h) Case 8, (i) Case 9, (j) Case 10.

Author Contributions: Conceptualization, H.S. and P.G.O.; Methodology, S.B., N.S., A.K.K., H.S. and P.G.O.; Formal Analysis, S.B., N.S. and P.G.O.; Investigation, S.B., N.S., A.K.K. and P.G.O.; writing—original draft preparation, S.B., N.S. and P.G.O.; writing—review and editing, S.B., N.S. and P.G.O.; supervision, P.G.O.; funding acquisition, P.G.O. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Mitacs Globalink Research Internship (GRI) Program and by the Natural Sciences and Engineering Research Council of Canada (NSERC, RGPIN-2017-05987).

Acknowledgments: Assistance provided by Abdallah Alshantaf in fabricating the Trombe wall prototype experimental setup is greatly appreciated. The authors also thank Atousa Pirvaram for assistance with the optical measurements.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Cozzi, L.; Gould, T. *World Energy Outlook 2022*; International Energy Agency: Paris, France, 2022.
2. Löf, G.O.G.; Duffie, J.A.; Smith, C.O. World distribution of solar radiation. *Solar Energy* **1966**, *10*, 27–37. [CrossRef]
3. Vaisi, S.; Pilla, F.; McCormac, D.; McCormack, S.J. Towards urban energy density (UED) definition. In Proceedings of the Energy for Sustainability 2015—Sustainable Cities: Designing for People and the Planet, Coimbra, Portugal, 14–15 May 2015.
4. Howard, B.; Parshall, L.; Thompson, J.; Hammer, S.; Dickinson, J.; Modi, V. Spatial distribution of urban energy consumption by end use. *Energy Build.* **2012**, *45*, 141–151. [CrossRef]
5. Güneralp, B.; Zhou, Y.; Urge-Vorsatz, D.; Gupta, M.; Sha, Y.; Patel, P.L.; Fragkias, M.; Li, X.; Seto, K.C. Global scenarios of urban density and its impacts on building energy use through 2050. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 8945–8950. [CrossRef] [PubMed]
6. Daigle, Q.; O'Brien, P.G. Heat Generated Using Luminescent Solar Concentrators for Building Energy Applications. *Energies* **2020**, *13*, 5574. [CrossRef]
7. Saadatian, O.; Sopian, K.; Lim, C.H.; Asim, N.; Sulaiman, M.Y. Trombe walls: A review of opportunities and challenges in research and development. *Renew. Sustain. Energy Rev.* **2012**, *16*, 6340–6351. [CrossRef]
8. Hu, W.; He, J.; Ji, J.; Zhang, S. A review on the application of Trombe wall system in buildings. *Renew. Sustain. Energy Rev.* **2017**, *70*, 976–987. [CrossRef]
9. Sergei, K.; Shen, C.; Jiang, Y. A review of the current work potential of a Trombe wall. *Renew. Sustain. Energy Rev.* **2020**, *130*, 109947. [CrossRef]
10. Duffie, J.A.; Beckman, W.A. *Solar Engineering of Thermal Processes*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 1991; pp. 250–330.
11. Mokni, A.; Lashin, A.; Ammar, M.; Mhiri, H. Thermal analysis of a Trombe wall in various climatic conditions: An experimental study. *Solar Energy* **2022**, *243*, 247–263. [CrossRef]
12. Yang, L.; Dhahad, H.A.; Chen, M.; Huang, Z.; Anqi, A.E.; Rajhi, A.A.; Qader, D.N. Transient analysis of buildings with Trombe wall in a southern envelope and strengthening efficacy by adding phase change material. *J. Build. Eng.* **2022**, *55*, 104670. [CrossRef]
13. Liu, H.; Li, P.; Yu, B.; Zhang, M.; Tan, Q.; Wang, Y. The performance analysis of high-efficiency dual-channel Trombe wall in winter. *Energy* **2022**, *253*, 124087. [CrossRef]
14. Shen, J.; Lassue, S.; Zalewski, L.; Huang, D. Numerical study of classical and composite solar walls by TRNSYS. *J. Therm. Sci.* **2007**, *16*, 46–55. [CrossRef]
15. Chan, H.Y.; Riffat, S.B.; Zhu, J. Review of passive solar heating and cooling technologies. *Renew. Sustain. Energy Rev.* **2010**, *14*, 781–789. [CrossRef]
16. Stazi, F.; Mastrucci, A.; di Perna, C. The behaviour of solar walls in residential buildings with different insulation levels: An experimental and numerical study. *Energy Build.* **2012**, *47*, 217–229. [CrossRef]
17. Stazi, F.; Mastrucci, A.; di Perna, C. Trombe wall management in summer conditions: An experimental study. *Solar Energy* **2012**, *86*, 2839–2851. [CrossRef]
18. Gan, G. A parametric study of Trombe walls for passive cooling of buildings. *Energy Build.* **1998**, *27*, 37–43. [CrossRef]
19. Jaber, S.; Ajib, S. Optimum design of Trombe wall system in mediterranean region. *Solar Energy* **2011**, *85*, 1891–1898. [CrossRef]
20. Torcellini, P.; Pless, S. *Trombe Walls in Low-Energy Buildings: Practical Experiences*; National Renewable Energy Lab.: Golden, CO, USA, 2004.
21. Mohamad, A.; Taler, J.; Oclon, P. Trombe Wall Utilization for Cold and Hot Climate Conditions. *Energies* **2019**, *12*, 285. [CrossRef]
22. Singh, H.; O'Brien, P.G. Semi-transparent water-based Trombe walls for passive air and water heating. *Buildings* **2022**, *12*, 1632. [CrossRef]
23. Skoplaki, E.; Palyvos, J.A. On the temperature dependence of photovoltaic module electrical performance: A review of efficiency/power correlations. *Solar Energy* **2009**, *83*, 614–624. [CrossRef]
24. Dubey, S.; Sarvaiya, J.N.; Seshadri, B. Temperature dependent photovoltaic (PV) efficiency and its effect on PV production in the world–a review. *Energy Procedia* **2013**, *33*, 311–321. [CrossRef]
25. Temaneh-Nyah, C.; Mukwekwe, L. An investigation on the effect of operating temperature on power output of the photovoltaic system at University of Namibia Faculty of Engineering and IT campus. Proceeding of the 2015 Third International Conference on Digital Information, Networking, and Wireless Communications (DINWC), Moscow, Russia, 3–5 February 2015; pp. 22–29.
26. Rahman, M.M.; Hasanuzzaman, M.; Abd Rahim, N. Effects of operational conditions on the energy efficiency of photovoltaic modules operating in Malaysia. *J. Clean. Prod.* **2017**, *143*, 912–924. [CrossRef]
27. Perraki, V.; Kounavis, P. Effect of temperature and radiation on the parameters of photovoltaic modules. *J. Renew. Sustain. Energy* **2016**, *8*, 013102. [CrossRef]
28. Hu, Z.; He, W.; Hu, D.; Lv, S.; Wang, L.; Ji, J.; Chen, H.; Ma, J. Design, construction and performance testing of a PV blind-integrated Trombe wall module. *Appl. Energy* **2017**, *203*, 643–656. [CrossRef]
29. Koyunbaba, B.K.; Yilmaz, Z.; Ulgen, K. An approach for energy modeling of a building integrated photovoltaic (BIPV) Trombe wall system. *Energy Build.* **2013**, *67*, 680–688. [CrossRef]
30. Jie, J.; Hua, Y.; Gang, P.; Jianping, L. Study of PV-Trombe wall installed in a fenestrated room with heat storage. *Appl. Therm. Eng.* **2007**, *27*, 1507–1515. [CrossRef]
31. Ahmed, O.K.; Hamada, K.I.; Salih, A.M. Performance analysis of PV/Trombe with water and air heating system: An experimental and theoretical study. *Energy Sour. Part A Recovery Util. Environ. Eff.* **2022**, *44*, 2535–2555. [CrossRef]

32. Cuce, E.; Cuce, P.M.; Bali, T. An experimental analysis of illumination intensity and temperature dependency on photovoltaic cell parameters. *Appl. Energy* **2013**, *111*, 374–382. [CrossRef]
33. NREL. Solar Resource Maps and Data. Available online: https://www.nrel.gov/gis/solar-resource-maps.html (accessed on 27 January 2023).
34. Singh, P.; Ravindra, N.M. Temperature Dependence of Solar Cell Performance—An Analysis. *Solar Energy Mater. Solar Cells* **2012**, *101*, 36–45. [CrossRef]
35. Cotfas, D.T.; Cotfas, P.A.; Machidon, O.M. Study of temperature coefficients for parameters of photovoltaic cells. *Int. J. Photoenergy* **2018**, *1*, 1–12. [CrossRef]
36. Marszal-Pomianowska, A.; Jensen, R.L.; Pomianowski, M.; Larsen, O.K.; Jørgensen, J.S.; Knudsen, S.S. Comfort of domestic water in residential buildings: Flow, temperature and energy in draw-off points: Field study in two danish detached houses. *Energies* **2021**, *14*, 3314. [CrossRef]
37. Water Temperature and Burns/Scalds. Available online: https://www.canada.ca/en/public-health/services/water-temperature-burns-scalds.html (accessed on 17 October 2022).
38. Schwingshackl, C.; Petitta, M.; Wagner, J.E.; Belluardo, G.; Moser, D.; Castelli, M.; Zebisch, M.; Tetzlaff, A. Wind effect on PV module temperature: Analysis of different techniques for an accurate estimation. *Energy Procedia* **2013**, *40*, 77–86. [CrossRef]
39. Van Sark, W.G.J.H.M. Feasibility of photovoltaic–thermoelectric hybrid modules. *Appl. Energy* **2011**, *88*, 2785–2790. [CrossRef]
40. Svetozarevic, B.; Begle, M.; Jayathissa, P.; Caranovic, S.; Shepherd, R.F.; Nagy, Z.; Hischier, I.; Hofer, J.; Schlueter, A. Dynamic photovoltaic building envelopes for adaptive energy and comfort management. *Nat. Energy* **2019**, *4*, 671–682. [CrossRef]
41. Tabadkani, A.; Roetzel, A.; Li, H.X.; Tsangrassoulis, A. Design approaches and typologies of adaptive facades: A review. *Autom. Constr.* **2021**, *121*, 103450. [CrossRef]
42. Maduru, V.R.; Shaik, S.; Cuce, E.; Afzal, A.; Panchal, H.; Cuce, P.M. UV coated acrylics as a substitute for generic glazing in buildings of Indian climatic conditions: Prospective for energy savings, CO_2 abatement, and visual acceptability. *Energy Build.* **2022**, *268*, 112231. [CrossRef]
43. Ghosh, A.; Sundaram, S.; Mallick, T.K. Investigation of thermal and electrical performances of a combined semi-transparent PV-vacuum glazing. *Appl. Energy* **2018**, *228*, 1591–1600. [CrossRef]
44. Bailey-Salzman, R.F.; Rand, B.P.; Forrest, S.R. Semitransparent organic photovoltaic cells. *Appl. Phys. Lett.* **2006**, *88*, 233502. [CrossRef]
45. Yang, Y.; O'Brien, P.G.; Ozin, G.A.; Kherani, N.P. See-through amorphous silicon solar cells with selectively transparent and conducting photonic crystal back reflectors for building integrated photovoltaics. *Appl. Phys. Lett.* **2013**, *102*, 221109. [CrossRef]
46. O'Brien, P.G.; Chutinan, A.; Mahtani, P.; Leong, K.; Ozin, G.A.; Kherani, N.P. Selectively transparent and conducting photonic crystal rear-contacts for thin-film silicon-based building integrated photovoltaics. *Opt. Express* **2011**, *19*, 17040–17052. [CrossRef] [PubMed]
47. Alrashidi, H.; Issa, W.; Sellami, N.; Sundaram, S.; Mallick, T. Thermal performance evaluation and energy saving potential of semi-transparent CdTe in Façade BIPV. *Solar Energy* **2022**, *232*, 84–91. [CrossRef]
48. Uddin, M.M.; Wang, C.; Zhang, C.; Ji, J. Investigating the energy-saving performance of a CdTe-based semi-transparent photovoltaic combined hybrid vacuum glazing window system. *Energy* **2022**, *253*, 124019. [CrossRef]
49. Sarkin, A.S.; Ekren, N.; Sağlam, S. A review of anti-reflection and self-cleaning coatings on photovoltaic panels. *Solar Energy* **2020**, *199*, 63–73. [CrossRef]
50. Chunxue, J.; Liu, W.; Bao, Y.; Chen, X.; Yang, G.; Wei, B.; Yang, F.; Wang, X. Recent applications of antireflection coatings in solar cells. *Photonics* **2022**, *9*, 906.
51. Loh, J.Y.Y.; Puzzo, D.P.; O'Brien, P.G.; Ozin, G.A.; Kherain, N.P. Enhancing photovoltaics with broadband high-transparency glass using porosity-tuned multilayer silica nanoparticle anti-reflective coatings. *RSC Adv.* **2014**, *4*, 31188–31195. [CrossRef]
52. Vázquez-Guardado, A.; Boroumand, J.; Franklin, D.; Chanda, D. Broadband angle-independent antireflection coatings on nanostructured light trapping solar cells. *Phys. Rev. Mater.* **2018**, *2*, 035201. [CrossRef]
53. Zapałowicz, Z.; Wojnicki, O. Estimation of total solar transmittance for twin-wall polycarbonate sheet with rectangular structure on the basis of experimental research. *Energies* **2022**, *15*, 1360. [CrossRef]
54. Brown, K.W.; Gessesse, B.; Butler, L.J.; MacIntosh, D.L. Potential effectiveness of point-of-use filtration to address risks to drinking water in the United States. *Environ. Health Insights* **2017**, *11*. [CrossRef]
55. Cescon, A.; Jiang, J. Filtration process and alternative filter media material in water treatment. *Water* **2020**, *12*, 3377. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

2022 IEEE 2nd International Conference on Mobile Networks and Wireless Communications (ICMNWC)

Design and Modal Analysis of Optical Fibers with Multiple Cores and Multiple Cladding Fiber

Ankit Kumar
Dept. of Electronics and communication engineering
Delhi technological university
New Delhi, India
Ankitgaur176@gmail.com, Ankitkumar_2k18ceec04@dtu.ac.in

2022 IEEE 2nd International Conference on Mobile Networks and Wireless Communications (ICMNWC) | 978-1-6654-9111-2/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICMNWC56175.2022.10031958

***Abstract*—Multi-core fiber (MCF) and multi-layer fiber (MLF) concepts are an attractive designing model that will enhance the next generation of high-speed fiber optical transmission (FOT) system. So, in this paper we have introduced symmetric type MCF and MLF fiber designing models and simulated with various parameter sweeps and analysis the mode profile of both types of fiber models.**

Keywords—MCF and MLF, designing and simulation, mode profile analysis.

I. INTRODUCTION

We know that the fiber technology is a backbone of the communication network and we also accept that the fiber technology is the most expensive and has a limited transmission capacity of SMF (100Tb/s) [1]. But nowadays we show that the transmission capacity gradually increases with time, so we have to require a novel fiber model, which is enhance our fiber technology [2].

Multi-Core Fiber (MCF): In this fiber model, we have many numbers of cores in single cladding, so we call it multi-core fiber (MCF) [3]. The multi-core fiber behaves like a bundle of single mode fiber since the core is meant to be single mode but multiple numbers of fiber cores exist in a single cladding due to this design. The result is enhanced the signal bandwidth as compared to SMF.

Multi-Layer Fiber (MLF): A multilayer optical fiber model is defined as a new type of fiber model which is modified from single mode fiber [4]. Here modified part is that double layer of cladding that is use in the model.

So, the basic purpose of this model is to reduce the absorption losses, and here we know that the reflection index of core and cladding is depends upon the degree of mode representation.

For that reason, we varied some steps of wavelength, radius, and effective mode index, respectively. Demonstrate how the core has evolved. As illustrated here, many forms of analysis are displayed in various modes with various settings [5].

So, that way we want to introduce a new type of optical fiber model, which is known as multi-core and multi-layer fiber model and analyze the different types of mode profile through COMSOL application to change the various core radios, effective mode index and wavelength.

In this scenario, we expect to introduce a unique fiber model which is use as a maximum data speed with large bandwidth and a large propagation length at minimum loss.

II. RELATED WORKS

Every ten years, the demand for internet data traffic increases by roughly 100 times. The information is carried through a variety of modern technologies [6].

Current single-mode single-core capacity (SM-SC) 100 Tb/s fiber has been achieved. In the foreseeable future, the SM-SC fiber may not be able to meet the fast-expanding demand. Multicore fiber is now in use. The thrust study area in the domain is (MCF) technology [7]. Because of the presence of fiber optic technology, if there are many cores in the same limited cladding region, the fiber's ability to convey information will be improved. as a result, it can outperform any potential capacity [8].

So, the basic motivation point is that single mode fiber has a limit of data carrying capacity with long propagation length, such as the 10G Base SE-SR at 10km (for OS2), and the Multimode fiber has the 10G Base SE-SR at 300m (for OM3), 400m (for OM4), and 300m(OM5) [9]. So that region we have design a different type of fiber model which is contain few numbers of same geometries of core and include a single cladding which diameter same as reference of multimode fiber [10].

III. LITERATURE SURVEY

Lots of research papers have been published on date with different way to designing of optical fiber models and analytical terms. We now understand that all respective papers demonstrate some important role with different type of technology, such as spacing and dimension factor, which play a major role between two and more than two cores fiber model [11].

For this reason, MCFs are a promising solution for extending the optical fiber-link information capacity into the space-division multiplexing dimension, and for this reason they have received a lot of attention from research groups as well as standard telecom-fiber manufacturers [12,13].

IV. METHODS

Designing and Simulation

A. Multi-Core Fiber (MCF)

In the simulation and design, we use COMSOL software for simulate of design and analyze the mode profile with various parameter sweeps. Here we use as parameter specification is diameter of core 1.65um and cladding diameter is 2.6um with Pitch area between core and cladding is 3.3um and we varied the operating frequency with respect to wavelength.

978-1-6654-9111-2/22/$31.00 ©2022 IEEE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on February 13,2023 at 06:38:16 UTC from IEEE Xplore. Restrictions apply.

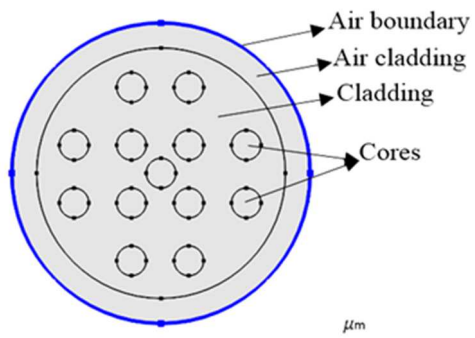


Fig. 1. Cross sectional view of Multicore fiber model

RI of Air boundary and Air cladding is 1, RI of Inner air cladding is 1.48 and RI of multiple-core is 1.5.

In multicore fiber, addressing each core independently is difficult due to the narrow channel spacing as shown in fig-1. It is difficult to address each core independently in multicore fiber [1,14]. Here we show the multiple parameter variations for simulation and analyze the mode profile.

The dimensions given below in table-1 show the dimensions of the multi-core fiber model:

TABLE I. DIMENSIONS OF MLC

Notation	Description	Dimensions
Dco	diameter core	1.65[um]
A	pitch	3.3[um]
Dcl	cladding diameter	2.6[um]
Wl	operating wavelength	1.55[nm]

B. *Multi-Layer Fiber (MLF)*

The multi-layer designing concept is based upon the layering of cladding on a core. These phenomena are based on the step indexing concept and in that way, we expect the absorption losses to be reduced due to reflecting to the cladding layer [10]. Layering of cladding with decreasing reflection index bends the rays toward the core, and this phenomenon we use to reduce the absorption loss which is absorbed by the outer lather coating. But here we only analyze the mode profile at various parameter sweeps and we check the appropriate range of the mode profile.

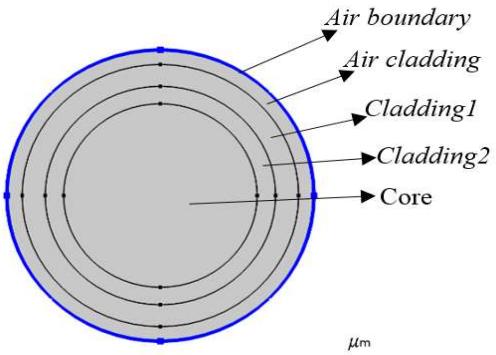


Fig. 2. Cross sectional view of the Multi-Layer Fiber model

The RI of Air boundary and Air cladding is 1, and the RI of the inner air cladding is respective…

cladding1(lightly doped silica) RI=1.45, Cladding 2 (highly dopped silica) RI = 1.48 and Core(silicon)RI=1.5

MLF can be enhance the signal capacity due to reflected rays from different RI of cladding layer so as using an optimization algorithm, multi-layer optical fibers with low dispersion values and a smooth dispersion slope in the wavelength range of 1.5-1.6 m were produced. For that reason, we provide more than one cladding concept as above given (fig-2) [6].

The dimensions given below table-2 show the dimensions of multi-layer fiber model:

TABLE II. DIMENSIONS OF MLF

Notation	Description	Dimensions
dcl	Air cladding diameter	125 [um]
Dcl 1	cladding diameter 1	118.3 [um]
Dcl 2	cladding diameter 2	111.6 [um]
dco	core diameter	105 [um]
wl	operating wavelength	1.55[um]

V. RESULTS AND DISCUSSION

The given matrix results in figure of multi-core and multi-layer fiber are showing the mode profile of model with various parameter sweeps. The COMSOL Multi Physics software is used here for mode profile analysis in various scenarios, but the COMSOL simulator can only analyze the single core of mode profile at a time.

A. *Result Summary of MCF*

In the results given below, the final results are represented in table-3 and table-4 with various parameter sweeps. the table's first column represents the mode profile with effects of various modifications in the effective mode index; the second column represents the mode profile at various core diameter; and the third column represents the mode profile at wavelength sweep.

In the matrix form of representation, showing the single parameter sweep at a time, the rest of the respective parameters are fixed.

B. *Result Discussion*

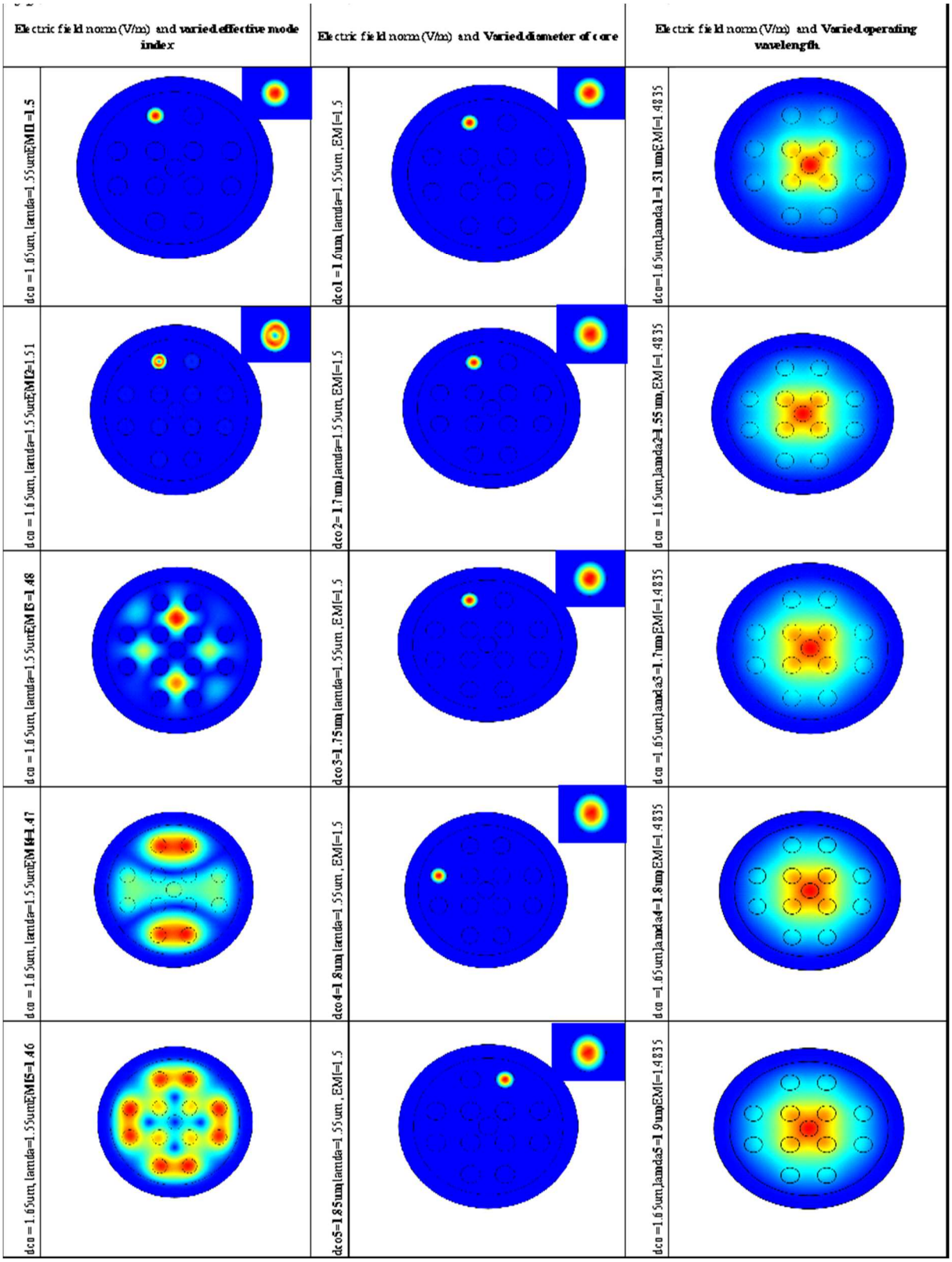
In the respective mode profile analysis, the multi-core fiber is mostly used for enhancing the signal carrying capacity with higher propagation length [15]. The MCF model is also an important topic which have require to analyze the power dissipation and dispersion losses. In future, we will implement the whole factor for designing purpose and also analyze each and every parameter such as losses and design parameters.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on February 13,2023 at 06:38:16 UTC from IEEE Xplore. Restrictions apply.

TABLE III. SIMULATION RESULT OF THE MODE PROFILE OF MULTI-CORE FIBER MODEL.

THE VARIATION OF EFFECTIVE MODE INDEX IN FIGURE-A AS SHOWING LP 01, FIG-B AS LP 02, FIG-C AS LP 21, FIG-D AS LP 11 AND FIG-E AS SHOWING LP 22,

IN THE (FIG-F TO FIG-J) VARIATION OF CORE DIAMETER, LP 01 MODE IS STABLE FOR ALL NEAREST VALUES, AND WHEN WE VARIED (FIG-K TO FIG-O) AS GIVEN VALUE THE WAVE LENGTH THE FIBER IS EXCITED ON CLADDING WITH LP 01



Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on February 13,2023 at 06:38:16 UTC from IEEE Xplore. Restrictions apply.

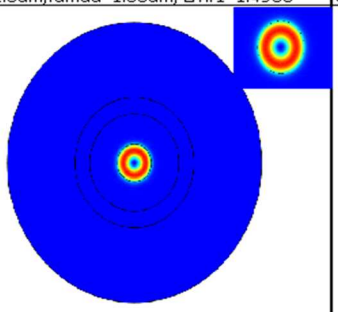
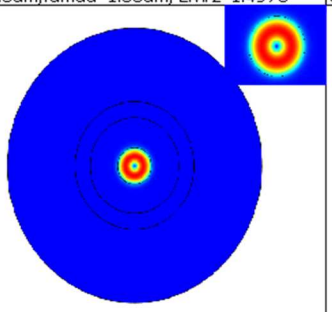
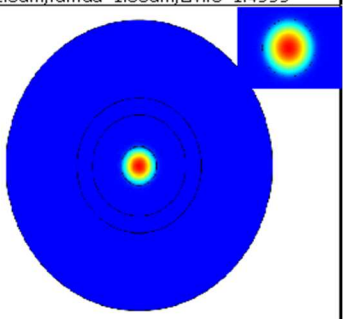
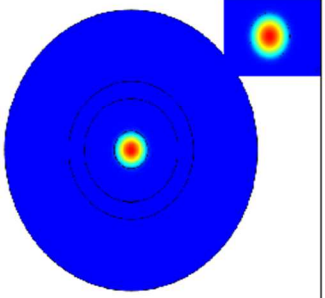
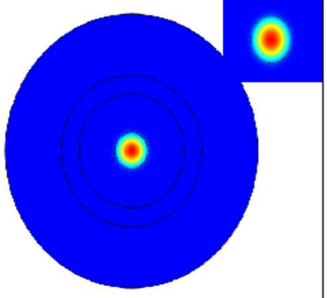
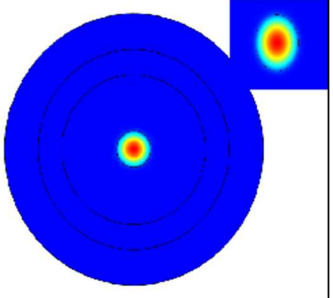
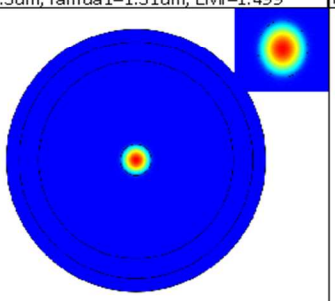
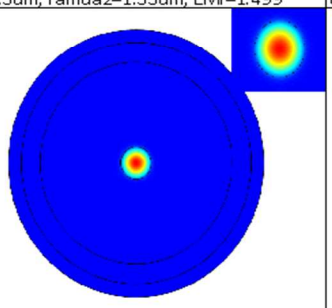
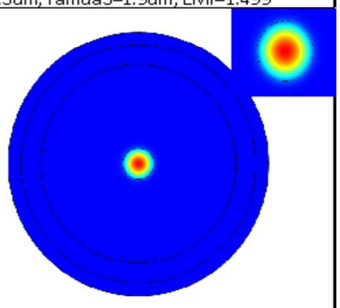
C. Result Summary of MLF

In the simulation, we evaluate the various parameter sweeps, and all simulated results are listed in the result table-4. In this simulation we analyze the mode profile is most of time the stable because the core diameter is fixed.

Here First row of fig (a) to fig (c) are showing the minor variation in mode profile according to various RI and other variation of radius and wavelength are approximately constant formation for mode profile in this multilayer fiber model.

TABLE IV. SIMULATION RESULT OF THE MODE PROFILE OF THE MULTI-LAYER FIBER MODEL

WITH THE VARIATION OF EFFECTIVE MODE INDEX IN FIG-A AND FIG-B AS SHOWING LP 02, FIG-C AS LP 01 AND AS SHOWN IN THE FIGURE ALL OTHER GIVEN VALUES WITH VARIATION OF CORE DIAMETER AND WAVELENGTH, THE LP 01 MODE IS STABLE BECAUSE THE CORE IS ALWAYS FIXED.

Miltilayer Fiber simulation model Mode Analysis Of Various parameter			
1	dcl=52.5um,lamda=1.55um, EMI1=1.4956	dcl=52.5um,lamda=1.55um, EMI2=1.4976	dcl=52.5um,lamda=1.55um,EMI3=1.4999
Electric field norm (V/m) varied effective mode index	 Figa	 Figb	 Figc
2	dcl1 =55.5um, lamda=1.55 um, EMI=1.499	dcl2=62.5um, lamda=1.55um,EMI=1.499	dcl3=183um, lamda=1.55um , EMI=1.499
Electric field norm (V/m) Varied diameter of core	 Figd	 Fige	 Figf
3	dcl=52.5um, lamda1=1.31um, EMI=1.499	dcl=52.5um, lamda2=1.55um, EMI=1.499	dcl=52.5um, lamda3=1.9um, EMI=1.499
Electric field norm (V/m) Varied operating wavelength	 Figg	 Figh	 Figi

D. Result Discussion

In this simulation, we analyze that the multi-layer fiber model has the minimum variation in mode profile. Due to constant core diameter and cladding not involve in mode profile representation [16]. So far, the multi-layer model is just an introductory part, but in the future, we will analyze the dispersion losses and power calculation and we expect that the MLF model effectively reduces the absorption losses [17].

VI. CONCLUSION

Here we show that both of fiber models have some specific mode profile at a particular value and here we analyze the mode profile at the various conditions with respect to the change of effective mode index, core diameter and operating wavelength.

In this analysis process, we show that the mode profile on some particular values is showing symmetric responses and, in some cases, they are responding to different types of mode profile. So, we conclude this type of fiber model is sufficient

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on February 13,2023 at 06:38:16 UTC from IEEE Xplore. Restrictions apply.

to achieve the appropriate response of mode profile as compared to the single mode, but few parameters affect the mode profile like effective mode variation and a little bit operating wavelength. And other hand we analyze the multi-layer fiber but we show that not a major different on mode profile due to fixed core diameter. Here basic motive for designing is that it could be reduce the absorption loses due to multi-layering. So, in future we analyze the losses with respect to data carrying capacity and environmental effects. And I justify that, this model design is more sufficient use as compared to single mode and multimode fiber.

REFERENCES

[1] B. Zhang et al., "Theory for mode coupling in perturbed fibers," Opt.Common., vol. 463, no. 125355, pp. 1–12, May 2020.

[2] Multi-Core Fiber Technology for SDM: Coupling Mechanisms and Design, Kunimasa Saitoh, Member, IEEE, Member, OSA JOURNAL OF LIGHTWAVE TECHNOLOGY, VOL. 40, NO. 5, MARCH 1, 2022 San Diego, CA, USA, 2020, Art. no. M4C.5.

[3] Design and Optimization of a Single-Mode Multi-Core Photonic Crystal Fiber with the Nanorod Assisted Structure to Suppress the Crosstalk'' Han Zhang, Guorui Wang, Jiwei Zhang, Fang Wang, Xin Yan, Xuenan Zhang, and Tonglei Cheng, IEEE PHOTONICS JOURNAL, VOL. 13, NO. 4, AUGUST 2021.

[4] H. Xiao, H. Li, B. Wu, Y. Dong, S. Xiao, and S. Jian, ``Elliptical hollowcore ptical bers for polarization-maintaining few-mode guidance," Opt. Fiber Technol., vol. 48, pp. 7_11, Mar. 2019.

[5] S. Chen, S. Li, L. Fang, A.Wang, and J.Wang, "OAM mode multiplexing in weakly guiding ring-core fiber with simplified MIMO-DSP," Opt. Exp.,

[6] F. Ren, ``Design of 20-polarization-maintaining-mode `pseudo-rectangle' elliptical-core _ber for MIMO-less MDM networks," Opt. Fiber Technol.,vol. 50, pp. 87_94, Jul. 2019.

[7] B. Batagelj, V. Janyani, and S. Tomazic: Research challenges in optical communications towards 2020 and beyond, Informacije MIDEM, ISSN 0352-9045.

[8] Y.Yamada et al., "Spatial mode dispersion control in a coupledMCFusing high density cabling parameters," in Proc. Opt. Fiber Commun. Conf.,

[9] K. Yoshida, T. Fujisawa, T. Sato, and K. Saitoh, "Optimum index profile of few-mode coupled multicore fibers for reducing the group delay spread," Opt. Express, vol. 27, no. 11, pp. 16281–16295, May 2019.

[10] Effects of the Intramode-Group Coupling Induced by the Elliptical Deformation of Orbital Angular Momentum Fibers Hu Zhang, Member, IEEE, Hui Li, Xiaoguang Zhang, Member, IEEE, Jingxuan Yang, Ze Chen, Lixia Xi, Wenbo Zhang and Xianfeng Tang IEEE PHOTONICS JOURNAL, VOL. 14, NO. 1, FEBRUARY 2022.

[11] Dispersion characteristics analysis of asymmetric multi-core fibers ling Gao1a, Xia Zhanga, Lei Shia, Weipeng Shia, Yongqing Huanga, Xiaomin Renaa State Key Laboratory of Information Photonics and Optical Communications (2011).

[12] H. Zhang, M. Bigot-Astruc, L. Bigot, P. Sillard, and J. Fatome, "Multiple modal and wavelength conversion process of a 10-Gbit/s signal in a 6-LPmode fiber," Opt. Exp., vol. 27, no. 11, pp. 15413–15425, May. 2019.

[13] R. Zhang et al., "Utilizing adaptive optics to mitigate intra-modal-group power coupling of graded-index few-mode fiber in a 200-gbit/s mode division - multiplexed link," Opt. Lett., vol. 45, no. 13, pp. 3577–3580, Jul. 2020.

[14] Seven-Core Optical Fiber Design and Fabrication for Space-Division Multiplexing Optimized for Low Crosstalk Ahmed Samir1, 2 and Bostjan Batagelj1, Member, IEEE, ICTON 2016. vol. 27, no. 26, pp. 38049–38060, Dec. 2019.

[15] J. Zhang et al., "Mode division multiplexed transmission of WDM signals over 100-km single-span oam fiber," Photon. Res., vol. 8, no. 7, pp. 1236–1242, Jul. 2020.

[16] R. Bhandari, "Analytic expressions for orbital angular momentum modal crosstalk in a slightly elliptical fiber," IEEE Photon. J., vol. 12, Dec. 2020, Art. no. 7202621.

[17] Design and Characterization of Nanopore-Assisted Weakly-Coupled Few-Mode Fiber for Simpler MIMO Space Division Multiplexing YUHENG XIE1,2, LI PEI1,2, JINGJING ZHENG1 2020.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on February 13,2023 at 06:38:16 UTC from IEEE Xplore. Restrictions apply.

Journal of Applied Fluid Mechanics, Vol. 16, No. 4, pp. 739-749, 2023.
Available online at www.jafmonline.net, ISSN 1735-3572, EISSN 1735-3645.
https://doi.org/10.47176/jafm.16.04.1450

Design and Simulation of the Spiral Micromixer with Chaotic Advection

R. Prakash†, M. Zunaid and Samsher

Department of Mechanical Engineering, Delhi Technological University, Delhi, 110042, India

†*Corresponding Author Email: ranjanprakash_2k18phdme16@dtu.ac.in*

(Received July 30, 2022; accepted November 26, 2022)

ABSTRACT

In recent years, the microfluidics technique has screwed up rising attention and be in progress a fascinating topic. Species mixing is a compelling part of any microfluidic system that abides by the major challenge. In this research, a relative explication of mixing quality of microchannels of two cross-sections square and circular in spiral form is presented by numerical simulation. To perform the visitation, geometric parameters like axial length of microchannels and hydraulic diameter are taken equal for both cases. Computational Fluid Dynamics codes unriddle the Continuity equation, Navier-Stokes equation, and Convection-Diffusion equation. Explication of fluid flow and mixing have been gone through with an extensive limit of Reynolds numbers 1 to 125. The results explicate that the circular section spiral microchannel affords a higher mixing quality admixed to the square section spiral microchannel. Furthermore, in the circumstance of circular section spiral micromixer, mixing index esteem has attained 92% at Re =125. For both section micromixer, the esteem of mixing index enhancement be contingent on Reynolds numbers. For both cases, pressure downfall has been computed for microchannels of similar lengths. The esteem of pressure downfall in square section spiral mixer is excess than circular section spiral mixer. The simulation outcomes exhibited that the circular section spiral micromixer is an effective design for microfluidic devices like Lab on a chip (LOC).

Keywords: Micromixing; Micromixer; Chaotic advection; Microfluidics; Simulation investigation.

NOMENCLATURE

c	mass fraction	V	velocity
D	diffusion coefficient	μ	fluid viscosity
M	mixing index	ρ	fluid density
p	pressure	σ	concentration variance
Re	Reynolds number		

1. INTRODUCTION

Microfluidics is the science that deals with the demeanor of fluids using micromixers and the technology of generating microminiaturized devices like Lab on a chip (LOC), giving various important advantages such as the use of the small volume of reagents compared to the large conventional mixer that results in huge saving of costly reagents per analysis, improved portability, low manufacturing cost, accurate measurement. Micromixer is the basic parts of microfluidics devices, which have a huge influence on biomedical diagnosis, medicine evolution, and chemical diligence for lab research like DNA Test, sample creation and study, cell separation, enzyme analysis, and protein folding. Also, a micromixer plays a vital role in the mixing of more than one sample for analysis. A nicely designed mixer can decrease the analysis time and the footmark of a lab on a chip device. The mixer integrated with a microfluidic device like organ on chip will be used in the treatment of chronic diseases in the future.

Micromixers can be ranked as passive and active micromixers. Passive micromixers do not solicit outer impediments to raise mixing. Passive mixing relies on diffusion, split and recombination, and chaotic advection. Based on the order of mixed phases, passive mixing can be again ranked as serial lamination, parallel lamination, injection, droplet mixing, and chaotic advection. Active micromixers use outer impediments to speed up the mixing. Based on kinds of impediments, an active mixer can be ranked in temperature-induced, pressure-driven, electro-hydrodynamic, magneto-hydrodynamic,

electrokinetic, dielectrophoretic, acoustic concepts, electro-osmotic effect (arrangement of zeta-potential on walls), and piezo-electrically force. In a temperature-induced active mixer, fluids at varying temperatures are mixed in microchannel and a temperature gradient was created using two Peltier modules in the junction of a microchannel. Pressure-driven active micromixers are simple in design and construction. In this design, inject the fluid into the mixing channel alternately with the help of two micropumps generating a pulsatile flow. Making this arrangement rises the interfacial area outcoming in enhanced mixing quality. In electro-hydrodynamic (EHD), mixers use the movement of electrically charged fluids in an either Alternating Current (AC) electric field or a Direct Current (DC) electric field to unsettle the interface of mixing fluids that enhance diffusion species transport. In a magneto-hydrodynamic micromixer, an external magnetic field applied dc voltages on the electrodes can create Lorentz forces, that can fold and roll the fluids in mixing channel and make the fluids mix completely. An electrokinetic mixer utilizes electrokinetic flow to transport fluid. Electro-kinetic mixing is affected through a unique voltage source accompanied by channels dimensioned to take the hoped-for voltage segmentation. A novel dielectrophoretic micromixer is utilizing parallel microelectrodes to exploit the polymer particles and generate a vortex that can bustle the fluid to obtain mixing. Acoustically driven micromixer utilizes Surface Acoustic Waves (SAW). The acoustic wave is passed to the solid surface of mixer. The sound stream enhances the bustling effect which improves the mixing quality. In electro-osmotic micromixing, analysis of flow and mixing performance in a two- dimensional microchannel having nonuniform zeta potential on channel walls is accomplished. A piezo-electrical force-driven micromixer needs the erection of an interposed piezoelectric actuator and the wiring of a piezoelectric actuator for the external power supply.

Efficient and intense mixing is much important for these exercises. The mixing process is a difficult process for low Reynolds number flow inside the micromixer. Simple T-junction and Y-junction micromixing mainly depend on molecular diffusion, which is much slower. So it is essential to utilize especially drafted micromixers to speed up the mixing process.

In this context, Azmi *et al.* (2021) deciphered the effect of the hybrid composition ratio for wire coil flow. Balasubramaniam *et al.* (2017) reported the enhancement of mixing quality in spiral micromixers of different cross-sections (rectangular, square, semi-circular, trapezoidal) and hydraulic diameters. Chen and Chen (2019) collated mixing index of circular micromixer and topology-based passive micromixers. Cao *et al.* (2021) developed a model to describe liquid-gas flow in an inclined and horizontal pipe. Dong *et al.* (2020) visited the impression of DC and AC electric fields on the mixing index in a microchannel. Farahinia *et al.* (2021) presented a simulation analysis of electro-osmotic flow in heterogeneous micromixers and also examined micromixer performance with combined effects of electro-osmotic and pressure-driven flows. Farahinia *et al.* (2020a) contributed a comprehensive review of various separation systems of Circulating Tumor Cells engendering from the primary tumor or cancer tissue. Nowadays some researchers are using microfluidic devices for Circulating Tumor Cells separation from the blood cells. Lab on a chip (LOC) and then Lab on a CD (LOCD) are the recent trends. Farahinia *et al.* (2020b) presented two cross-sections of T –Junction microchannel namely circular cross-section and rectangular cross-section and studied the effects of input angles on mixing efficiency. Farahinia and Zhang (2019) studied the optimal design of a passive T micromixer having assigned patterns of blockades and grooves with various geometries. Ghazimirsaeed *et al.* (2021) explored the impact of rectangular cross-section serpentine microchannel with three aspect ratios and another three cross-section geometry like circular, triangular, and trapezoidal on mixing performance. Hasanah *et al.* (2022) designed micromixers having Koch Fractal Obstacle and studied the effect of two different inlet conditions and resistance angles. Hossain *et al.* (2009) formulated the problem of three different microchannels and analyzed the pressure downfall and mixing quality of these microchannels. Jiang *et al.* (2022) explicated the various shapes of cross-sections like triangular and elliptical and inserted different obstacles like hexagonal, triangular, and cylindrical in variable radius spiral microchannel. Kockmann *et al.* (2003) analyzed liquid mixing in simple static mixers having different cross-sections like rectangular, square, and trapezoidal.

Li *et al.* (2012) explicated the impression of a small width outlet and turning angle in a zigzag microchannel. Liu *et al.* (2022) proposed a microchannel transacted by two piezoelectric micropumps having multi-stage mixing. This micromixer is combination of novel texture and piezoelectric drive technique to create a novel mixing technique, which provides excellent mixing quality. Liu *et al.* (2018) designed and created a self-circulation microchannel having high-frequency vibration rely on the piezoelectric drive technique. Piezoelectric lead-zirconate-titanate (PZT) micropumps integrated with microchannel are used to generate self-circulation reflux in the mixing chamber and give a better working performance with compelling frequency of 80-120 Hz. Mondal *et al.* (2021) created a serpentine microchannel in a square waveform by the use of the WEDM process. Nezhad *et al.* (2018) offer baffles in three different forms in the microchannel and observed these baffles effects on the mixing index. Nouri *et al.* (2017) explicated the influence of ferrofluid and water mixture on mixing index by use of a magnetic field. Prakash *et al.* (2021) inspected the impression of a T-bend mixing channel at various radii. Pang *et al.* (2020) presented a T- form microfluidic device with a neck at the junction and evaluate droplet volume under several flow rates. Quiyoom *et al.* (2017) deciphered the free surface effect on liquid-phase mixing in shallow vessels. Redapangu *et al.* (2021) analyzed the effect of changing density and viscosity of two immiscible liquids in a 3-D inclined

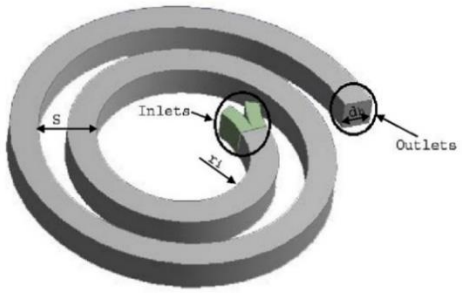
microchannel. Ruijin *et al.* (2017) likened mixing quality of three mixers Baker, Smale, and Helical. Rouhi *et al.* (2021) studied the cross-sections effects on mixing quality of spiral microchannel in Reynolds number limit from 0.001 to 50. Santana *et al.* (2019) examined the impose of circular form obstructions and triangular form baffles in the micromixer with ethanol and oil mixture. Seo and Kim (2012) designed a passive and active combination micromixer and collated the mixing index esteem of passive micromixer with various obstacles and active micromixer under an external electric field. Saleel *et al.* (2020) analyzed the impose of changing the height of a rectangular block in a micromixer in the concern of Electro-osmotic flow. Shinde *et al.* (2021) examined mixing quality of SARM having U-form, V-form, and square C-form pillars located in a microchannel. Shimizu and Uetsuji (2022) proposed two drive modes single and double Metal-capped bimorph actuator (MC-BMP) and compared the mixing index and flow rate with conventional BMP (Bimorph). Tokas *et al.* (2021) monitored the effect of working fluid as blood in a T-helical microchannel. Tony *et al.* (2021) generalized the concept of Integrated Microfluidic Circuit (IMC) to the conception designated soft microfluidic system (SMS). It includes a soft robotic microfluidic system, which can show robotic behaviors, like movement, control, sensing, and actuation. Umadevi *et al.* (2021) introduced a study about blood and copper nanoparticles in a water mixture in an inclined artery under an external magnetic field. Vatankhah and Shamloo (2018) calibrated the mixing quality of helical-type microchannels.

Wu *et al.* (2021) anatomized the effect of swirl motion in a jet mixer for liquid mixing. Xia *et al.* (2016) explored the impose of gaps and baffles at different positions in a planar micromixer. Xu *et al.* (2011) read out the fluid motion and pressure drop in a staggered oriented ridges micromixer. Yang *et al.* (2022) reviewed Microfluidic Point of Care (POC) instruments and their uses. Microfluidic POC tools can hurriedly retrace diseases at a minimum cost. This technique is used to find out diseases in immature regions to decrease the impacts of disease and enhance the potency of life in these regions. Zhang *et al.* (2019) created fractal baffles in a T microchannel and likened the mixing quality of primary and secondary fractals.

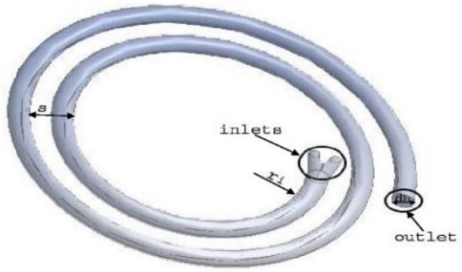
Contemplating the literature exploration, in this paper square section spiral micromixer and circular section spiral micromixer are explicated and similitude the mixing index of square section spiral micromixer and circular section spiral micromixer. Then the impose of Reynolds number is examined for both cross-sections. The outcomes report that a circular section spiral mixer is more verseed than a square section spiral mixer. The advantages of micromixer proposed in this study are that better mixing quality than simple T junction, Y junction, and serpentine micromixers. Also presented circular section spiral micromixer is a better configuration than a square section spiral micromixer for microfluidic devices namely Lab on a chip (LOC).

2. DELINEATION OF THE MICROMIXER

Square section spiral micromixer and circular section spiral micromixer are explicated in Fig. 1. Two configurations are square and circular with two inlets in which water is flowing from one inlet and the water-soluble dye is flowing through the second inlet. For the relative study, hydraulic diameter of the mixing channel and inlets channel of both configurations was elected equal. Geometry descriptions are provided in Table 1.



(a)



(b)

Fig.1 Microchannels of two cross-sections (a) square (b) circular.

Table 1 Configuration mensuration of square section spiral micromixer and circular section spiral micromixer.

Type	Square section spiral micromixer	Circular section spiral micromixer
Hydraulic diameter of mixing channel	0.2 mm	0.2 mm
Hydraulic diameter of inlets channel	0.133 mm	0.133 mm
The distance between two successive turnings, s	0.4 mm	0.4 mm
initial microchannel radius, r_i	1 mm	1 mm
Number of turns, N	2	2
The total spiral length	15 mm	15 mm

In this paper, the spiral curvature of varied cross-sections is introduced so in this case, secondary flow creates, which tends to interact between centrifugal and inertia forces. The centrifugal forces poke the liquid from the centre of the microchannel to an outer wall, where the liquid has less velocity. At the outside wall, a double vortex creates by squeezing the liquids and parenting the liquids move in downward and upward directions. On the other side, the viscid wall friction moistens both vortexes by working as opposed to centrifugal force, which relies on average liquid velocity and curvature radius. As stated by Dean, the flow dominations and creation of vortices in a laminar region with a small Reynolds number in curvature rely on the dimensionless number named Dean Number. It is interpreted as:

$$\mathrm{De} = \mathrm{Re}\sqrt{\frac{D_h}{2R}} \tag{1}$$

Where R is the microchannel curvature radius, Re is the Reynolds number, and D_h is the hydraulic diameter. If the value of De is less than one, the mixing of two fluids is primely obtained by diffusion. If Dean number is greater than ten, the first double vortex creates and improves micro mixing.

Reynolds number is demonstrated as the proportion of inertia force to the viscous force. Mathematically, it is expressed as:

$$\mathrm{Re} = \frac{\rho v\, D_h}{\mu} \tag{2}$$

Where ρ is density of fluid, v is velocity of fluid, and μ is absolute viscosity of fluid.

The hydraulic diameter D_h is evaluated by equation (3), as exhibited below:

$$D_h = \frac{4A}{P} \tag{3}$$

Where A is a cross-section area and P is the perimeter of the cross-section of the mixing channel.

3. SIMULATION ACCESSORY

To anatomize the mixing process and flow in the microchannels, the Computational Fluid Dynamics (CFD) tool Ansys Fluent has been utilized. By application of finite volume technique, Fluent code elucidates the momentum and continuity equations. For computing the fluid motion, Fluent-solver will figure out the adequate mean values of the characteristic of every control mass in the stream zone. These mean esteems rely on magnitudes of constituent characteristics and the rates of every constituent which is extant in the control mass. In the concern of two or more fluids, the retention is rendered that fluids are mingled at the molecular scale and characteristics of fluids rely on similitudes of components. The embracement is as well created that mass fraction pullulates through diffusion and convection. The differential movement of particular constituents in the species mixture is numerated by genitive mass flux. This mass flux is imitated by the impression of concentration gradients, pressure gradients, etc. The governing equations are illustrated in Shinde *et al.* (2021) and indicated in equations (4) to (6) as:

$$\nabla v = 0 \tag{4}$$

$$\rho\left[\frac{\partial v}{\partial t} + (v.\nabla)v\right] = -\nabla p + \mu\nabla^2 v \tag{5}$$

$$\frac{\partial c}{\partial t} = D\nabla^2 c - v.\nabla c \tag{6}$$

Where, viscosity μ, density ρ, diffusion coefficient D, velocity v, pressure p, mass fraction c, and Del operator∇.

$$\nabla = \frac{\partial}{\partial x}\hat{\imath} + \frac{\partial}{\partial y}\hat{\jmath} + \frac{\partial}{\partial z}\check{\mathrm{k}}$$

A scant quantity of dye diluted in water and pure water is utilized as two operative fluids for mixing. Since the dye is diluted in water in a very small quantity, so characteristics of both water and water dye will be approx. same. The characteristics of water and water-dye have been selected at 20°C and values are taken as, the diffusivity for mixture (D) is 10^{-10} m^2/s, the viscosity of water (μ) is 0.001Ns/m^2, the density of water (ρ) is 1000 kg/m^3. The same inlet velocity and zero gauge outlet pressure are elaborated as the boundary conditions. Water and water with dye are allocated to inlet1 and inlet2 with mass fraction zero and one respectively. The no-slip circumstances were enforced on the parapets. The results are recognized to have reached convergence when the residual esteem is maximally 10^{-8}.

The mixing quality is admeasured by numerating the concentration variance of composition in a microchannel. To estimate intensity of tincture in a microchannel, a variance of concentration of composition at cross-section is determined as presented in Hossain *et al.* (2009):

$$\sigma = \sqrt{\frac{1}{N}\sum(C_i - C_m)^2} \tag{7}$$

In above elucidation, N is total sampling point at the cross-section, C_i is the concentration at sampling point, i, and C_m is mean of the concentrations. In this affair, N on every plane was 500 to maintain high accuracy. To examine the mixing efficiency of microchannel, mixing index is determined as:

$$\mathrm{M} = 1 - \sqrt{\frac{\sigma^2}{\sigma_{max}^2}} \tag{8}$$

Where σ is standard deviation value at cross-section and σ_{max} is maximum standard deviation. A supreme mixing index denotes a supreme mixing quality. For separate streams mixing index esteem is zero and for fully mixed streams mixing index esteem is one.

4. RESULTS AND DISCUSSION

The mixing index is enhanced by geometry transformation. For this goal, two microchannels of

varied cross-sections square and circular have been simulated and admixed the mixing index results of both channels square and circular cross-sections. As well scrutinize the impose of Reynolds numbers on pressure drop and mixing index.

4.1 Mesh Generation and Grid Sensitivity Analysis

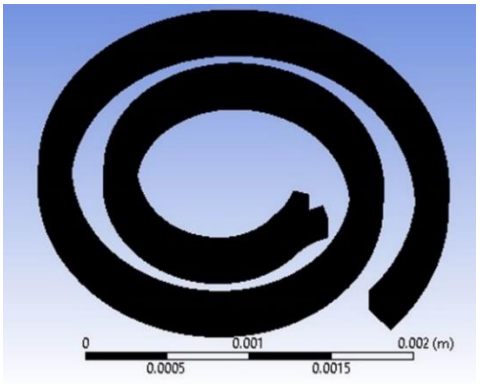
Mesh generation in a square cross-section and circular cross-section spiral micromixer with tetrahedral type mesh is presented in Fig. 2(a) and Fig. 2(b). To search out the best grid elements, this is essential to calibrate that the rectification is grid-independent. Five diverse mesh systems 1.9 million, 3.9 million, 5.1 million, 5.5 million, and 6.3 million were tested for both microchannels. Numerated mixing index at the outlet is revealed in Fig. 2(c), as can be looked over in this diagram, the numerated mixing index reduces with increasing the number of grid elements. This is due to less molecular diffusion in the case of increasing the number of grid cells. However, mixing index numerated with 5.5 million and 6.3 million give no more significance difference. So from grid independency, 5.5 million was elected as the best grid for all simulations.

4.2 Validation of computational work

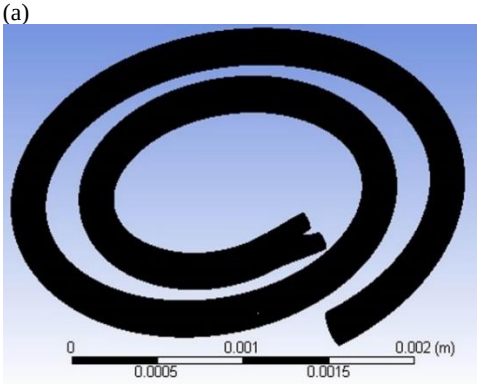
The existing simulation work is admixed with the numerical outgrowths of Vatankhah and Shamloo (2018) for the same Reynolds number and microchannel length as explicated in Fig. 3. The mixing index versus Reynolds number graph has been plotted for the outlet of a microchannel. The outgrowths exhibit compatible consent between the existing workout and Vatankhah and Shamloo (2018) work with a maximum percentage error of 5.7.

4.3 Sequel of modified geometry

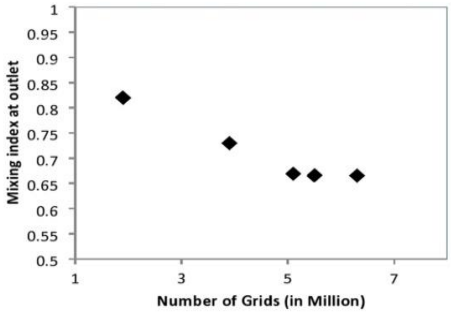
To interpret the simulation results, mixing index of simulations is compared at the outlet of square and circular cross-section spiral micromixer. The dispersions of concentration of microchannel of square cross-section and circular cross-section at Reynolds number 125 are explicated in Fig. 4. This concentration contour explicates that mixing quality in circular cross-section microchannel is of high quality than square cross-section microchannel. To illustrate the influence of circular cross-section on mixing quality, the power of vortices created inside a cross-section can be scaled by swirling strength, which is defined as the proportion of the tangential velocity of the fluid to the main velocity along a plane. Because of creation of vortices inside micromixers, velocity created at the centric region isolating the two vortices is much improved, outcoming in supreme swirling strength in a spiral or curved microchannels in similitude to a straight microchannel. It is got that circular cross-section microchannel obtain supreme mixing quality than square cross-section microchannel because of their enhanced swirling strength. Also, the mass fraction variations in circular cross-section microchannel at



(a)



(b)



(c)

Fig. 2: (a) Tetrahedral mesh of square cross-section spiral micromixer (b) Tetrahedral mesh of circular cross-section spiral micromixer (c) Numerated mixing quality at outlet and number of the grid element.

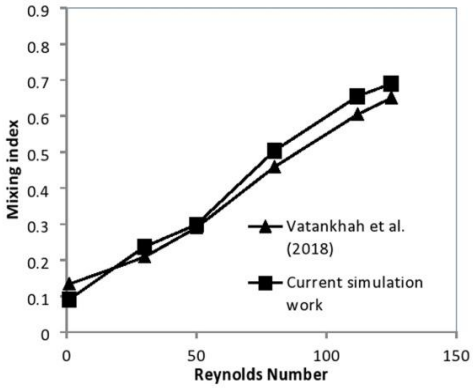
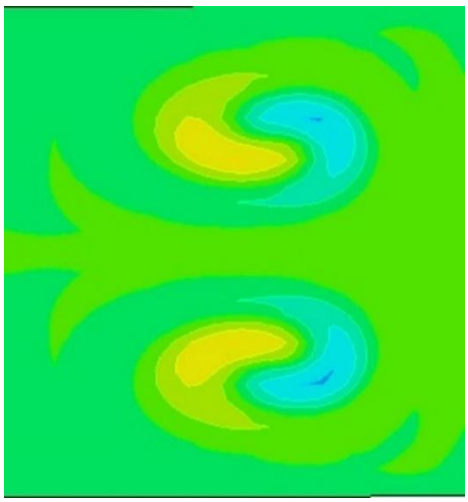
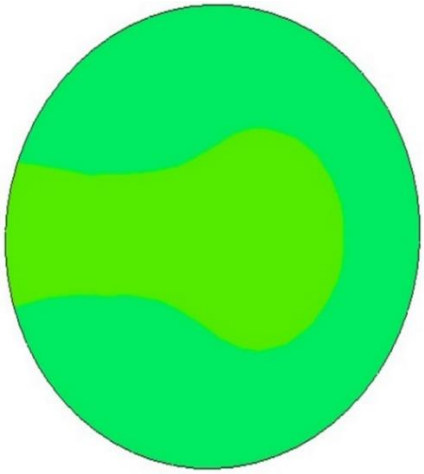


Fig. 3 Collation of mixing index of current square section microchannel and previously published data.



(a)



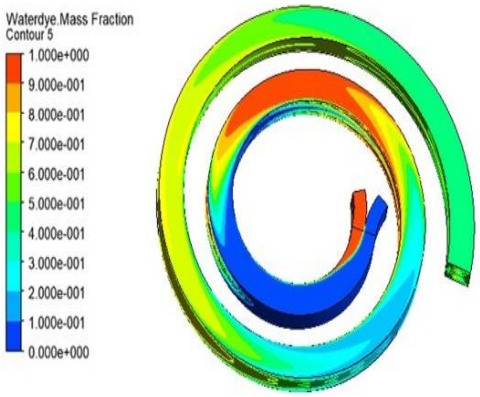
(b)

Fig. 4 Concentration contours of microchannel (a) square cross-section (b) circular cross-section at Re = 125.

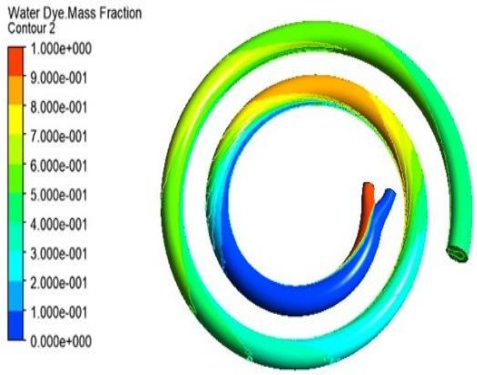
an outlet achieves near to optimal mixing mass fraction esteem i.e 0.5 mol/m^3 but in the case of square cross-section microchannel, mass fraction variations at an outlet are in a wide range. So we require extra microchannel length and extra time to get the same outcomes as circular cross-section microchannel. The worth of mixing index at an outlet for a circular cross-section mixing channel is (M= 92.95) while for a square cross-section mixing channel is (M = 84.20).

Figure 5 explicates a mass fraction contour along the helical length in square cross-section and circular cross-section microchannel. The figure displays that at this high Reynolds number, mixing rises consistently from the junction point to the outlet. The red color reports one fluid (water) and the blue color reports another fluid (dye diluted in water). When mixing has occurred, delineates green color.

The mixing is more influenced through secondary flow generated due to centrifugal force in the case of circular cross-section microchannel as compared to square cross-section microchannel. So chaotic advection is more effective in circular cross-section spiral micromixer as compared to square cross-section spiral micromixer, which prominently improves mixing. This incidence can be seen by streamlines rely on mass fraction. The streamlines rooted on mass fraction ubiquitously the helical length in square and circular cross-section microchannel are explicated in Fig. 6. These streamlines illustrate the deviation in mixing quality pattern executed amid square and circular section microchannel. A more advanced perusal was executed to scrutinize the alteration of mixing quality of square and circular cross-section microchannels at different part length along the direction of flow. At Reynolds number = 125, the mixing quality rapidly increases from 14.6 % at a distance of 1mm to 77.2 % at a distance of 10mm in the concern of square cross-section spiral micromixer and 22.6% at a distance of 1mm to 90.3 % at a distance of 10mm in concern of circular cross-section spiral micromixer intimating the benefit of implementing circular section microchannel for high Reynolds number. The concentration contours at four positions starting from the junction point for square cross-section microchannel and circular cross-section microchannel are explicated in Fig. 7 and Fig. 8 delineates the velocity vector plots at outlet of microchannel square and circular cross-section. At Reynolds number = 125, the flow sample is nearly identical in both channels, and velocity vectors are curvilinear to a microchannel wall throughout the cross-section.

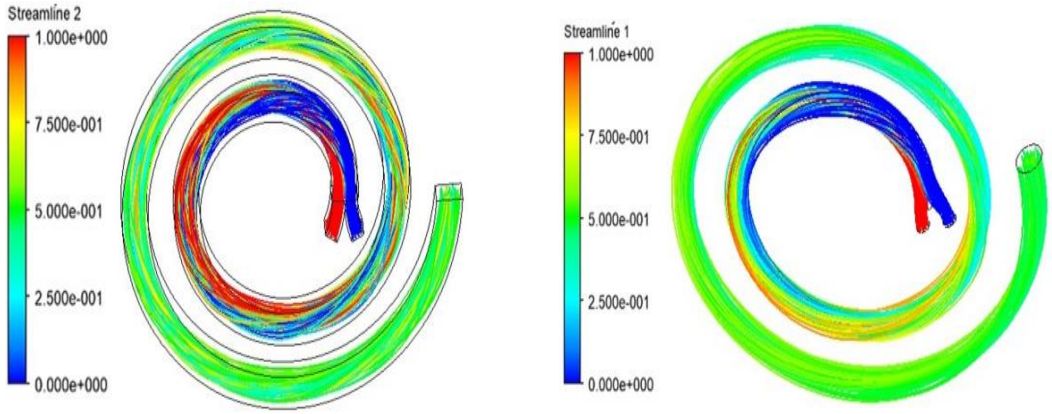


(a)



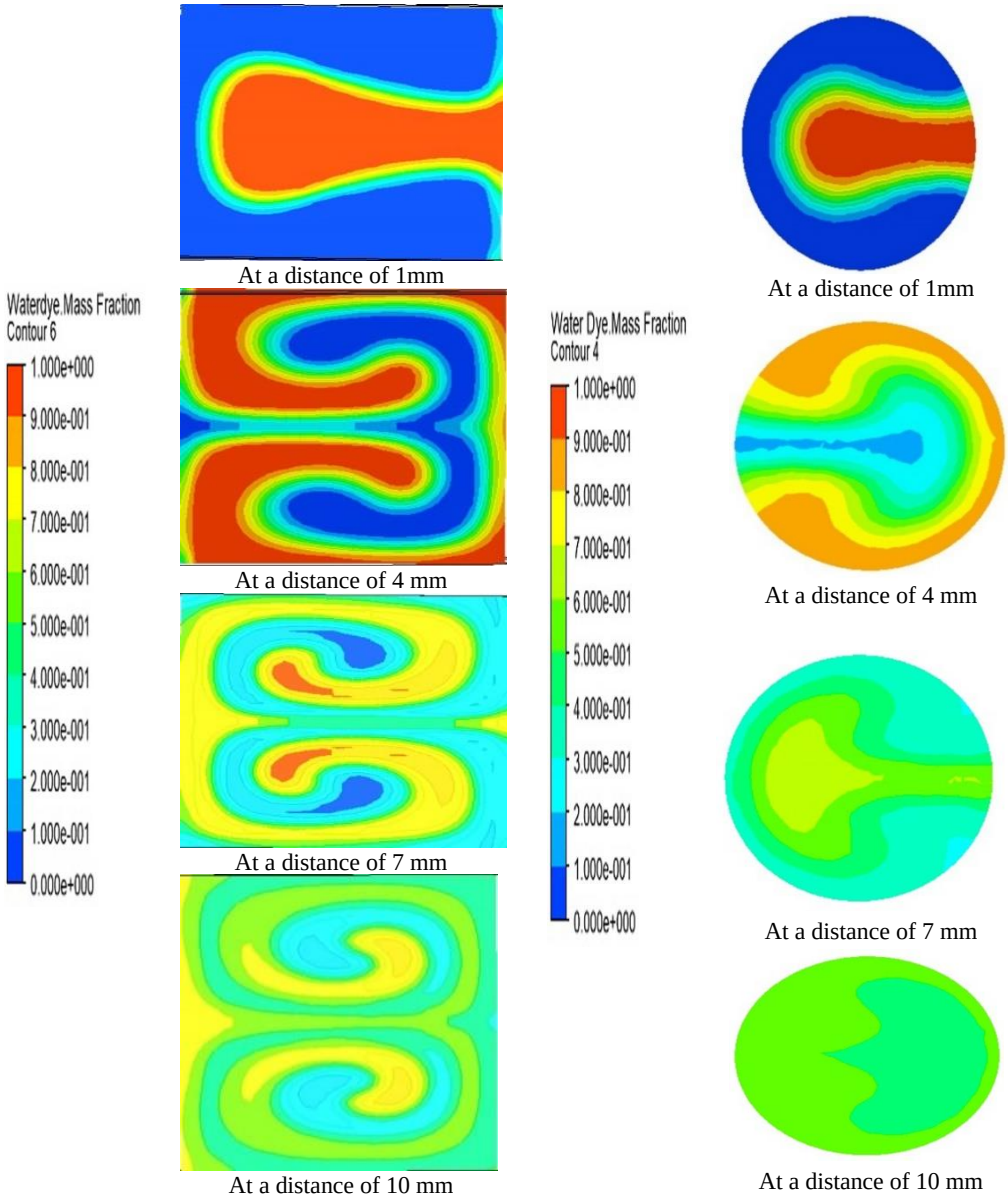
(b)

Fig. 5 Mass fraction contour along helical length for (a) square cross-section microchannel (b) circular cross-section microchannel at Re = 125.



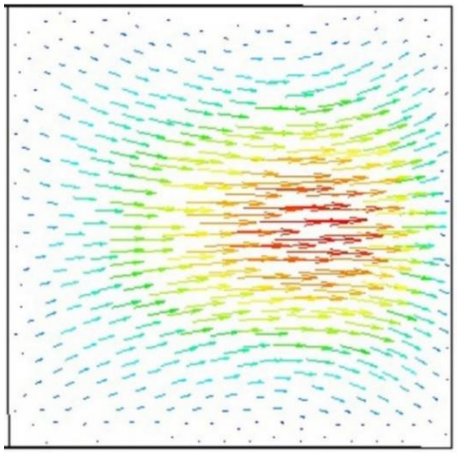
(a) (b)

Fig. 6 Streamlines rooted on species concentration for microchannel of (a) square cross-section (b) circular cross-section at Re = 125.

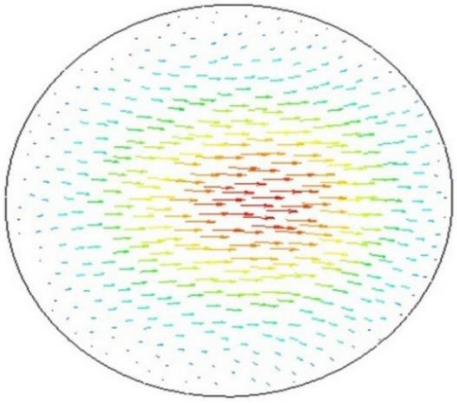


(a) (b)

Fig. 7 Contour of species concentration at four different locations starting from junction point for microchannel of (a) square cross-section (b) circular cross-section at Re = 125.



(a)



(b)

Fig. 8 Velocity vectors plot for microchannel of (a) square cross-section (b) circular cross-section at Re = 125.

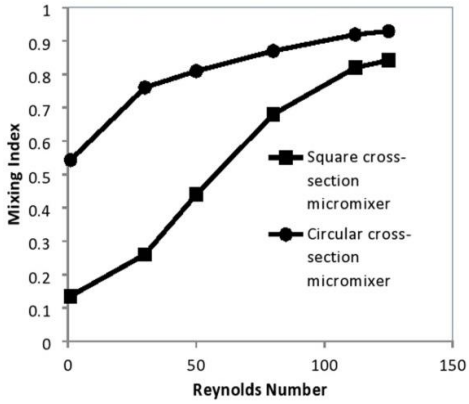


Fig. 9 Mixing index values for square and circular cross-section microchannel at various Reynolds numbers.

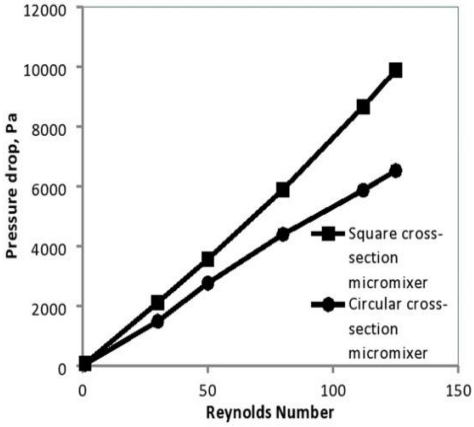


Fig. 10 Alteration of the pressure fall with Reynolds number for both geometry.

4.4 Imposingness of Reynolds Number

The Reynolds number imposingness on the mixing quality in both microchannel square and circular cross-sections was deciphered. The simulation process was performed with Reynolds numbers (1, 30, 50, 80, 112, and 125) for square cross-section microchannel and circular cross-section microchannel having a fixed length.

Figure 9 explicates the plots between mixing index and Reynolds numbers in which circular cross-section microchannel have better mixing quality.

The interpretation of mixing continual solicits the perusal of pressure drop amid inlet and outlet of mixing channel and this happens more valuable in concern of passive mixers because the pressure falls instants an opinion of the power which will be needed to pump the fluids from inlet to outlet. So, it is necessary to examine the pressure drop in a microchannel. The mixing process uses energy inputs in terms of pressure drops. Figure 10 delineates alteration of pressure downfall for the two configurations at multifarious Reynolds numbers. For both concerns, the pressure falls have been worked out for microchannels with the same streamwise lengths. The pressure falls rise suddenly with the Reynolds number for both the microchannel. The circular cross-section microchannel explicates lesser pressure drop than the square cross-section microchannel because there are no sharp edges in the circular cross-section microchannel. Figure 11 explicates dispensations of mixing index along streamwise length at Reynolds number 125. The mixing rises well-orderly along microchannel length as delineated in this diagram.

5. CONCLUSIONS

In this work, the numerical perusal of mixing in square and circular cross-section microchannels has been exploited. Dye diluted in water solution and pure water have been elected as working fluids for mixing. Navier-Stokes in three-dimensional explores flow field and mixing for Reynolds numbers 1 to 125. It was educed that circular cross-section microchannel explicates a superior mixing quality than square cross-section microchannel. In both microchannels, the pressure downfalls suddenly rise with the Reynolds number. The circular cross-section microchannel shows smaller pressure drops than the square cross-section microchannel. The tenor of this exploration is to get two objectives which are

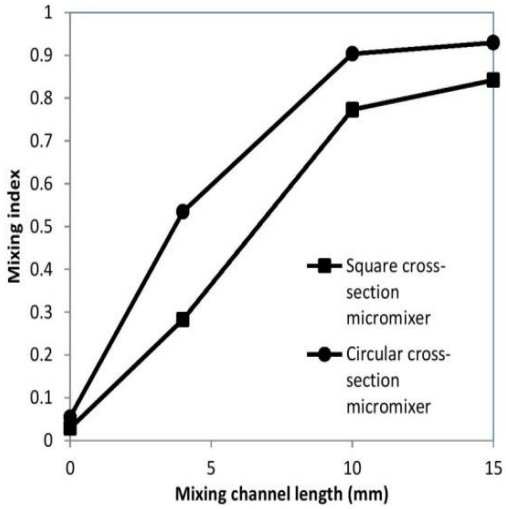


Fig. 11 Dispensations of mixing index along streamwise length at Re =125.

supreme mixing quality and small pressure drops in the microchannels. The circular cross-section mixing channel has fulfilled both criteria, so this cross-section is better for fluids mixing. In actual practice, the operation of a Lab on a chip system is not affected through pressure drop. The design plays an important role, so this circular cross-section mixing channel has a good potential for use in lab on a chip device. To design the next generation micromixer, the phenomenon of cavitation inside the micromixer will be considered in the future. Also, future studies will be focused on computational analysis of mixing of different fluids like blood mixture, water-ethanol mixture, water-ink mixture, etc because mixing quality depends on the diffusion coefficient of fluids mixture.

ACKNOWLEDGEMENTS

This work was encouraged by Delhi Technological University, Delhi.

CONFLICT OF INTEREST

The authors reveal that there is no conflict of interest regarding the publication of this paper.

REFERENCES

Azmi, W. H., K. A. Hamid, A. I. Ramadhan and A. I. M. Shaiful (2021). Thermal hydraulic performance for hybrid composition ratio of Tio2- Sio2 nanofluids in a tube with wire coil inserts. *Case Studies in Thermal Engineering* 25, 100899.

Balasubramaniam, L., R. Arayanarakool, S. D. Marshall, B. Li, P. S. Lee and P. C. Y. Chen (2017). Impact of cross-sectional geometry on mixing performance of spiral microfluidic channels characterized by swirling strength of Dean-vortices. *Journal of Micromechanics and Microengineering* 27, 095016.

Cao, X., K. Yang, T. Li, N. Xiong, W. Yang and J. Bian (2021). Numerical simulation of hydrate particle behaviors in gas-liquid flow for horizontal and inclined pipeline. *Case Studies in Thermal Engineering* 27, 101294.

Chen, Y. and X. Chen (2019). An improved design for passive micromixer based on topology optimization method. *Chemical Physics Letters* 734, 136706.

Dong, S., P. F. Geng, D. Dong and C. X. Li (2020). Mixing enhancement of electroosmotic flow in microchannels under DC and AC electric field. *Journal of Applied Fluid Mechanics* 13(1), 79-88.

Farahinia, A., J. Jamaati, H. Niazmand and W. Zhang (2021). The effect of heterogeneous surface charges on mixing in a combined electroosmotic/pressure-driven micromixer, *Journal of the Brazillian Society of Mechanical Sciences and Engineering* 43, 497.

Farahinia, A., W. J. Zhang and I. Badea (2020a). Circulating Tumor Cell Separation of Blood Cells and Sorting in Novel Microfluidic Approaches: a review. *Journal of Science: Advanced Materials and Devices* doi: 10.20944/preprints202010.0622.v1.

Farahinia, A. and W. J. Zhang (2020b). Numerical analysis of a microfluidic mixer and the effects of different cross-sections and various input angles on its mixing performance. *Journal of the Brazillian Society of Mechanical Sciences and Engineering* 42, 190.

Farahinia, A. and W. J. Zhang (2019). Numerical investigation into the mixing performance of micro T-mixers with different patterns of obstacles. *Journal of the Brazillian Society of Mechanical Sciences and Engineering* 41, 491.

Ghazimirsaeed, E., M. Madadelahi, M. Dizani and A. Shamloo (2021). Secondary flows, mixing, and chemical reaction analysis of droplet based flow inside serpentine microchannels with different cross-sections. *Langmuir* 37(17), 5118-5130.

Hasanah, L., D. F. Nurvadila, R. E. Pawinanto, B. Mulyanti, C. Wulandari, A. Aminudin and J. Yunas (2022). A passive micromixer with Koch Snowflakes Fractal obstacle in microchannel. *Journal of Applied Fluid Mechanics* 15(5), 1333-1344.

Hossain, S., M. A. Ansari and K. Y. Kim (2009). Evaluation of the mixing performance of three passive micromixers. *Chemical Engineering Journal* 150, 492-501.

Jiang, Y., B. Zhang, Y. Tian, Y. Zhang and Y. Chen (2022). Analysis of different shapes of cross-sections and obstacles in variable-radius spiral micromixers on mixing efficiency. *Chemical Engineering and Processing Process Intensification* 171, 108756.

Kockmann, N., M. Engler, C. Fo''ll and P. Woias (2003). Liquid mixing in static micromixers with various cross-sections. In *Ist International Conference on Microchannels and Minichannels* 911-918.

Li, Y., D. Zhang, X. Feng, Y. Xu and B. F. Liu (2012). A microsecond microfluidic mixer for characterizing fast biochemical reactions. *Talanta* 88, 175-180.

Liu, G., M. Wang, P. Li, X. Sun, L. Dong and P. Li (2022). A micromixer driven by two valveless piezoelectric pumps with multistage mixing characteristics. *Sensors and Actuators: A. Physical* 333, 113225.

Liu, G., X. Ma, C. Wang, X. Sun and C. Tang (2018). Piezoelectric driven self-circulation micromixer with high frequency vibration. *Journal of Micromechanics and Microengineering* 28, 085010.

Mondal, B., S. Pati and P. K. Patowari (2021). Serpentine square wave microchannel fabrication with WEDM and soft lithography. *Materials Today: Proceedings* 46, 8513-8518.

Nezhad, J. R. and S. A. Mirbozorgi (2018). An immersed boundary-lattice Boltzmann method to simulate chaotic micromixers with baffles. *Computers and Fluids* 167, 206-214.

Nouri, D., A. Z. Hesri and M. P. Fard (2017). Rapid mixing in micromixers using magnetic field. *Sensors and Actuators A* 255, 79-86.

Pang,Y., Q. Zhou, X. Wang, Y. Lei, Y. Ren, M. Li, J. Wang and Z. Liu (2020). Droplets generation under different flow rates in T-junction microchannel with a neck. *AIChE Journal* 66(10), 16290.

Prakash, R., M. Zunaid and Samsher (2021). Simulation analysis of mixing quality in T-junction micromixer with bend mixing channel. *Materials Today: Proceedings* 47, 3833-3838.

Quiyoom, A., V. V. Buwa and S. K. Ajmani (2017). Role of free surface on gas-induced liquid mixing in a shallow vessel. *AIChE Journal* 63(8), 3582-3598.

Redapangu, P. R., T. G. Kidan and K. Berhane (2021). Mixing and interpenetration in a three dimensional Buoyancy-Driven flow of two immiscible liquids: A GPU based LBM Approach. *Journal of Applied Fluid Mechanics* 14(2), 601-613.

Rouhi, O., S. R. Bazaz, H. Niazmand, F. Mirakhorli, S. M. hafi, H. A. Amiri, M. Miansari and M. E. Warkiani (2021). Numerical and experimental study of cross-sectional effects on the mixing performance of the spiral microfluidics. *Micromachines* 12, 1470.

Ruijin, W., L. Beiqi, S. Dongdong and Z. Zefei (2017). Investigation on the splitting-merging passive micromixer based on Baker's transformation. *Sensors and Actuator B* 249, 395-404.

Saleel, C. A., A. Algahtani, I. A. Badruddin, T. M. Yunus Khan, S. Kamangar and M. A. H. Abdelmohimen (2020). Pressure-Driven Electro-Osmotic Flow and Mass Transport in Constricted Mixing Micro-channels. *Journal of Applied Fluid Mechanics* 13(2), 429-441.

Santana, H. S., J. L. Silva Jr and O. P. Taranto (2019). Optimization of micromixer with triangular baffles for chemical process in millidevices. *Sensors and Actuators B: Chemical* 281, 191-203.

Seo, H. S. and Y. J. Kim (2012). A study on the mixing characteristics in a hybrid type microchannel with various obstacle configurations. *Materials Research Bulletin* 47, 2948-2951.

Shimizu, H. and Y. Uetsuji (2022). Fluid-structure and electric coupled analysis of a valveless microfluidic system using metal-capped piezoelectric actuator. *Sensors and Actuators A: Physical* 333, 113232.

Shinde, A. B., A. V. Patil and V. B. Patil (2021). Enhance the mixing performance of water and ethanol at micro level using geometrical modifications. *Materials Today: Proceedings* 46, 460-470.

Tokas, S., M. Zunaid and M. A. Ansari (2021). Non-Newtonian fluid mixing in a Three Dimensional spiral passive micromixer. *Materials Today: Proceedings* 47, 3947-3952.

Tony, A., A. Rasouli, A. Farahinia, G. Wells, H. Zhang, S. Achenbach, S. M. Yang, W. Sun and W. Zhang (2021). Toward a Soft Microfluidic System: Concept and Preliminary Developments. In *27th International Conference on Mechatronics and Machine Vision in Practice*, 755-759.

Umadevi, C., M. Dhange, B. Haritha and T. Sudha (2021). Flow of blood mixed with copper nanoparticles in an inclined overlapping stenosed artery with magnetic field. *Case Studies in Thermal Engineering* 25, 100947.

Vatankhah, P. and A. Shamloo (2018). Parametric study on mixing process in an in-plane spiral micromixer utilizing chaotic advection. *Analytica Chimica Acta* 1022, 96-105.

Wu, B., C. Li, M. Zhand and P. Luo (2021). Liquid mixing intensification by adding swirling flow in the transverse jet mixer. *AIChE Journal* 67(8), 17276.

Xia, G. D., Y. F. Li, J. Wang and Y. L. Zhai (2016). Numerical and experimental analyses of planar micromixer with gaps and baffles based on field synergy principle. *International Communications in Heat and Mass Transfer* 71, 188-196.

Xu, Z., C. Li, D. Vadillo, X. Ruan and X. Fu (2011). Numerical simulation on fluid mixing by effects of geometry in staggered oriented ridges micromixers. *Sensors and Actuator B* 153, 284-292.

Yang, S. M., S. Lv, W. Zhang and Y. Cui (2022). Microfluidic Point-of-Care (POC) Devices in Early Diagnosis: A Review of Opportunities and Challenges. *Sensors* 22, 1620.

Zhang, S., X. Chen, Z. Wu and Y. Zheng (2019). Numerical study on stagger Koch fractal baffles micromixer. *International Journal of Heat and Mass Transfer* 133, 1065-1073.

The current issue and full text archive of this journal is available on Emerald Insight at:
https://www.emerald.com/insight/1477-7266.htm

Does workplace toxicity influence turnover intentions among Indian healthcare employees? Investigating the moderating role of gratitude

Received 10 May 2022
Revised 26 September 2022
Accepted 12 January 2023

Naval Garg
University School of Management and Entrepreneurship, Delhi Technological University, Delhi, India
Manju Mahipalan
NIT, Calicut, India, and
Nidhi Sharma
Kirori Mal College, University of Delhi, Delhi, India

Abstract

Purpose – The study examined the relationship between workplace toxicity and turnover intentions among Indian healthcare employees. It also explored the role of gratitude as a moderator in the workplace toxicity–turnover intentions relationship.

Design/methodology/approach – The study is based on a cross-sectional research design. The sample comprises 315 employees from the Indian healthcare sector. Approximately, 400 employees are approached both through email and office visits. Responses were received from 336 participants, and 21 incomplete questionnaires were discarded. The relationships between four variables of workplace toxicity and turnover intentions are examined using correlation and hierarchical regression. The moderation effect of gratitude is studied using the PROCESS macro in SPSS 21.

Findings – The results revealed that workplace toxicity could explain 45.8% variations in employees' turnover intentions. It also reported significant negative regression coefficients between all four dimensions of workplace toxicity and turnover intentions. It suggested that toxic health organizations may promote turnover intentions among healthcare employees. Also, findings recommended a significant moderating effect of gratitude amid the relationships of four dimensions of workplace toxicity and turnover intentions.

Practical implications – Hospital administrators must ensure that health professionals have the necessary support to remain effective in the field by providing a conducive working environment emerging from sound human resource practices that promote respect, collegial relationships, teamwork and collaboration. The present research demonstrates gratitude as one such factor that could act as a catalyst within the workplace.

*The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

*Informed consent of respondents was taken.

* All authors contributed equally in the conceptualization and writing of the study

Declaration: We hereby declare that the present article is compiled with Ethical Standards as below.*

1. Disclosure of potential conflicts of interest.

2. No conflict of interest.

3. Research involving human participants and/or animals- Informed consent of human respondents taken.

4. *Funding*: Not applicable.

5. Data is not shared owing to the promise of confidentiality with respondents.

6. *Ethical approval*: All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Journal of Health Organization and Management
© Emerald Publishing Limited
1477-7266
DOI 10.1108/JHOM-08-2022-0233

Practitioners could achieve a healthy work environment by developing complementary relief measures that build organizational capacities and improve its culture while sponsoring programs for individual employees that instill positivity through awareness of gratitude in everyday life.

Originality/value – This study offered a comprehensive understanding of workplace toxicity by investigating its four dimensions. Also, it is one of the pioneer studies that evaluate the role of gratitude in restricting workplace toxicity-induced turnover intentions.

Keywords Healthcare workers, Workplace toxicity, Gratitude, Turnover intentions

Paper type Research paper

Introduction

The instances of toxic behavior are on the rise in modern workplaces (Rasool *et al.*, 2021). A toxic work environment is dysfunctional for employees, as they are exposed to psychosocial hazards which harm their well-being and diminish productivity. Quite often, the ruinous effects of workplace toxicity have far-reaching implications. Unpleasant work dynamics can affect employees beyond work and even more so during the ongoing COVID-19 pandemic (Abbas Khan, 2022). A toxic work environment may go against the organization's fundamental interests and undermine its principles (Tiwari and Jha, 2022). Anjum and Ming (2018) define a toxic workplace as negative interpersonal behaviors within the organization. Used as an embracive term, toxicity includes a variety of behaviors, including bullying, narcissism, ostracism, aggression, harassment and hostility in the workplace. Tastan (2017) identified two categories of workplace toxins, namely behavioral and contextualizing, an attempt to create a broad spectrum of toxic workplaces based on the characteristics, traits and behaviors. Appelbaum and Roy-Girard (2007) explored the ideas of a toxic leader, toxic management and toxic culture as several entry points for toxicity in organizations. Deery *et al.* (2011) noted that harassment and bullying are common workplace occurrences. They have a corrosive effect on morale and lead to undesirable consequences, including absenteeism, turnover, unethical practices and unproductiveness (Agrawal and Pandey, 2021).

The repercussions of toxic workplace behavior are well documented. Boudrias *et al.* (2021), upon reviewing relevant literature, found that the effect of toxic behaviors such as bullying broadly falls into three categories; health impairments, occupational outcomes and personal and family domain outcomes. The psychological impact of toxicity on employees persists for a long time owing to the escalation of sadness, anger and fear, which culminates in constant emotional fatigue and a drain on mindfulness states (Niven *et al.*, 2021; Said and Tanova, 2021). In short, in light of the pervasive effect of workplace toxicity, a significant number of companies are concentrating more on tackling the problem due to the following reasons. First, toxic behaviors can result in higher stress levels, poor physical and mental health, burnout and reduced engagement, productivity and innovation, all of which are detrimental to an organization's progress. Second, it can potentially shift people's sense of what is normal and acceptable. It could desensitize the employees, propelling the vicious cycle of toxicity. Finally, the idea of work and career is changing. Employees might not stay with the organization for long when faced with workplace toxins and jump ship to better alternatives.

Toxic behavior manifestation and response to such behaviors can differ across countries due to differences in cultural settings. For instance, different cultural dimensions may promote or discourage bullying behavior, as well as influence the preferred forms of bullying. Garg *et al.* (2022) indicate that the majority of workplace toxicity studies have been conducted in western settings and that there is a need to investigate the phenomenon in non-western contexts in order to broaden knowledge. Researchers have noticed that in collectivist and high-power distance cultural contexts, damaging behaviors like bullying are not extensively studied (Majeed and Naseer, 2021; Said and Tanova, 2021). Within this frame of reference, the present study attempts to understand the impact of a range of behavioral and contextual

factors on the turnover intentions of employees. Understanding turnover intentions is crucial as it is the best predictor of actual turnover. Employee turnover, particularly the departure of high-performing employees, can be costly and damaging to an organization's health (Lazzari *et al.*, 2022). Additionally, the study examines the moderating role of gratitude on the toxicity-turnover intention linkage. By providing an upward spiral of emotions (Jans-Beken *et al.*, 2018), gratitude can become beneficial in toning down the harmful consequences. To sum up, with the mental and physical health of health care professionals becoming a global concern (Hofmeyer and Taylor, 2021; Kriakous *et al.*, 2020; Santabárbara *et al.*, 2021), the study revolves around the objectives of

(1) understanding and examining the relationship between workplace toxicity and turnover intentions in the Indian context and

(2) the role of gratitude as a moderator in the workplace toxicity-turnover intentions relationship.

Theoretical framework and hypothesis development

Although a handful of studies have addressed the problem of workplace toxicity in various Indian contexts (D'Cruz*et al.*, 2016; Rai and Agarwal, 2017, 2018; Srivastava and Dey, 2019), the said phenomenon remains underexplored in the Indian healthcare sector. One of the largest sectors in terms of revenue and employment, healthcare poses several challenges for its employees, which have been aggravated due to the pandemic (Chatterjee *et al.*, 2021). Limited studies have pointed out the effect of specific forms of toxic behaviors on the intention to quit among healthcare professionals, such as abusive supervision (Pradhan and Jena, 2018a, b). However, other forms of toxicity, such as controlling behavior, manipulation, bullying and humiliation, remain un-investigated. The Indian healthcare system is central to the country's well-being and progress and is driven forward by a talented and committed workforce. It is imperative to know the psychological problems that can potentially plague the workplace, thus inhibiting the ability to function optimally as a healthcare worker in the Indian healthcare system. A recent study reported that burnout due to workplace stress is highly prevalent among Indian healthcare professionals for several reasons (Kesarwani *et al.*, 2020). It means grave implications for professionals' wellness, institutions' performance and patient outcomes. The profession's technical, emotional and time demands can leave people unhappy and unable to connect empathetically, and workplace toxicity can aggravate this state of affairs.

Understanding the behavioral and contextual factors that magnify workplace toxins is critical in determining areas of improvement within the organization in order to reduce actual employee turnover. When it comes to people, the primary goal of any organization is to retain talented employees in order to ensure the organization's long-term viability. Also, not many researchers have analyzed the possibility of gratitude awareness and practices protecting healthcare professionals against instances of stress, burnout and other unpleasant experiences at work.

Workplace toxicity

Tastan (2017) noted that literature about workplace toxicity describes toxic behavior as "a psychological self-centred disconnect from humanity, and a subsequent severing of empathetic ties to other people" (p. 4). Such workplaces can bring in more hostility and leave employees feeling undervalued with diminished morale and productivity levels. Toxicity can manifest in several behaviors, such as the tendency to control, manipulate, bully and humiliate and the exhibition of narcissism (Kusy and Holloway, 2009). Additionally, aggressiveness, devious politics, negativity, passivity and lack of credibility have been cited

as examples of toxicity (Anderson, 2013; Brightman, 2013; Saxena *et al.*, 2020). Other studies pinpointed specific toxic behaviors which can manifest in workplaces, such as ostracism, incivility and harassment (Anjum and Ming, 2018). In a stretch of time, researchers have thrown light on the common characteristics of a toxic workplace such as rewarding mediocre performance instead of merit-based performance, high employee turnover, managers who tend to lose their temper and are overly demanding, extreme focus only on self-advancement, negative emotional contagion and high stress among others (Colligan and Higgins, 2006; Gilbert *et al.*, 2012; Sarkar *et al.*, 2022). Frostand Robinson (1999) posited that personality traits, career setbacks, resistance to change, autocratic work environments, downsizing and other management actions spawned toxic behaviors and workspaces, and Wilde (2016) outlined the macro sources of toxicity which can contaminate the organization, classifying them into structural, cognitive, behavioral and symbolic sources. However, a closer look at the factors contributing to the formation of a toxic workplace reveals that the role of ego-centric leaders/managers fairly takes the central focus (Bhandarker and Rai, 2019; Gilbert *et al.*, 2012; Goldman, 2006; Green, 2014; Lipman-Blumen, 2010; Smith and Fredricks-Lowman, 2020; Webster *et al.*, 2016; Williams, 2018; Garg, 2022).

For the present study, the area of inquiry apropos of workplace toxicity factors is based on the holistic view of the toxic workplace environment developed by Tastan (2017). The strength of this framework comes from the quantitative and qualitative studies, which resulted in two distinct categories of workplace toxins encompassing most of the characteristics, traits and behaviors within the organizational settings. Consequently, the human side of workplace toxicity or 'behavioral toxics' includes managers' toxic behaviors and coworkers' toxic behaviors. The second category of 'contextual toxics' is factored in the contextual elements such as toxic social-structural factors and toxic climate. This schema of workplace toxicity is overarching to a great extent because it identifies toxicity as a phenomenon eventualized by the continuous interactions between negative actions, situations, perceptions and undesirable organizational attributes.

Effects of workplace toxicity

Appelbaum and Roy-Girard (2007) posited that toxic organizations are likely to underperform because of the hostile work environment characterized by feelings of despair, anger, low morale, poor communication and lack of trust. Such workplaces could also witness high levels of deviant behavior where the employee willingly and voluntarily violates the organizational norms, thus posing a threat to overall organizational well-being and goals. Toxic interpersonal, occupational and organizational conditions were found to have a significant adverse effect on various job outcomes, including satisfaction, commitment and experience of meaningfulness at work (Chamberlain and Hodson, 2010). A toxic environment impedes project success by rousing workplace stress (Wang *et al.*, 2020), and toxic leadership can sabotage performance at the individual, team and organizational levels, decreasing motivation, commitment and innovation (Williams, 2018). Apart from these, researchers have been able to study the ramifications of toxic workplace behaviors discretely. For example, meta-analytic reviews have outlined the psychosomatic, psychological, behavioral and attitudinal problems associated with workplace bullying across various cultures and work contexts (Moayed *et al.*, 2010; Nielsen and Einarsen, 2012; Verma *et al.*, 2022). Likewise, extensive research spanning two decades explains the damaging concomitants of workplace incivility, including reduced well-being, decreased job satisfaction, job productivity, work engagement, performance and elevated levels of employee cynicism (Aljawarneh and Atan, 2018; Vasconcelos, 2020). Similarly, workplace ostracism as a stressor adversely impacts employees' task performance, interpersonal behaviors and well-being and instigates deviant behavior (Chung, 2018; Howard *et al.*, 2020). More recent research has addressed the mental

distress caused by bullying and harassment due to extended periods of emotional exhaustion or burnout (Anasori *et al.*, 2019; Nielsen *et al.*, 2020; Rossiter and Sochos, 2018; Mehta *et al.*, 2022). Conway *et al.* (2021) note that bullying can be a powerful stress factor capable of inducing enduring trauma by shattering people's beliefs of themselves and their environment. Furthermore, workplace mistreatment can weaken employee voice as a result of a breach of psychological contract (Rai and Agarwal, 2018). Researchers have emphasized the far-reaching repercussions of toxic behaviors that extend beyond the apparent perpetrator-target relationship, as individuals who witness bullying or harassment experience a violation of psychological safety at work (Salin and Notelaers, 2018).

Workplace toxicity and turnover intentions

The present research attempts to understand the impact of multiple toxic traits and behaviors on a single outcome: employee intention to quit; turnover intentions. Turnover intention indicates the employee's behavioral intention to leave the organization and can be viewed as the final step in the decision-making process (Bothma and Roodt, 2013). Carmeli and Weisberg (2006) detailed the three elements associated with the withdrawal cognition process: thoughts of quitting the job, intention to search for a different job and the intention to quit. Turnover intention is a much-studied phenomenon as it carries significant implications for organizations. Turnover intentions are detrimental to the organization's health, and they do not exist in a vacuum (Rahman and Nas, 2013; Sharma *et al.*, 2022). Though there is a difference between turnover intention and actual turnover, researchers agree that turnover intentions can be used as a valid proxy for substantial turnover as it is a reliable determinant and has an immediate causal effect on the turnover decision of the employee (Chen *et al.*, 2010; Korunka *et al.*, 2008; Liu and Onwuegbuzie, 2012; Sjöberg and Sverke, 2000; Steensma *et al.*, 2003). Researchers have attempted to understand the plausible causes of turnover intentions using several models, including the widely known Job resources-demand (JD-R) model (Schaufeli and Bakker, 2004). Studies demonstrate that employees' intentions to quit are intensified by workplace stress, emotional exhaustion, burnout, higher level of job demands, lower levels of satisfaction and commitment, organizational politics, and poor management (Applebaum *et al.*, 2010; Beecroft *et al.*, 2008; Chen *et al.*, 2010; Kim, 2015; Liu and Onwuegbuzie, 2012; Poon, 2004; Rahman and Nas, 2013; Slattery and Rajan Selvarajan, 2005; Garg *et al.*, 2021).

Although no studies have investigated the impact of the entire spectrum of workplace toxins on turnover intentions, clear evidence exists for a coherent link between toxic behaviors and intentions to quit. For instance, employees considered quitting the organization due to concerted workplace aggression by way of verbal harassment as they believed it to be the reflection of the underlying organization's social norms and viewed exiting the organization as a more rational and viable option (Deery *et al.*, 2011). When subjected to covert and overt uncivil behaviors, employees decide to leave the organization on account of significant distress (Sguera *et al.*, 2011). Commenting from a gender perspective, Gardner (2012) reported the structural, social and cultural toxins such as lack of resources, quality leadership, work-life balance policies and overall negative institutional and departmental environment, which compelled the women employees to leave their organization. Abusive leaders exhibiting hostile behaviors and rudeness aggravate employees' desire to withdraw from the organization (Labrague *et al.*, 2020; Mathieu and Babiak, 2016; Tillman *et al.*, 2018; Garg and Sarkar, 2020). Sim *et al.* (2021) concur with this finding stating that toxic leadership can amplify organizational intimidation and manipulate employees' sense of job security, thus contributing to turnover intentions. Ahmad and Kaleem (2019) examined the direct and indirect effects on turnover intentions. They found that constant exposure to harmful bullying could create jarring experiences for the victim,

thus severely undermining their well-being. It prompts the individual to re-evaluate the decision to stick with the organization. Toxic social stressors like workplace ostracism, where the individual is dismissed or ignored in a group, can induce feelings of isolation and heighten the tendencies to withdraw from the organization (Glazer *et al.*, 2021; Garg and Saxena, 2020). Evidence suggests that workplace bullying tends to exacerbate employees' intention to quit by worsening occupational burnout (Shen Hsiao *et al.*, 2021; Srivastava and Agarwal, 2020; Garg, 2020). Constant exposure to toxic workplace behaviors reduces employees' perceptions of fairness and respect and raises their perceptions of injustice, spurring them to leave the organization (Reknes *et al.*, 2021). Based on the extant literature, following hypotheses are formulated for the present study:

H1. Toxic behaviors of coworkers are negatively associated with employees' turnover intentions.

H2. Toxic behaviors of managers are negatively associated with employees' turnover intentions.

H3. Toxic social-structural factors are negatively associated with employees' turnover intentions.

H4. Toxic climate is negatively associated with employees' turnover intentions.

The moderating role of gratitude

Gratitude, or a feeling of thankfulness, has been understood as an essential and effective predictor of individual well-being (Chen *et al.*, 2021; Watkins, 2004). Gratitude plays a critical role in building human relationships and shaping various work attitudes (Lanham *et al.*, 2012). The usefulness of gratitude at the workplace to combat workplace toxicity arises from the fact that individuals with higher levels of gratitude bring positive strategies to cope with challenges at work (Fehr *et al.*, 2017) instead of a toxic coping strategy. Building on Fredrickson's (2004) theory of positive emotions, Alkozei *et al.* (2018) proposed that gratitude broadens the cognitive and emotional resources of a person to deal with stressors by allowing the individual to reframe negative or challenging situations positively and allocating more attention to the positive stimuli in the environment. It illustrates a way in which gratitude is associated with emotional regulation, wherein it supplements coping strategies and thus contributes to the overall well-being of individuals. Recent research findings corroborate this postulation (please see Guan and Jepsen, 2020; Krejtz *et al.*, 2014; Mărean *et al.*, 2019; Nezlek *et al.*, 2019; Portocarrero *et al.*, 2020; Garg, 2020; Garg and Sarkar, 2020). Another way gratitude could buffer the harmful effects of workplace toxicity is by strengthening social connectedness and building high-quality relationships at work. High-quality relationships enhance psychological safety and work motivation while lessening emotional exhaustion (Carmeli *et al.*, 2009; Carmeli and Gittell, 2009; Fernet *et al.*, 2010; Garg and Mehak, 2021; Garg *et al.*, 2021). Workplace friendships impelled by gratitude can prompt individuals to explore, collaborate and negotiate with fellow beings (Badri *et al.*, 2022). Such interactions might downscale the harmful effect of toxic behaviors by forming significant social and psychological resources. Davis *et al.* (2021) note that when people feel appreciated, the motivation to work enhances and reduces the turnover intent. Although the role of gratitude as a moderator concerning turnover intentions has not been examined categorically, it is plausible that higher levels of gratitude can obviate individuals' tendency to withdraw from the organization despite the prevalence of toxic behaviors. Accordingly, the following four hypotheses are proposed.

H5. Gratitude moderates the relationship between toxic behaviors of coworkers and employees' turnover intentions.

H6. Gratitude moderates the relationship between toxic behaviors of managers and employees' turnover intentions.

H7. Gratitude moderates the relationship between toxic social-structural factors and employees' turnover intentions.

H8. Gratitude moderates the relationship between a toxic climate and employees' turnover intentions.

Methodological framework

Sample and data collection

The study is based on a cross-sectional research design with participants approached only once for data collection. Also, convenient sampling was used to collect the data. The sample comprises 315 employees from the Indian healthcare sector. Approximately 400 employees are approached both through email and office visits. Responses were received from 336 participants, and 21 incomplete questionnaires were discarded. A minimum sample size of 275 was required as per general rule prescribed by Hair *et al.* (2010). According to Hair *et al.* (2010), minimum sample size should be five times the number of statements in the questionnaire. The total number of variables in the study was 55 (49 of workplace toxicity, 6 of gratitude, 6 of turnover intentions and 4 of control variables). A response rate of 70% was assumed based on previous studies. Thus, 400 employees were approached to get data from 280 participants. The sample includes 160 male (51%) and 155 female (49%) respondents. 201 employees are married (64%), and 114 are unmarried (36%). Also, 72(23%), 130(41%), 65 (21%) and 48(15%) employees belong to the age category of "below 25 years" "26–35 years" "36–45 years" and "above 45 years" respectively. The average experience of the employees is 12.9 years.

Measures

Workplace toxicity. Workplace toxicity is measured using Tastan's (2017) toxic workplace environment questionnaire. The scale explores two environmental toxins named behavioral toxics and contextual toxics. These workplace toxins are categorized into four variables-toxic behaviors of coworkers, toxic behaviors of managers, toxic social-structural factors and toxic climate. The questionnaire has 49 items that measure workplace toxicity on a six-point rating scale ranging from 1 (*never*) and 6 (*always*). A few items of the scale are "I feel as if each day adds new stress on employees," "I witness a manager constantly act in a harassing and terrorizing way towards a subordinate," and "I occasionally witness a coworker ignore and not return greetings upon meeting others."

Turnover intentions. Roodt's (2004) turnover intention scale is used to measure the employees' turnover intentions (TIS). It comprises six items that rate turnover intentions on a five-point rating scale ranging from 1 (*never*) to 5 (*always*). A few items of the scale are "How satisfying is your job in fulfilling your personal needs?" and "How often are you frustrated when not given the opportunity at work to achieve your personal work-related goals?"

Gratitude. McCullough *et al.*'s (2002) gratitude questionnaire-6 (GQ-6) is used to capture one's proneness to experience gratitude in daily life. GQ-6 is a six-item scale that measures gratitude on a seven-point scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). A few items of the scale are "I have so much in life to be thankful for" and "If I had to list everything that I felt grateful for, it would be a very long list."

Control variables. Gender, age, experience and marital status are taken as control variables in the study. These variables are controlled because they influence the experience of toxicity in the workplace (Howards *et al.*, 2020; Kin, 2015; Garg *et al.*, 2020).

Analysis of data

The collected data is subjected to reliability, validity, multicollinearity and common method variance explorations. The relationships between four variables of workplace toxicity and turnover intentions are examined using correlation and hierarchical regression. A two-stage regression model is preferred. Model 1 explores the control variables, and model 2 discusses the significance of both control and independent variables. Later, the moderation effect of gratitude is studied using the PROCESS macro in SPSS 21.

Results

Reliability and validity

Cronbach's alpha (CA) and composite reliability (CR) estimates are calculated to assess internal consistency reliability. Acceptable values of both reliability estimates are 0.70 and above (Hair *et al.*, 1998). Table 1 reports all values of CA and CR, which confirms the reliability of the scales. Average Variance Explained (AVE) values are estimated to evaluate convergent validity. Again acceptable values are 0.70 and above (Fornell and Larcker, 1981). All AVE values, as indicated in Table 1, are greater than 0.70, which confirms convergent validity. Since the workplace toxicity questionnaire measures several toxins of work, discriminate validity is also assured. Table 2 states the results of the Fornell–Larcker test. It confirms that AVE values of each dimension of workplace toxicity exceed the squared correlation involving the dimensions. The results conclude discriminate validity of the workplace toxicity scale.

Multicollinearity and common method variance

The study is based on more than one independent variable. Thus, it is imperative to evaluate multicollinearity among coworkers' toxic behaviors, managers' toxic behaviors, toxic social-structural factors and toxic climate. The issue of multicollinearity is examined using the variance inflation factor (VIF) value. Table 1 reports that all VIF values are less than five, confirming the absence of multicollinearity among independent variables.

Variables	Mean	S.D	CA	CR	AVE	VIF
Toxic behaviors of coworkers	3.46	0.968	0.783	0.911	0.591	1.018
Toxic behaviors of managers	3.83	0.969	0.790	0.924	0.603	2.530
Toxic social-structural factors	3.38	0.981	0.831	0.931	0.618	2.381
Toxic climate	4.44	0.652	0.765	0.899	0.578	1.894
Gratitude	4.68	0.813	0.876	0.963	0.622	–
Turnover intention	3.33	0.882	0.724	0.860	0.588	–

Note(s): CA- Cronbach'alpha, CR- Composite reliability, AVE- Average Variance Explained, VIF- Variance Inflation Index
Source(s): Primary data

Table 1. Descriptive statistics

Variables	1	2	3	4
Toxic behaviors of coworkers	1 *(0.591)*			
Toxic behaviors of managers	0.218	1 *(0.603)*		
Toxic social-structural factors	0.110	0.243	1 *(0.618)*	
Toxic climate	0.461	0.592	0.481	1 *(0.578)*

Note(s): AVE values are in italic (diagonally), rest values are square of the correlation
Source(s): Primary data

Table 2. Discriminant validity (Results of the Fornell-Larcker test)

Further, since data about independent and dependent variables are collected simultaneously, there might be some issues with common method variance. Both procedural and statistical remedies are used to address the variance. First, responses are sought anonymously, i.e. the names of the respondents are not collected. Also, participants are promised strict confidentiality of data. They are ensured that data would be used for academic purposes only and would not be shared with any other party. Podsakoff *et al.* (2012) suggested that anonymity and confidentiality motivate respondents to respond more accurately and freely. Second, two cover stories are used to psychologically separate independent and dependent variables. Through these stories, participants are ensured that workplace toxicity and turnover intentions scales are part of two unrelated studies. Podsakoff *et al.* (2012) suggested that psychological separation reduces the probability of the latent linkage between the independent and dependent variables. However, respondents are made aware of the truth and reason for cover stories after collecting data. Third, independent and dependent variables are measured using a different rating scale. It helps in addressing the common method variance. And lastly, Harman's single factor method is applied to assess common method variance. In Harman's approach, all items of the study are subjected to factor analysis. No signal factor is reported to explain the majority of the variation, which according to Chang *et al.* (2010), negates the possibility of significant common method variance. The largest factor could explain 21.7% of the variance.

Analysis of data

The relationships between four variables of workplace toxicity and turnover intentions are examined using correlation and hierarchical regression. In hierarchical regression, the first model explored the effect of four variables (gender, experience, age and marital status). And in the second model, the effects of control variables and four variables of workplace toxicity were investigated. The moderation effect of gratitude is studied using the PROCESS macro in SPSS 21.

Results

Correlation and hierarchical regressions

Table 3 reports the correlation matrix of all the variables of the study. It highlights significant positive relationships between four dimensions of workplace toxicity and the employees' turnover intentions. It suggests that both toxicity at work and turnover intention move in the same direction. Also, statistically significant negative correlations are reported between four forms of workplace toxicity and gratitude. Similarly, a significant negative association is observed between gratitude and turnover intentions. Table 4 represents the results of two-stage hierarchical regression, and Table 5 describes the model summary. In the first stage,

Variables	1	2	3	4	5	6
Toxic behaviors of coworkers	1					
Toxic behaviors of managers	0.467*	1				
Toxic social-structural factors	0.332*	0.493*	1			
Toxic climate	0.679*	0.770*	0.694*	1		
Gratitude	−0.350*	−0.466*	−0.379*	−0.564*	1	
Turnover intention	0.344*	0.554*	0.591*	0.653*	−0.422*	1

Note(s): * Sig. at 0.05
Source(s): Primary data

Table 3.
Correlation matrix

Table 4.
Results of hierarchical multiple regression (DV- Turnover Intentions)

Model	Variable	Unstandardized coefficient B	Unstandardized coefficient Std. error	Standardized coefficient Beta	T	Sig.
1	Constant	3.801	0.249		15.271	0.00*
	Gender	0.075	0.099	0.069	2.209	0.373
	Age	0.022	0.111	0.041	0.366	0.715
	Experience	0.176	0.086	0.158	1.832	0.068
	Marital status	0.186	0.100	0.186	1.868	0.063
2	Constant	0.703	0.279		2.523	0.012*
	Gender	0.096	0.073	0.069	2.332	0.520
	Age	0.002	0.082	0.004	0.047	0.963
	Experience	0.072	0.064	0.065	1.017	0.310
	Marital status	0.065	0.074	0.065	0.886	0.376
	Toxic behaviors of coworkers	0.190	0.053	0.198	1.675	0.049*
	Toxic behaviors of managers	0.221	0.059	0.233	2.034	0.033*
	Toxic social-structural factors	0.424	0.053	0.449	4.190	0.000*
	Toxic climate	0.491	0.132	0.512	4.490	0.000*

Note(s): * Sig. at 0.05
Source(s): Primary data

Table 5.
Model summary

Model	R	R^2	Adjusted R^2	Change statistics F changes	Change statistics Sig F changes	ANOVA F	ANOVA Sig.
1	0.179	0.032	0.019	2.560	0.39*	2.560	0.39*
2	0.700	0.490	0.477	68.672	0.00*	36.733	0.00*

Note(s): * Sig. at 0.05
Source(s): Primary data

only control variables are regressed. It is revealed that no control variable is a significant predictor of turnover intentions; however overall model is significant ($F = 2.560$, $p < 0.05$). Now, in the second stage, all control variables and independent variables (four dimensions of workplace toxicity) are regressed over employees' turnover intentions. All regression coefficients of four independent variables are significant. The most competent predictor of turnover intentions of the employees is the toxic climate ($\beta = 0.512$, $p < 0.05$), followed by the toxic social-structural factors ($\beta = 0.449$, $p < 0.05$). The least significant predictor is the toxic behavior of coworkers ($\beta = 0.198$, $p < 0.05$). These results indicate acceptance of the study's first four hypotheses (H1–H4). The comparative evaluation of model 1 and model 2 clarifies that the coefficient of determination (R^2) has improved from 0.032 to 0.49. It means control variables explain only 3.2% of variations in the dependent variables, and the remaining 45.8% of variations are explained by workplace toxicity.

Moderation analysis using the PROCESS macro

Tables 6–9 elaborate on the results of the moderation analysis. Table 6 reports the moderating effect of gratitude amid the relationship between toxic behavior of coworkers and employees' turnover intentions. It is observed that the regression weight of toxic behavior of coworkers reduces from 0.729 to 0.100 on interacting with gratitude. In simple words, employees' gratuitous behavior minimizes the effect of toxic behavior of coworkers on turnover intentions. Also, F-value concludes the significant moderating effect of gratitude in

Table 6. Table 7 investigates the moderating effect of gratitude amid the relationship between the toxic behavior of managers and employees' turnover intentions. Again, it is observed that the regression weight reduces from 0.505 to 0.032 when the toxic behavior of the manager interacts with gratitude. Furthermore, F-value indicates that the moderation effect of gratitude is significant in Table 7. Similar results could also be observed in Tables 8 and 9. These results conclude the significant moderating effect of gratitude and thereby suggest acceptance of hypotheses 6 to 8.

Model	Coefficient	S.E	*t* value	*p*	LLCI	ULCI
Constant	4.7327	0.586	8.069	0.000	3.578	5.886
TBC	0.729	0.138	5.254	0.000*	0.200	0.346
G	−0.668	0.194	−3.946	0.041*	−1.151	−0.385
Interaction (G*TBC)	0.100	0.047	2.126	0.034*	0.007	0.194

Model	R^2-Change	F-value	*p*
Interaction (G*TBC)	0.112	4.522	0.034*

Note(s): G-Gratitude, TBC- Toxic behavior of workers, * Sig at 0.05
Source(s): Primary Data

Table 6. Results of moderation analysis (IV- Toxic behaviors of coworkers)

Model	Coefficient	S.E	*t* value	*p*	LLCI	ULCI
Constant	2.140	0.598	3.577	0.004*	0.963	3.318
TBM	0.505	0.138	3.666	0.003*	0.234	0.777
G	−0.098	0.192	−2.178	0.017*	−0.280	−0.477
Interaction (G*TBM)	0.032	0.045	1.949	0.029*	0.058	0.122

Model	R^2-Change	F-value	*p*
Interaction (G*TBM)	0.091	7.394	0.001*

Note(s): G-Gratitude, TBM- Toxic behavior of managers, * Sig at 0.05
Source(s): Primary Data

Table 7. Results of moderation analysis (IV- Toxic behaviors of managers)

Model	Coefficient	S.E	*t* value	*p*	LLCI	ULCI
Constant	2.702	0.518	5.214	0.000*	1.682	3.772
TSSF	0.394	0.122	3.222	0.001*	0.153	0.636
G	−0.334	0.172	−2.347	0.022*	−0.674	−0.006
Interaction (G*TSSF)	0.021	0.423	2.849	0.017*	0.061	0.104

Model	R^2-Change	F-value	*p*
Interaction (G*TSSF)	0.154	8.748	0.041*

Note(s): G-Gratitude, TSSF- Toxic social-structural factors, * Sig at 0.05
Source(s): Primary data

Table 8. Results of moderation analysis (IV- Toxic social-structural factors)

Discussion

The present exploration covers a significant ground in the subject research by providing for the first time comprehensive empirical evidence of the relationship between workplace toxicity and turnover intentions with gratitude as a moderator of such association in the Indian healthcare sector. Literature on workplace toxicity enumerates its negative impact on critical workplace factors, including job satisfaction, commitment, motivation, organizational citizenship behavior, job engagement, well-being, trust and goal integration, among others (Bakker *et al.*, 2007; Poulose and Sudarsan, 2017; Pradhan and Jena, 2018; Surya *et al.*, 2017; Taştan, 2017; Zhou *et al.*, 2020), that ultimately have a bearing on an employee's intention to stay or quit (Martinko *et al.*, 2013; McNall *et al.*, 2009; Mosadeghrad *et al.*, 2008; Ng *et al.*, 2021; Taştan, 2017). Such workplace attitudes form an integral part of successful performance and organizational effectiveness, especially in the healthcare sector (Afolabi *et al.*, 2018; Goldstein, 2003; Lambrou *et al.*, 2010; Mosadeghrad and Ferdosi, 2013; Ng *et al.*, 2021; Shantz *et al.*, 2016; Top *et al.*, 2015; Garg, 2020). Moreover, studies have indicated an intention to quit and turnover as severe issues in the healthcare sector (Bharath, 2021) and within the Indian cultural context (Dasgupta, 2015; Srinivasan and Chandwani, 2014; Garg, 2017). Together they are a clarion call to action for researchers, making the present examination of workplace toxicity, its impact on turnover intentions and ways for managing it highly relevant in this sector.

Implications for theory

The healthcare sector is known for its inherently demanding and challenging work environment (Greenglass *et al.*, 2001; McVicar, 2003; Nielsen *et al.*, 2009; Perez *et al.*, 2015; Surya *et al.*, 2017). The scenario is no different in the Indian context (Chhugani and James, 2017; Dhiman and Arora, 2018; Nayak *et al.*, 2018; Poulose and Sudarsan, 2017), where healthcare professionals experience high levels of emotional labor, exhaustion and burnout (Afolabi *et al.*, 2018; Aiken *et al.*, 2002; Maslach *et al.*, 1996; Pandey and Singh, 2015; Srinivasan and Chandwani, 2014). A toxic work culture would only add to the already stressful environment and amplify the emotional disconnect of employees with their workplace. Such withdrawal cognitions and resultant intentions to quit are worrisome phenomena since they often precede quitting. Through strong direct associations between workplace toxicity dimensions and employees' turnover intentions, the results of this study provide concrete evidence of this effect.

While the study results indicate a need for timely intervention aimed at eradicating workplace toxicity to reduce the associated detrimental outcomes, it may be unlikely to eliminate toxicity from the workplace fully. Thus, investigating the relationship between

Model	Coefficient	S.E	*t* value	*p*	LLCI	ULCI
Constant	0.742	0.675	1.098	0.272	−0.587	2.072
TC	0.819	0.158	5.177	0.000*	0.508	1.130
G	−0.491	0.202	−3.450	0.037*	−0.690	−0.307
Interaction (G*TC)	0.001	0.048	4.030	0.041*	−0.841	−0.097

Model	R^2-Change	F-value	*p*
Interaction (G*TC)	0.118	10.132	0.022*

Note(s): G-Gratitude, TC- Toxic Climate, * Sig at 0.05
Source(s): Primary Data

Table 9.
Results of moderation analysis (IV- Toxic climate)

workplace toxicity and employees' intention to quit may not be sufficient in itself without the knowledge of potential moderators, especially those that could neutralize the harmful effects of workplace toxicity. The teachings of Rāmakrishna Paramahamsa, the rejuvenator of Hinduism in the middle of the nineteenth century, in the Indian cultural context echo a similar philosophical foundation (Chidbhavananda, 1983, p. 100).

> Provocation by the external forces cannot be prevented. But refusing to be provoked is entirely in one's own hands. The mind should be so tuned that it maintains its own placidity in all eventualities. The path is full of thorns. One method is to cleanse it and another to put on shoes.

While interventions for effective management and mitigation of workplace toxicity would be akin to "cleansing the path," promoting mindfulness or contemplative techniques integrated with positive psychology variables such as gratitude would be analogous to "putting on shoes." The positive psychology literature and JD-R theory also advocate in favor of gratitude interventions to replenish the emotional resources expended while wading off toxicity at the workplace. The present study has generated empirical evidence supporting such assertions and validated gratitude as a powerful tool that moderates the negative impact of workplace toxicity on turnover interventions in the Indian healthcare sector. Since Indian society has roots in a collectivist culture (Shabir *et al.*, 2022), social connectedness at the workplace could also alleviate the undesirable outcomes of workplace toxicity. Gratitude that encourages a balanced understanding and positive thoughts in the face of unpleasant circumstances (Jans-Beken and Wong, 2019) could facilitate this social glue necessary for encouraging high-quality work relations.

Implications for practice

This study reveals that a massive 45.8% of the variation in turnover intentions is explained by workplace toxicity. It indicates that overall management of workplace toxicity should help create an environment where employees feel less inclined to quit or look for alternative employment opportunities. Furthermore, we find that contextual dimensions like toxic climate and socio-structural factors have a more significant impact on turnover intentions compared to the behavioral dimensions like toxic behavior of managers and coworkers. This signifies that healthcare workers are far more affected by the adversities rooted within the organization's culture than untoward instances arising from individual-level interactions. These results are in tune with previous findings, which show that skilled labor in the healthcare sector could be challenging to retain despite higher pay being offered to them by their employer if they are poorly treated (Mahadi *et al.*, 2020). Since healthcare organizations are highly labor-intensive (Afolabi *et al.*, 2018; Latino, 2008), the quality of their human resources is indispensable for effective service delivery. Given this, the Indian healthcare sector is in a critical state as it falls below international benchmarks and standards existing in comparable developing countries (Talib and Rahman, 2013). Here, the health and well-being of healthcare employees are often ignored. Healthcare workers deal with excessive physical and psychological pressure due to heavy workloads, workplace hazards and occupational stress (Nayak *et al.*, 2018).

As one of the largest service industries in employment, with a significant bearing on economic growth (Srinivasan and Chandwani, 2014), it is imperative for the Indian healthcare industry to ensure lower turnovers. It needs to invest in strategic HR practices focused on institutionalized change initiatives that focus on transforming the organizational values and ethos. Additionally, with increasing complexities and multidisciplinary treatment approaches, an effective healthcare system is dependent upon harmonious relationships and teamwork within and across the departments (Chhugani and James, 2017; Srinivasan and Chandwani, 2014). Therefore, hospital administrators must ensure that health professionals have the necessary support in order to remain effective in the field by providing a conducive

working environment emerging from sound human resource practices that promote respect, collegial relationships, teamwork and collaboration. The present research demonstrates gratitude as one such factor that could act as a catalyst within the workplace. Practitioners could achieve a healthy work environment by developing complementary relief measures that build organizational capacities and improve its culture while simultaneously sponsoring programs for individual employees that instill positivity through awareness of gratitude in everyday life. Such investments should reap benefits through high-quality and compassionate healthcare delivery.

Limitations and scope for future research

The study has certain limitations that stem from factors inherent in its research design and environment. The results of this study are based on a cross-sectional research design. Its framework could be used for future longitudinal studies that could help assess the validity of the outcomes from a long-term perspective. Additionally, examining the relationship in other sectors and geographies should broaden the appeal of the findings. Exploring additional factors like optimism, hope and resilience as moderators of the relationship between workplace toxicity and turnover intentions would help create a dossier of positive psychology variables that could aid practitioners in designing interventions most suited to their organizational contexts for managing workplace toxicity. One could also undertake comparative analyses of organizations with and without a history of promoting mindfulness practice modeled around positive psychology, which would be similar to performing scientific experiments under controlled environments and significantly enhance the stature of the research subject. Furthermore, since the current study relies on convenience sampling for data collection, future researchers are advised to use random sampling for better generalizability of the results. Finally, the results of this study may be supplemented with qualitative research techniques that capture dynamic processes and expand our understanding of unexpressed respondent perspectives.

Conclusion

This study reported that workplace toxicity is a significant predictor of employee turnover. Healthcare organizations are required to look at toxic variables to ensure employee retention. The present study also reports that employees' gratefulness moderates the linkage between toxicity and turnover intentions.

References

Abbas Khan, N. (2022), "Determinants of proactive work behavior of employees during the COVID-19 crisis: a perspective on toxic leadership in the virtual work setting", *European Journal of Psychology Open*, Germany: Hogrefe Publishing, Vol. 80 Nos 1-2, p. 77.

Afolabi, A., Fernando, S. and Bottiglieri, T. (2018), "The effect of organizational factors in motivating healthcare employees: a systematic review", *British Journal of Healthcare Management*, Vol. 24 No. 12, pp. 603-610.

Agrawal, A. and Pandey, A. (2021), "Types, antecedents and outcomes of organizational dysfunctional behaviors-A review", *Indian Journal of Industrial Relations*, Vol. 57 No. 2, pp. 239-252.

Ahmad, S. and Kaleem, A. (2019), "Zooming in on the workplace bullying and turnover intentions pathway: the role of well-being and a cultural boundary condition", *Personnel Review*, Emerald Publishing, Vol. 49 No. 2, pp. 425-444.

Aiken, L.H., Clarke, S.P., Sloane, D.M., Sochalski, J. and Silber, J.H. (2002), "Hospital nurse staffing and patient mortality, nurse burnout, and job dissatisfaction", *Journal of the American Medical Association*, Vol. 288 No. 16, pp. 1987-1993.

Aljawarneh, N.M.S. and Atan, T. (2018), "Linking tolerance to workplace incivility, service innovative, knowledge hiding, and job search behavior: the mediating role of employee cynicism", *Negotiation and Conflict Management Research*, John Wiley & Sons, Vol. 11 No. 4, pp. 298-320.

Alkozei, A., Smith, R. and Killgore, W.D.S. (2018), "Gratitude and subjective well-being: a proposal of two causal frameworks", *Journal of Happiness Studies*, Vol. 19, pp. 19-42.

Anasori, E., Bayighomog, S.W. and Tanova, C. (2019), "Workplace bullying, psychological distress, resilience, mindfulness, and emotional exhaustion", *The Service Industries Journal*, Vol. 40 Nos 1-2, pp. 65-89, Routledge.

Anderson, A. (2013), "Coping in a toxic work environment", *Forbes*, available at: http://www.forbes.com/sites/amyanderson/2013/06/17/coping-in-a-toxic-work-environment/

Anjum, A. and Ming, X. (2018), "Combating toxic workplace environment: an empirical study in the context of Pakistan", *Journal of Modelling in Management*, Emerald Group Publishing, Vol. 13 No. 3, pp. 675-697.

Appelbaum, S.H. and Roy-Girard, D. (2007), *Toxins in the Workplace: Affect on Organizations and Employees*, Corporate Governance, Vol. 7 No. 1, pp. 17-28.

Applebaum, D.D., Fowler, D.S., Fiedler, D.N., Osinubi, D.O. and Robson, D.M. (2010), "The impact of environmental factors on nursing stress, job satisfaction, and turnover intention", *The Journal of Nursing Administration*, NIH Public Access, Vol. 40 No. 0, p. 323.

Badri, S.K.Z., Yap, W.M. and Ramos, H.M. (2022), "Workplace affective well-being: gratitude and friendship in helping millennials to thrive at work", *International Journal of Organizational Analysis*, Vol. 30 No. 2, pp. 479-498, doi: 10.1108/IJOA-04-2020-2148.

Bakker, A.B., Hakanen, J.J., Demerouti, E. and Xanthopoulou, D. (2007), "Job resources boost work engagement, particularly when job demands are high", *Journal of Educational Psychology*, Vol. 99 No. 2, pp. 274-284.

Beecroft, P.C., Dorey, F. and Wenten, M. (2008), "Turnover intention in new graduate nurses: a multivariate analysis", *Journal of Advanced Nursing*, John Wiley & Sons, Vol. 62 No. 1, pp. 41-52.

Bhandarker, A. and Rai, S. (2019), "Toxic leadership: emotional distress and coping strategy", *International Journal of Organization Theory and Behavior*, Emerald Group Publishing, Vol. 22 No. 1, pp. 65-78.

Bharath, M. (2021), "Something beyond paycheque to boosting employee retention: evidence from a South Indian hospital", *Vilakshan - XIMB Journal of Management*, Vol. ahead-of-print No. ahead-of-print, doi: 10.1108/XJM-03-2021-0072.

Bothma, C.F.C. and Roodt, G. (2013), "The validation of the turnover intention scale", *SA Journal of Human Resource Management*, AOSIS, Vol. 11 No. 1, doi: 10.4102/SAJHRM.V11I1.507.

Boudrias, V., Trépanier, S.G. and Salin, D. (2021), "A systematic review of research on the longitudinal consequences of workplace bullying and the mechanisms involved", *Aggression and Violent Behavior*, Vol. 56, p. 101508, doi: 10.1016/j.avb.2020.101508.

Brightman, B. (2013), "How to overcome the 6 most toxic employee behaviors", Fast Company, available at: http://www.fastcompany.com/ 3023318/leadership-now/how-to-overcome-the-6-most-toxicemployee-behaviors

Carmeli, A. and Gittell, J.H. (2009), "High-quality relationships, psychological safety, and learning from failures in work organizations", *Journal of Organizational Behavior*, John Wiley & Sons, Vol. 30 No. 6, pp. 709-729.

Carmeli, A. and Weisberg, J. (2006), "Exploring turnover intentions among three professional groups of employees", *Human Resource Development International*, Vol. 9 No. 2, pp. 191-206.

Carmeli, A., Brueller, D. and Dutton, J.E. (2009), "Learning behaviors in the workplace: the role of high-quality interpersonal relationships and psychological safety", *Systems Research and Behavioral Science*, John Wiley & Sons, Vol. 26 No. 1, pp. 81-98.

Chamberlain, L.J. and Hodson, R. (2010), "Toxic work environments: what helps and what hurts", *Sociological Perspectives*, Vol. 53 No. 4, pp. 455-477.

Chang, S.J., van Witteloostuijn, A. and Eden, L. (2010), "From the editors: common method variance in international business research", *Journal of International Business Studies*, Vol. 41 No. 2, pp. 178-184.

Chatterjee, S.S., Chakrabarty, M., Banerjee, D., Grover, S., Chatterjee, S.S. and Dan, U. (2021), "Stress, sleep and psychological impact in healthcare workers during the early phase of COVID-19 in India: a factor analysis", *Frontiers in Psychology*, Frontiers, Vol. 0, p. 473.

Chen, M.-F., Lin, C.-P. and Lien, G.-Y. (2010), "Modelling job stress as a mediating role in predicting turnover intention", *The Service Industries Journal*, Routledge, Vol. 31 No. 8, pp. 1327-1345.

Chen, H., Yang, X., Xia, W., Li, Y., Deng, Y.L. and Fan, C. (2021), "The relationship between gratitude and job satisfaction: the mediating roles of social support and job crafting", *Current Psychology*, doi: 10.1007/s12144-021-01658-y.

Chhugani, M. and James, M.M. (2017), "Challenges faced by nurses in India-The major workforce of the healthcare system", *Nurse Care Open Access Journal*, Vol. 2 No. 4, pp. 112-114.

Chidbhavananda, S. (1983), *Daily Divine Digest*, Sri Ramakrishna TapovanamTirupparaitturai.

Chung, Y.W. (2018), "Workplace ostracism and workplace behaviors: a moderated mediation model of perceived stress and psychological empowerment", *Anxiety, Stress and Coping: An International Journal*, Routledge, Vol. 31 No. 3, pp. 304-317.

Colligan, T.W. and Higgins, E.M. (2006), "Workplace stress: etiology and consequences", *Journal of Workplace Behavioral Health*, Vol. 21 No. 2, pp. 89-97.

Conway, P.M., Høgh, A., Balducci, C. and Ebbesen, D.K. (2021), *Workplace Bullying and Mental Health", Pathways of Job-Related Negative Behaviour*, Springer, Singapore, pp. 101-128.

Dasgupta, P. (2015), "Turnover intentions among nurses in private hospitals: antecedents and mediators", *Journal of Strategic Human Resource Management*, Vol. 4 No. 3, pp. 1-15.

Davis, R.S., Stazyk, E.C. and Dickman, Z.T. (2021), "Advantages of feeling appreciated: an examination of how receipt of gratitude influences the linkages between PSM and behaviour", *Public Administration*, John Wiley & Sons, Vol. 99 No. 4, pp. 723-739.

Deery, S., Walsh, J. and Guest, D. (2011), "Workplace aggression: the effects of harassment on job burnout and turnover intentions", *Work, Employment and Society*, Vol. 25 No. 4, pp. 742-759.

Dhiman, N. and Arora, N. (2018), "Exploring the relationship of psycap and turnover intentions: a study among health professionals", *Amity Business Review*, Vol. 19 No. 1, pp. 106-118.

D'Cruz, P., Paull, M., Omari, M. and Guneri-Cangarli, B. (2016), "Target experiences of workplace bullying: insights from Australia, India and Turkey", *Employee Relations*, Emerald Group Publishing, Vol. 38 No. 5, pp. 805-823.

Fehr, R., Fulmer, A., Awtrey, E. and Miller, J.A. (2017), "The grateful workplace: a multilevel model of gratitude in organizations", *Academy of Management Review*, Academy of Management, Vol. 42 No. 2, pp. 361-381.

Fernet, C., Gagné, M. and Austin, S. (2010), "When does quality of relationships with coworkers predict burnout over time? The moderating role of work motivation", *Journal of Organizational Behavior*, John Wiley & Sons, Vol. 31 No. 8, pp. 1163-1180.

Fornell, C.G. and Larcker, D.F. (1981), "Evaluating structural equation models with unobservable variables and measurement error", *Journal of Marketing Research*, Vol. 18 No. 1, pp. 39-50.

Fredrickson, B.L. (2004), "The broaden-and-build theory of positive emotions", *Philosophical Transactions of the Royal Society B: Biological Sciences*, Royal Society, Vol. 359, pp. 1367-1377.

Frost, P. and Robinson, S. (1999), "The toxic handler organizational hero-and casualty", *Harvard Business Review*, Vol. 77 No. 4, pp. 96-106.

Gardner, S.K. (2012), "'I couldn't wait to leave the toxic environment': a mixed methods study of women faculty satisfaction and departure from one research institution", *NASPA Journal About Women in Higher Education*, Vol. 5 No. 1, Informa UK Limited, available at: DOI: 10.1515/NJAWHE-2012-1079.

Garg, N. (2017), "Workplace spirituality and organizational commitment: an empirical exploration. *Purushartha -*", *A Journal of Management, Ethics and Spirituality*, Vol. 10 No. 2, pp. 8-24.

Garg, N. (2020), "Exploring role of gratitude in developing teacher leadership in Indian universities", *International Journal of Educational Management*, Vol. 34 No. 5, pp. 881-901.

Garg, N. (2022), "Gratitude and work-family enrichment among Indian female workforce: exploring the mediating role of psychological capital", *International Journal of Work Organization and Emotion*, Vol. 13 No. 1, pp. 1-15.

Garg, N., Jain, A. and Punia, B.K. (2020), "Gratitude, social intelligence, and leadership among university teachers: mediation and moderation analysis", *International Journal of Organizational Analysis*, Vol. 29 No. 2, pp. 368-388.

Garg, N., Katiyar, N. and Mehak (2021), "Gratitude Questionnaire-6 (GQ6): psychometric properties in India", *Journal of Religion and Health*, Vol. 60 No. 5, pp. 3716-3731.

Garg, N. and Mehak (2021), "Gratitude resentment and appreciation scale (GRAT-16): analyzing psychometrics properties in the Indian context", *Current Psychology*, doi: 10.1007/s12144-021-02169-6.

Garg, N., Rohtagi, A., Mittal, C. and Sharma, D. (2022), *Combating Workplace Toxicity with the Help of Workplace Spirituality: A Study of Indian Public Sector Oil and Gas Companies*, International Journalof Public Administration, Routledge, pp. 1-13.

Garg, N. and Sarkar, A. (2020), "Vitality among university students: exploring the role of gratitude and resilience", *Journal of Organizational Effectiveness: People and Performance*, Vol. 7 No. 3, pp. 321-337.

Garg, N. and Saxena, A. (2020), "Analyzing the inter-relation between workplace spirituality and constructive deviance", *Asian Journal of Business Ethics*, Vol. 9 No. 1, pp. 121-141.

Gilbert, J.A., Carr-Ruffino, N., Ivancevich, J.M. and Konopaske, R. (2012), "Toxic versus cooperative behaviors at work: the role of organizational culture and leadership in creating community-centered organizations", *International Journal of Leadership Studies*, Vol. 7 No. 1, pp. 29-47.

Glazer, S., Farley, S.D. and Rahman, T.T. (2021), "Performance consequences of workplace ostracism", in Liu, C., M.J. (Ed.), *Workplace Ostracism*, Palgrave Macmillan, Cham, pp. 159-188.

Goldman, A. (2006), "Personality disorders in leaders: implications of the DSM IV-TR in assessing dysfunctional organizations", *Journal of Managerial Psychology*, Vol. 21 No. 5, pp. 392-414.

Goldstein, S.M. (2003), "Employee development: an examination of service strategy in a high-contact service environment", *Production and Operations Management*, Vol. 12 No. 2, pp. 186-203.

Green, J.E. (2014), "Toxic leadership in educational organizations", *Education Leadership Review*, Vol. 15 No. 1, pp. 18-33.

Greenglass, E.R., Burke, R.J. and Fiksenbaum, L. (2001), "Workload and burnout in nurses", *Journal of Community and Applied Social Psychology*, Vol. 11 No. 3, pp. 211-215.

Guan, B. and Jepsen, D.M. (2020), "Burnout from emotion regulation at work: the moderating role of gratitude", *Personality and Individual Differences*, Elsevier, Vol. 156, 109703.

Hair, J.F., Anderson, R.E., Tatham, R.L. and Black, W.C. (1998), *Multivariate Data Analysis*, 5th ed., Prentice-Hall, Upper Saddle River, NJ.

Hair, J.F., Black, W.C., Babin, B.J. and Anderson, R.E. (2010), *Multivariate Data Analysis*, 7th ed., Prentice-Hall, Upper Saddle River.

Hofmeyer, A. and Taylor, R. (2021), "Strategies and resources for nurse leaders to use to lead with empathy and prudence so they understand and address sources of anxiety among nurses practising in the era of COVID-19", *Journal of Clinical Nursing*, John Wiley & Sons, Vol. 30 Nos 1-2, pp. 298-305.

Howard, M.C., Cogswell, J.E. and Smith, M.B. (2020), "The antecedents and outcomes of workplace ostracism: a meta-analysis", *The Journal of Applied Psychology*, NLM (Medline), Vol. 105 No. 6, pp. 577-596.

Jans-Beken, L., Lataster, J., Peels, D., Lechner, L. and Jacobs, N. (2018), "Gratitude, psychopathology and subjective well-being: results from a 7.5-month prospective general population study", *Journal of Happiness Studies*, Vol. 19 No. 1, pp. 1673-1689, doi: 10.1007/s10902-017-9893-7.

Jans-Beken, L. and Wong, P.T. (2019), "Development and preliminary validation of the existential gratitude scale (EGS)", *Counselling Psychology Quarterly*, Vol. 34 No. 1, pp. 72-86.

Kesarwani, V., Husaain, Z.G. and George, J. (2020), "Prevalence and factors associated with burnout among healthcare professionals in India: a systematic review and meta-analysis", *Indian Journal of Psychological Medicine*, Indian Psychiatric Society South Zonal Branch, Vol. 42 No. 2, p. 108.

Kim, J. (2015), "What increases public employees' turnover intention?", *Public Personnel Management*, SAGE Publications, Vol. 44 No. 4, pp. 496-519.

Korunka, C., Hoonakker, P. and Carayon, P. (2008), "Quality of working life turnover intention in information technology work", *Human Factors and Ergonomics In Manufacturing*, Vol. 18 No. 4, pp. 409-423.

Krejtz, I., Nezlek, J.B., Michnicka, A., Holas, P. and Rusanowska, M. (2014), "Counting one's blessings can reduce the impact of daily stress", *Journal of Happiness Studies 2014*, Springer, Vol. 17 No. 1, pp. 25-39.

Kriakous, S.A., Elliott, K.A., Lamers, C. and Owen, R. (2020), "The effectiveness of mindfulness-based stress reduction on the psychological functioning of healthcare professionals: a systematic review", *Mindfulness 2020*, Vol. 12 No. 1, pp. 1-28, 12:1, Springer.

Kusy, M. and Holloway, E. (2009), *Toxic Workplace!: Managing Toxic Personalities and Their Systems of Power*, John Wiley & Sons.

Labrague, L.J., Nwafor, C.E. and Tsaras, K. (2020), "Influence of toxic and transformational leadership practices on nurses' job satisfaction, job stress, absenteeism and turnover intention: a cross-sectional study", *Journal of Nursing Management*, John Wiley & Sons, Vol. 28 No. 5, pp. 1104-1113.

Lambrou, P., Kontodimopoulos, N. and Niakas, D. (2010), "Motivation and job satisfaction among medical and nursing staff in a Cyprus public general hospital", *Human Resources for Health*, Vol. 8 No. 1, pp. 1-9.

Lanham, M., Rye, M., Rimsky, L. and Weill, S. (2012), "How gratitude relates to burnout and job satisfaction in mental health professionals", *Journal of Mental Health Counseling*, Vol. 34 No. 4, pp. 341-354.

Latino, R.J. (2008), *Patient Safety: The PROACT Root Cause Analysis Approach*, Taylor and Francis - CRC Press.

Lazzari, M., Alvarez, J.M. and Ruggieri, S. (2022), "Predicting and explaining employee turnover intention", *International Journal of Data Science and Analytics 2022*, Springer, Vol. 14 No. 3, pp. 279-292.

Lipman-Blumen, J. (2010), "Toxic leadership: a conceptual framework", *Handbook of Top Management Teams*, Palgrave Macmillan, pp. 214-220.

Liu, S. and Onwuegbuzie, A.J. (2012), "Chinese teachers' work stress and their turnover intention", *International Journal of Educational Research*, Pergamon, Vol. 53, pp. 160-170.

Măirean, C., Turliuc, M.N. and Arghire, D. (2019), "The relationship between trait gratitude and psychological wellbeing in university students: the mediating role of affective state and the moderating role of state gratitude", *Journal of Happiness Studies*, Pergamon, Vol. 20 No. 5, pp. 1359-1377.

Mahadi, N., Woo, N.M.F., Baskaran, S. and Yaakop, A.Y. (2020), "Determinant factors for employee retention: should I stay?", *International Journal of Academic Research in Business and Social Sciences*, Vol. 10 No. 4, pp. 201-213.

Majeed, M. and Naseer, S. (2021), "Is workplace bullying always perceived harmful? The cognitive appraisal theory of stress perspective", *Asia Pacific Journal of Human Resources*, John Wiley and Sons, Vol. 59 No. 4, pp. 618-644.

Martinko, M.J., Harvey, P., Brees, J.R. and Mackey, J. (2013), "A review of abusive supervision research", *Journal of Organizational Behavior*, Vol. 34 No. 1, pp. 120-137.

Maslach, C., Jackson, S. and Leiter, M. (1996), *Burnout Inventory Manual*, Consulting Psychologists Press, Palo Alto, CA.

Mathieu, C. and Babiak, P. (2016), "Corporate psychopathy and abusive supervision: their influence on employees' job satisfaction and turnover intentions", *Personality and Individual Differences*, Vol. 91, pp. 102-106, Pergamon.

McCullough, M.E., Emmons, R.A. and Tsang, J. (2002), "The grateful disposition: a conceptual and empirical topography", *Journal of Personality and Social Psychology*, Vol. 82 No. 1, pp. 112-127.

McNall, L.A., Masuda, A.D. and Nicklin, J.M. (2009), "Flexible work arrangements, job satisfaction, and turnover intentions: the mediating role of work-to-family enrichment", *The Journal of Psychology*, Vol. 144 No. 1, pp. 61-81.

McVicar, A. (2003), "Workplace stress in nursing: a literature review", *Journal of Advanced Nursing*, Vol. 44 No. 6, pp. 633-642.

Mehta, P., Garg, N. and Gharib, M. (2022), "Social support during COVID-19: exploring the psychometric properties of the PSS-JSAS and its relationship with job search activities", *Socio-Economic Planning Sciences*, p. 101349.

Moayed, F.A., Daraiseh, N., Shell, R., Salem, S., Moayedy, F.A., Daraisehy, N. and Shelly, R. (2010), "Workplace bullying: a systematic review of risk factors and outcomes", *Theoretical Issues in Ergonomics Science*, Taylor & Francis Group, Vol. 7 No. 3, pp. 311-327.

Mosadeghrad, A.M. and Ferdosi, M. (2013), "Leadership, job satisfaction and organizational commitment in healthcare sector: proposing and testing a model", *Materia Socio-Medica*, Vol. 25 No. 2, pp. 121-126.

Mosadeghrad, A.M., Ferlie, E. and Rosenberg, D. (2008), "A study of the relationship between job satisfaction, organizational commitment and turnover intention among hospital employees", *Health Services Management Research*, Vol. 21 No. 4, pp. 211-227.

Nayak, T., Sahoo, C.K. and Mohanty, P.K. (2018), "Workplace empowerment, quality of work life and employee commitment: a study on Indian healthcare sector", *Journal of Asia Business Studies*, Vol. 12 No. 2, pp. 117-136.

Nezlek, J.B., Krejtz, I., Rusanowska, M. and Holas, P. (2019), "Within-person relationships among daily gratitude, well-being, stress, and positive experiences", *Journal of Happiness Studies*, Springer, Vol. 20 No. 3, pp. 883-898.

Ng, L.P., Choong, Y.O., Kuar, L.S., Tan, C.E. and Teoh, S.Y. (2021), "Job satisfaction and organizational citizenship behavior amongst health professionals: the mediating role of work engagement", *International Journal of Healthcare Management*, Vol. 14 No. 3, pp. 797-804.

Nielsen, M.B. and Einarsen, S. (2012), "Outcomes of exposure to workplace bullying: a meta-analytic review", *Work and Stress: An International Journal of Work, Health and Organizations*, Taylor & Francis Group, Vol. 26 No. 4, pp. 309-332.

Nielsen, K., Yarker, J., Randall, R. and Munir, F. (2009), "The mediating effects of team and self-efficacy on the relationship between transformational leadership, and job satisfaction and psychological well-being in healthcare professionals: a cross-sectional questionnaire survey", *International Journal of Nursing Studies*, Vol. 46 No. 9, pp. 1236-1244.

Nielsen, M.B., Christensen, J.O., Finne, L.B. and Knardahl, S. (2020), "Workplace bullying, mental distress, and sickness absence: the protective role of social support", *International Archives of Occupational and Environmental Health*, Springer, Vol. 93 No. 1, pp. 43-53.

Niven, K., Connolly, C., Stride, C.B. and Farley, S. (2021), "Daily effects of face-to-face and cyber incivility via sadness, anger and fear", *Work and Stress, Routledge*, Vol. 36 No. 2, pp. 147-163.

Pandey, J. and Singh, M. (2015), "Asakti-Anasakti as mediator of emotional labor strategies and burnout: a study on ASHA workers", *Indian Journal of Industrial Relations*, Vol. 51 No. 1, pp. 57-69.

Perez, G.K., Haime, V., Jackson, V., Chittenden, E., Mehta, D.H. and Park, E.R. (2015), "Promoting resiliency among palliative care clinicians: stressors, coping strategies, and training needs", *Journal of Palliative Medicine*, Vol. 18 No. 4, pp. 332-337.

Podsakoff, P.M., MacKenzie, S.B. and Podsakoff, N.P. (2012), "Sources of method bias in social science research and recommendations on how to control it", *Annual Review of Psychology*, Vol. 63 No. 1, pp. 539-569.

Poon, J.M.L. (2004), "Effects of performance appraisal politics on job satisfaction and turnover intention", *Personnel Review*, Emerald Group Publishing, Vol. 33 No. 3, pp. 322-334.

Portocarrero, F.F., Gonzalez, K. and Ekema-Agbaw, M. (2020), "A meta-analytic review of the relationship between dispositional gratitude and well-being", *Personality and Individual Differences*, Pergamon, Vol. 164, 110101.

Poulose, S. and Sudarsan, N. (2017), "Assessing the influence of work-life balance dimensions among nurses in the healthcare sector", *Journal of Management Development*, Vol. 36 No. 3, pp. 427-437.

Pradhan, S. and Jena, L.K. (2018), "Emotional intelligence as a moderator in abusive supervision-intention to quit relationship among Indian healthcare professionals", *Asia-Pacific Journal of Business Administration*, Vol. 10 No. 1, pp. 35-49.

Pradhan, S. and Jena, L.K. (2018a), "Emotional intelligence as a moderator in abusive supervision-intention to quit relationship among Indian healthcare professionals", *Asia-Pacific Journal of Business Administration*, Emerald Publishing, Vol. 10 No. 1, pp. 35-49.

Pradhan, S. and Jena, L.K. (2018b), "Abusive supervision and job outcomes: a moderated mediation study", *Evidence-Based HRM: A Global Forum for Empirical Scholarship*, Emerald Publishing, Vol. 6 No. 2, pp. 137-152.

Rahman, W. and Nas, Z. (2013), "Employee development and turnover intention: theory validation", *European Journal of Training and Development*, Emerald Publishing, Vol. 37 No. 6, pp. 564-579.

Rai, A. and Agarwal, U.A. (2017), "Workplace bullying among Indian managers: prevalence, sources and bystanders' reactions", *International Journal of Indian Culture and Business Management*, Inderscience Publishers, Vol. 15 No. 1, p. 58.

Rai, A. and Agarwal, U.A. (2018), "Workplace bullying and employee silence: a moderated mediation model of psychological contract violation and workplace friendship", *Personnel Review*, Emerald Group Publishing, Vol. 47 No. 1, pp. 226-256.

Rasool, S.F., Wang, M., Tang, M., Saeed, A. and Iqbal, J. (2021), "How toxic workplace environment effects the employee engagement: the mediating role of organizational support and employee wellbeing", *International Journal of Environmental Research and Public Health*, Vol. 18, p. 2294, doi: 10.3390/ijerph18052294.

Reknes, I., Glambek, M. and Einarsen, S.V. (2021), "Injustice perceptions, workplace bullying and intention to leave", *Employee Relations*, Emerald Group Holdings, Vol. 43 No. 1, pp. 1-13.

Roodt, G. (2004), *Turnover Intentions. Unpublished Document*, University of Johannesburg, Johannesburg.

Rossiter, L. and Sochos, A. (2018), "Workplace bullying and burnout: the moderating effects of social support", *Journal of Aggression, Maltreatment and Trauma, Routledge*, Vol. 27 No. 4, pp. 386-408.

Said, H. and Tanova, C. (2021), "Workplace bullying in the hospitality industry: a hindrance to the employee mindfulness state and a source of emotional exhaustion", *International Journal of Hospitality Management, Pergamon*, Vol. 96, 102961.

Salin, D. and Notelaers, G. (2018), "The effects of workplace bullying on witnesses: violation of the psychological contract as an explanatory mechanism?", *The International Journal of Human Resource Management*, Routledge, Vol. 31 No. 18, pp. 2319-2339.

Santabárbara, J., Bueno-Notivol, J., Lipnicki, D.M., Olaya, B., Pérez-Moreno, M., Gracia-García, P., Idoiaga-Mondragon, N. and Ozamiz-Etxebarria, N. (2021), "Prevalence of anxiety in health care professionals during the COVID-19 pandemic: a rapid systematic review (on published articles in Medline) with meta-analysis", *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, Elsevier, Vol. 107, 110244.

Sarkar, A., Garg, N. and Palo, S. (2022), "Workplace spirituality and psychological capital (PsyCap) of the employees: a cross-sectional study", *Global Business Review*, doi: 10.1177/09721509221111445.

Saxena, A., Garg, N., Punia, B.K. and Prasad, A. (2020), Exploring role of Indian workplace spirituality in stress management: a study of oil and gas industry. *Journal of Organizational Change Management*, Vol. 33 No. 5, pp. 779-803.

Schaufeli, W.B. and Bakker, A.B. (2004), "Job demands, job resources, and their relationship with burnout and engagement: a multi-sample study", *Journal of Organizational Behavior*, John Wiley & Sons, Vol. 25 No. 3, pp. 293-315.

Sguera, F., Bagozzi, R.P. and Boss, W. (2011), "Workplace incivility and turnover intentions: the efficacy of managerial interventions", *Academy of Management Proceedings*, Vol. 2011 No. 1, pp. 1-65.

Shabir, S., Khan, O.F. and Gani, A. (2022), "Work-life interference: a perpetual struggle for women employees", *International Journal of Organizational Analysis*, Vol. 30 No. 2, pp. 181-196.

Shantz, A., Alfes, K. and Arevshatian, L. (2016), "HRM in healthcare: the role of work engagement", *Personnel Review*, Vol. 45 No. 2, pp. 274-295.

Sharma, N., Pathak, A., Lavanya, B.L., Garg, N. and Lata, K. (2022). Exploring the psychometric properties of personal optimism and self-efficacy optimism-extended (POSO-E) scale among Indian teachers. *Benchmarking: An International Journal*, Vol. ahead-of-print No. ahead-of-print, doi: 10.1108/BIJ-01-2022-0054.

Shen Hsiao, S.T., Ma, S.C., Guo, S.L., Kao, C.C., Tsai, J.C., Chung, M.H. and Huang, H.C. (2021), "The role of workplace bullying in the relationship between occupational burnout and turnover intentions of clinical nurses", *Applied Nursing Research*, W.B. Saunders, Vol. 68 No. 1, 151483.

Sim, B.Y.H., Lee, M.C.C., Kwan, S.S.M. and Tuckey, M.R. (2021), "The relationship between toxic leadership, job insecurity, workplace bullying and turnover intention in the Malaysian context: a multilevel mediational perspective", *Asian Perspectives on Workplace Bullying and Harassment*, Springer, Singapore, pp. 181-210.

Sjöberg, A. and Sverke, M. (2000), "The interactive effect of job involvement and organizational commitment on job turnover revisited: a note on the mediating role of turnover intention", *Scandinavian Journal of Psychology*, John Wiley & Sons, Vol. 41 No. 3, pp. 247-252.

Slattery, J.P. and Rajan Selvarajan, T.T. (2005), "Antecedents to temporary employee's turnover intention", *Journal of Leadership and Organizational Studies*, SAGE Publications, Vol. 12 No. 1, pp. 53-66.

Smith, N. and Fredricks-Lowman, I. (2020), "Conflict in the workplace: a 10-year review of toxic leadership in higher education", *International Journal of Leadership in Education*, Routledge, Vol. 23 No. 5, pp. 538-551.

Srinivasan, V. and Chandwani, R. (2014), "HRM innovations in rapid growth contexts: the healthcare sector in India", *The International Journal of Human Resource Management*, Vol. 25 No. 10, pp. 1505-1525.

Srivastava, S. and Agarwal, S. (2020), "Workplace bullying and intention to leave: a moderated mediation model of emotional exhaustion and supervisory support", *Employee Relations*, Emerald Group Holdings, Vol. 42 No. 6, pp. 1547-1563.

Srivastava, S. and Dey, B. (2019), "Workplace bullying and job burnout: a moderated mediation model of emotional intelligence and hardiness", *International Journal of Organizational Analysis*, Emerald Publishing, Vol. 28 No. 1, pp. 183-204.

Steensma, H., Van Breukelen, W. and Sturm, M. (2003), "Studying employee turnover by splitting up the usual comparison group", *Journal of Individual Employment Rights*, Vol. 11 No. 3, pp. 211-227.

Surya, M., Jaff, D., Stilwell, B. and Schubert, J. (2017), "The importance of mental well-being for health professionals during complex emergencies: it is time we take it seriously", *Global Health: Science and Practice*, Vol. 5 No. 2, pp. 188-196.

Talib, F. and Rahman, Z. (2013), "Current health of Indian healthcare and hospitality industries: a demographic study", *International Journal of Business Research and Development*, Vol. 2 No. 1, pp. 1-17.

Taştan, S.B. (2017), "Toxic workplace environment: in search for the toxic behaviors in organizations with a research in healthcare sector", *Postmodern Openings*, Vol. 8 No. 1, pp. 83-109.

Tastan, S.B. (2017), "Toxic workplace environment: in search for the toxic behaviors in organizations with a research in healthcare sector", *Postmodern Openings, Editura Lumen, Department of Economics*, Vol. 8 No. 1, pp. 83-109.

Tillman, C.J., Gonzalez, K., Crawford, W.S. and Lawrence, E.R. (2018), "Affective responses to abuse in the workplace: the role of hope and affective commitment", *International Journal of Selection and Assessment*, John Wiley & Sons, Vol. 26 No. 1, pp. 57-65.

Tiwari, M. and Jha, R. (2022), "Narcissism, toxic work culture and abusive supervision: a double-edged sword escalating organizational deviance", *International Journal of Organizational Analysis*, Emerald Group Holdings, Vol. 30 No. 1, pp. 99-114.

Top, M., Akdere, M. and Tarcan, M. (2015), "Examining transformational leadership, job satisfaction, organizational commitment and organizational trust in Turkish hospitals: public servants versus private sector employees", *The International Journal of Human Resource Management*, Vol. 26 No. 9, pp. 1259-1282.

Vasconcelos, A.F. (2020), "Workplace incivility: a literature review", *International Journal of Workplace Health Management*, Emerald Group Holdings, Vol. 13 No. 5, pp. 513-542.

Verma, S., Garg, N. and Arumugam, T. (2022), "Being ethically resilient during COVID-19: a cross-sectional study of Indian supply chain companies", *The International Journal of Logistics Management*, Vol. ahead-of-print No. ahead-of-print. doi: 10.1108/IJLM-05-2022-0203.

Wang, Z., Zaman, S., Rasool, S.F., Zaman, Q.U. and Amin, A. (2020), "Exploring the relationships between a toxic workplace environment, workplace stress, and project success with the moderating effect of organizational support: empirical evidence from Pakistan", *Risk Management and Healthcare Policy*, Dove Press, Vol. 13, p. 1055.

Watkins, P.C. (2004), "Gratitude and subjective well-being", in Emmons, R.A. and McCullough, M.E. (Eds), *Psychology of Gratitude*, Oxford University Press, pp. 167-195.

Webster, V., Brough, P. and Daly, K. (2016), "Fight, flight or freeze: common responses for follower coping with toxic leadership", *Stress and Health*, John Wiley & Sons, Vol. 32 No. 4, pp. 346-354.

Wilde, J. (2016), "The macro sources of toxicity in organizations", *The Social Psychology of Organizations: Diagnosing Toxicity and Intervening in the Workplace*, 1st ed., Taylor & Francis, pp. 111-124.

Williams, K.R. (2018), "Toxic leadership in defense and federal workplaces: sabotaging the mission and innovation", *International Journal of Public Leadership*, Emerald Publishing, Vol. 14 No. 3, pp. 179-198.

Zhou, X., Rasool, S.F. and Ma, D. (2020), "The relationship between workplace violence and innovative work behavior: the mediating roles of employee well-being", *Healthcare*, Vol. 8 No. 3, pp. 332-347.

Further reading

Day, G., Robert, G. and Rafferty, A.M. (2020), "Gratitude in health care: a meta-narrative review", *Qualitative Health Research*, SAGE PublicationsSage CA: Los Angeles, CA, Vol. 30 No. 14, pp. 2303-2315.

Corresponding author
Naval Garg can be contacted at: naval.garg@dtu.ac.in

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**

ECG Signals- Early detection of Arrhythmia using Machine Learning approaches

Jahnavi Nandanwar
Department of Software Engineering
Delhi Technological University
Delhi, India
jahnavinandanwar_2k19se056@dtu.ac.in

Jasmeet Singh
Department of Software Engineering
Delhi Technological University
Delhi, India
jasmeetsingh_2k19se057@dtu.ac.in

Sanjay Patidar
Department of Software Engineering
Delhi Technological University
Delhi, India
sanjaypatidar@dtu.ac.in

2023 13th International Conference on Cloud Computing, Data Science & Engineering (Confluence) | 978-1-6654-6263-1/23/$31.00 ©2023 IEEE | DOI: 10.1109/Confluence56041.2023.10048810

***Abstract* — Heartbeats can be recorded in the form of electrical signals using a machine called an electrocardiogram. The cardiovascular system's performance can be tracked using an electrocardiogram (ECG), which can also be used to diagnose various cardiac diseases. The paper by Mohammad Kachuee *et al.*, entitled "ECG Heartbeat Classification: A Deep Transferable Representation" classifies heartbeats into five different classes according to AAMI EC57 standards. In this paper, we continue the work on their preprocessed dataset and train the model to detect arrhythmia, i.e., abnormalities in the heartbeat. We also propose a comparative study between different machine learning and deep learning models, namely random forest classification, support vector machines, and convolutional neural networks, for detecting abnormalities in heartbeats. The results we got after the application of all our classifiers showed the best accuracy score of 98.26% for the MIT-BIH dataset and a 100% accuracy score for the PTB Diagnostic dataset.**

Keywords— ECG, deep learning, arrhythmia, heartbeat classification, random forest, SVM, CNN.

I. Introduction

A heartbeat is generated when two parts of the heart work together to pump blood in about a second. To monitor the functionalities of these beats manually has become a daunting task for the cardiologists. So, ECG is a safe and reliable technique now in common practice to monitor cardiac health.

An ECG (electrocardiogram) is used to check the heart's rhythm and a device helps display the electrical activity. These electrical signals are detected by the sensors possibly attached to the upper surface of the body on skin. This helps in getting each of the heartbeat signals recorded within a timeframe. The signals are further viewed and examined by the doctor and it is on doctor's skill to detect any anomalies present in the chart or in any particular case the early detection of such diseases. The requested ECG is mainly collected by a cardiologist (heart's doctor) or a general practitioner (GP). These recorded signals can help in detecting any abnormalities in the heart's rhythm. We analyze the abnormalities i.e. arrhythmias according to five classes previously classified by Mohammad Kachuee *et al*. [1] in their paper "ECG Heartbeat Classifier: A Deep Transferable Representation".

They classified the heartbeats using deep convolutional neural networks, which resulted in accurately classifying five different arrhythmias according to AAMI EC57 standards. We continue our study on the preprocessed dataset and apply two different techniques to detect arrhythmias.

A. HeartBeats

The human heart makes a lub-dub sound which we refer to as a heartbeat. This is actually a part of the blood pumping process which takes almost around a second. A human heart beats with the intensity of 60 to 100/minute. So, this pumping action is a two step process namely, *Systole,* the period of contraction when blood pours out of the chambers, *follows diastole,* the period of relaxation when the chambers fill with blood. The heart is divided into four chambers.

In the right side of the heart the upper right chamber called the atrium receives the deoxygenated blood from the body. The blood then passes into the lower right chamber known as the ventricle through a valve, the tricuspid valve. This deoxygenated blood is pumped into the artery which takes it to the lungs to absorb oxygen. This artery is called pulmonary artery and the valve through which it is pumped is called the pulmonary valve. Now, this blood is oxygenated. This blood now reaches the upper left chamber, the left atrium using the pulmonary veins. Next, the oxygenated blood reaches the lower left chamber, the left ventricle through a valve known as the mitral valve. The left ventricle pushes blood into a large artery, the Aorta, through a valve, the aortic valve. This is the largest artery which takes the oxygen rich blood into the different parts of the body. The body takes up all the oxygen from the blood and the process continues. The oxygenated and deoxygenated blood in the heart are separated in the right and left sides by a valve called septum.

This pumping action of the heart needs a signal to detect when to contract or relax. This starts with the sinoatrial node at the atria (when the contract to pump blood into ventricles). The heart has muscles which receive electrical signals using a node, the atrio-ventricular node. This node represents a cell cluster in the area of atria and ventricles convergence. The system created by these signals is called the conducting system. This causes a heartbeat in a normal heart [15].

B. Arrhythmia

Any irregularity in the heartbeat or any deviation of beats from the normal behavior is called Arrhythmia. Defects in the signals of the heart can be the cause for having tachycardia (beating too fast), bradycardia (beating

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 01,2023 at 07:23:01 UTC from IEEE Xplore. Restrictions apply.

too slow), or not in a normal rhythm. Arrhythmia may be harmless at times such as during running or exercise or may be a cause of a serious issue. According to the Association for the Advancement of Medical Instrumentation (AAMI) EC57 the arrhythmia detection software must be able to discriminate five different types of heartbeat: Ventricular ectopic beat (V), supraventricular ectopic beat (S), ventricular fusion beat (F), ambiguous beat (Q), and other beats or non ectopic beats (N). In this paper, we have focussed our attention on the early detection and prediction of arrhythmia causes due to these five types of heartbeats [16].

C. *Electrocardiogram*

The graphical representation of the electrical activities of the heart is shown using a device called electrocardiogram (ECG). In a normal heart when the atria is full of blood, the SA node fires which is represented by a P-wave on ECG. A PQ segment is formed at 100 ms of P wave as the signal travels from SA node to AV node (Atrial contraction, Sibili). Now, the firing of AV nodes takes place and ventricular depolarization forms the QRS component. The Q wave represents the depolarization of interventricular septum. The R wave represents the depolarization of the main mass of ventricles. The last phase of ventricular depolarization forms the S wave at the base of the heart. Atrial repolarization takes place at the same time. After this the ventricles contract and pump blood and the S-T segment is formed. Ventricular repolarization takes place immediately before ventricular relaxation/ diastole. These electrical signals are recorded in a timely manner by using electrodes. The Fig. 1 below explains a single heartbeat of a normal human as presented on an ECG[17].

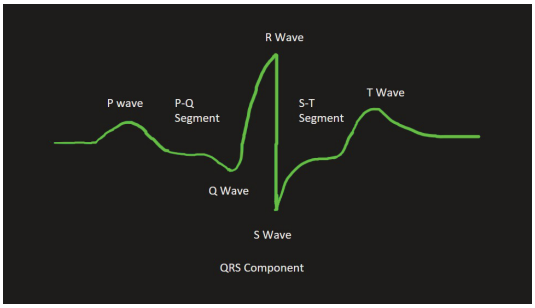


Fig. 1 Single heartbeat on ECG

D. *Selecting a Machine Learning model*

After selecting the attributes for creating a classification model, now we must make use of a machine learning technique for accomplishing this. Supervised learning methods offer higher performance over their unsupervised counterparts and hence we shall make use of supervised algorithms. The more popular of these techniques are decision trees, and random forest, which is an ensemble learning classification technique.

E. *Decision Trees Classification (DTC)*

The Decision Trees Classification (DTC) is a supervised learning model used in non-parametric classification and categorisation. It is one of the most effective and powerful techniques used for classification datasets.

A decision tree is represented by a flowchart-like structure which shows the ‘test’ and their ‘results’. Each leaf node represents a class label combining decisions of all attributes, a branch node represents individual outcome of a test and each internal node represents a test on a single attribute. Classification rules are represented by the routes from root to leaf.

Despite being a common technique for classification tasks, decision trees have drawbacks of their own, with overfitting being the most obvious. When a statistical model matches its training data exactly, overfitting occurs. Overfitting in decision trees happens when the tree is built to exactly fit every sample in the training data set. As a result, it produces branches that strictly adhere to sparse data criteria. As a result, this influences the precision of predictions for samples that are not included in the training set.

F. *Random Forest Classifier (RFC)*

Seeing how decision trees offer a great way for tackling classification problems however at an expense. In order to overcome these issues, another supervised machine learning technique which can be used is Random Forest Classifier (RFC). This technique is another supervised learning technique in the category of ensemble learning used for classification and regression or other tasks that works by building several decision trees during the training phase. The class that the majority of the trees chose is the output of the random forest for classification problems. This process can be understood by Fig. 2 shown below. It shows that the different kinds and varied forms of decision trees create a "forest" according to the results . This is also trained via bootstrap aggregation or bagging processes. Other training processes may be applied to increase the accuracy of the classifier.

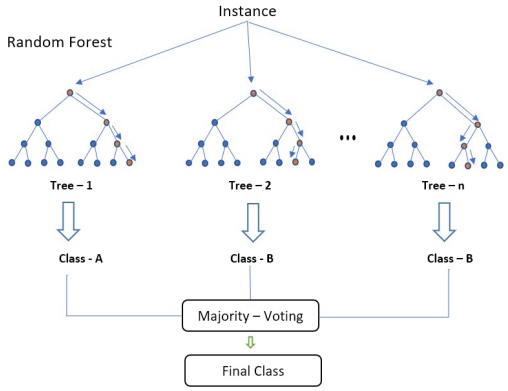


Fig. 2 Random Forest

The random forest algorithm chooses the outcome based on the decision trees' predictions. By averaging the outcomes from various trees, it provides predictions. The accuracy of the outcome improves with the rise in tree numbers. A random forest algorithm overcomes the drawbacks of the decision tree technique. It increases accuracy and reduces dataset overfitting [18],[19].

G. *Convolutional Neural Network (CNN)*

A Convolutional Neural Network (CNN) is a deep learning algorithm which extracts advanced features from an input image. It is mostly used for visual images. It is a

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 01,2023 at 07:23:01 UTC from IEEE Xplore. Restrictions apply.

system of a network of neurons that are fully connected to each other to present outcomes known as multilayer perceptrons. This technique seems to be a better version of other conventional machine learning models. CNN works by taking inputs in the first layer of the neurons, then attaching variances and biases in the hidden layers, passing one's results to the next one and finally outputs the results of the combined hidden perceptrons in the output layer. A CNN model requires substantially less preprocessing in comparison to the other classification techniques. While filters in outdated techniques are hand-engineered, with the right training, CNNs can pick up these filters and attributes.

A CNN structure established was under the influence of the way the human visual cortex is arranged and seems similar to the neural connectivity patterns in the human brain. Only the 'Receptive Field' is an area Individual neurons only respond to stimuli in this restricted region of the visual field, known as the Receptive Field. There are several overlapping fields like this that make up the total visual field [20],[21].

II. Objective

The objective of this report is to train different machine learning models on the preprocessed dataset by Mohammad Kachuee *et al.*[1] originally extracted from the PhysionNet's MIT-BIH and PTB Diagnostics datasets and see if we are able to detect any anomalies in the five different classes of heartbeats under study. We will also try to find the best results in the prediction of arrhythmia and if it is of serious type. We will aim towards asking the following research questions.

- What is the accuracy in prediction for random forest classifier?
- What is the accuracy in prediction for a support vector machine?
- What is the accuracy in prediction for Convolutional Neural Networks?

III. Related Work

Many have tried to work out their models for the early detection of Arrhythmia by using different machine learning or deep learning techniques. Here we focused our attention on the works including the physionet's MIT-BIH and PTB diagnostic datasets. Mohammad Kachuee, Shayan Fazeli, and Majid Sarrafzadeh conducted a study and proposed a method using CNN that can easily classify five different arrhythmias according to AAMI standards [1]. Their accuracies were about 93.4% and 95.9% for the proposed methods. Another study conducted by Yang Liu, Qince Li, Runnan He, Kuanquan Wang, Jun Liu, Yongfeng Yuan, *et al.,* introduced WSDL-AD (Weakly supervised deep learning framework for Arrhythmia Detection) framework which has shown around 10% increased accuracy than other pretrained models [2]. Another research was based on classification of micro-classes of Arrhythmia dataset (Mengze Wu, Yongdi Lu *et al.*)[3] and divided the train and test sets into more generalized subjects (Safa Sultan Qurraie *et al.*)[4]. Others tried on real time arrhythmia detection using several different techniques(Jing Cai *et al*)[5],(Sandeep Chandra Bollepalli *et al*)[6],(Vessela Krasteva *et al*)[7],(T. Petty *et al*)[8]. Other CNN techniques were used in (Pranav Rajpurkar *et al*)[9],(Ali Isin and Selen Ozdalili)[10]. Some other authors used nonlinear deep learning techniques (M. I. Owis, A. H. Abou-Zied *et al*)[11],(Zahra Ebrahimi, Mohammad Loni *et al*)[12],(Meng-Hsi Wu and Edward Y Chang)[13]. Others used 2D-CNN models.(Cuihua Tian, Yiping Zhang)[14].

IV. Experimental Design

Before we begin to classify heartbeats, first we must understand and recognize the various attributes on which classification can be performed and how these attributes are present in each class.

After completing this task, we must decide upon and select a machine learning model to create our design. And as we go forward, we shall see why we have chosen Random Forest Classifier (RFC) and Convolutional Neural Network (CNN) for this purpose.

Finally, we will present a comparison between the two models and draw a conclusion as to which model and validation technique is best suited for the task of classifying heartbeats and detecting Arrhythmia at early stages.

A. About the Datasets

The dataset contains the electrical activities recorded by the Electrocardiogram and the preprocessing done by the authors of [1] paper. The original dataset contains ECG Signals recorded and uploaded by PhysioNet entitled as MIT-BIH and PTB Diagnostic databases.

Further each signal was splitted into 10 sec windows and divided into five classes according to AAMI EC57 standards. The preprocessing done on the dataset is described in section III A of the above mentioned paper. The recordings were from 47 different subjects recorded at the sampling rate of 360Hz. Each beat was annotated by two cardiologists at least.

While selecting the correct attributes, we must keep in mind the closeness of an attribute with the different classes and how well it is able to distinguish between different classes independently i.e., without the availability of any other feature which may help in distinguishing between heartbeats of different classes. In addition, we must not take 2 or more highly inter-dependent attributes so that there is less redundancy and we can create a more efficient model.

Keeping this in mind, and taking into consideration the restrictions of the datasets that are available for use, we have selected five major attributes that shall be used for the classification process. This dataset is divided into five different classes in accordance with the beat annotations and AAMI EC57 categories.

N: Non-Ectopic Beats
S: Supraventricular Ectopic Beats
V: Ventricular Ectopic Beats
F: Fusion Beats
Q: Unknown Beats

The PTB Diagnostics dataset consists of ECGs with the matching measured Frank lead ECGs or the 549 conventional 12-lead resting ECGs are both present in the databases. The ECGs are digitized at 1000Hz (0.5 V/LSB; 16 Bit ADC) sampling rate.

This database may be made available at sample rates of up to 10,000Hz upon special request with the record of 290 subjects.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 01,2023 at 07:23:01 UTC from IEEE Xplore. Restrictions apply.

148 patients had myocardial infarction. 18 points for cardiomyopathy and heart failure. Branch block bundled for 15 points. Dysrhythmia 14 points. 7 points for myocardial hypertrophy. 6 points for valvular heart disease. Myocarditis scored four points. Inconsequential 5 points. A healthy population is 54 people.

The PTB Diagnostics dataset is further divided into PTB normal beats and PTB abnormal beats. In this project, we have analyzed the signals of each.

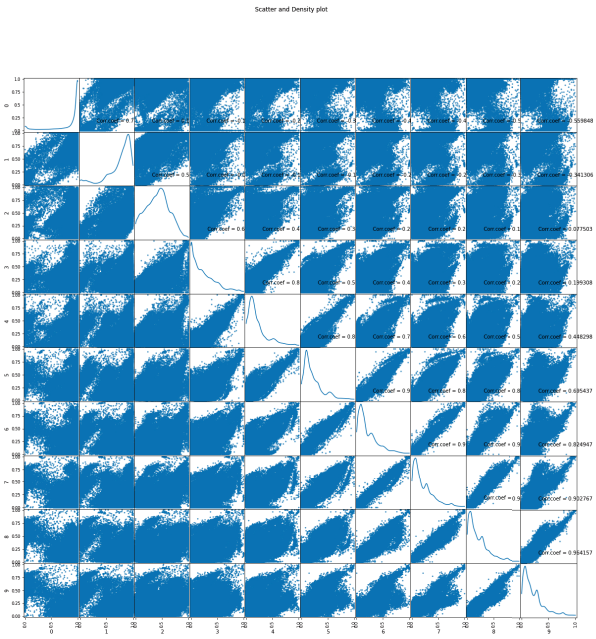
V. Research Methodology

Here we suggest the methodologies and processes we used during the whole process from visualizing, preprocessing to training and testing our models on the given dataset. We run the models on the dataset and check the results of the decision tree, random forest models and CNN algorithms. The results of all these models were compared using statistical tests which allowed us to decide which model should be most likely used.

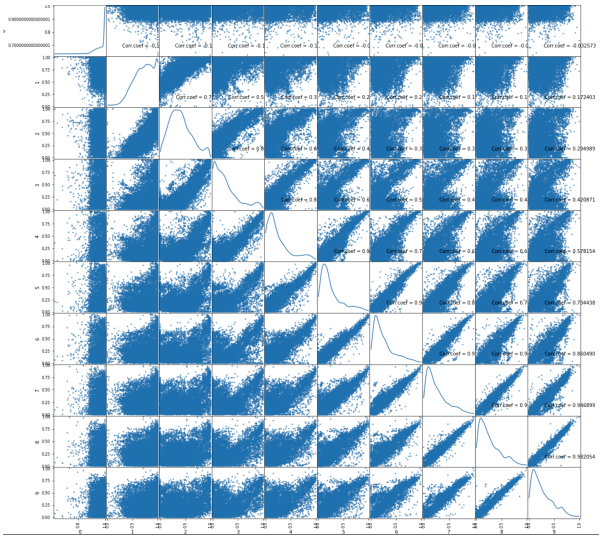
The code for our work is uploaded and available at https://github.com/JD771/ECG-Signals--Early-detection-of-Arrhythmia-using-Machine-Learning-approaches.

A. Analyzing the Dataset

First we analyze the data and plot different graphs to better understand our dataset. For this analysis we normally describe and name different data according to our convenience. Then we have created a function to plot scatter plots of combining all the rows and columns of a dataset. Similarly, one more function has been created for plotting a correlation matrix of all the values of these datasets. The Fig. 3a below shows the scatter plot of MIT_train and Fig. 3b shows the scatter plot of PTB_abnormal dataset.



(a) MIT train dataset scatterplot

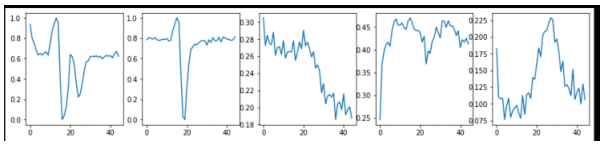


(b) PTB Abnormal Scatter Plot

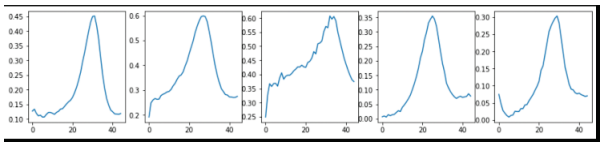
Fig. 3 Scatter plot of MIT_train and PTB_abnormal datasets

B. Balancing the Classes on MIT data and Removing Bias

On plotting the pie chart of different beats recorded in the MIT BIH dataset, there was a huge bias in the distribution of classes. So, to remove this bias and balance the classes, we first divided the training data into 4 different classes of equal distribution & plotted some signal frequencies with respect to normal beats. We then resampled the data values using Scikit Learn's resample module. Fig. 4 shows (a) the Abnormal heartbeat signals (10s window) recorded in ECG and plotted as a graph and (b) the Normal heartbeat signals plotted for five different classes of heartbeats which we worked on this project.



(a) PTB Abnormal Beats

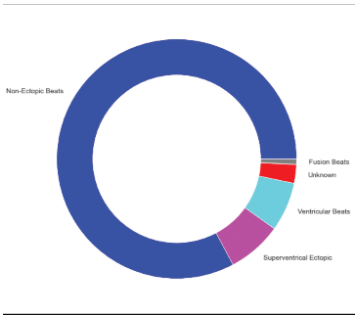


(b) PTB Normal Beats

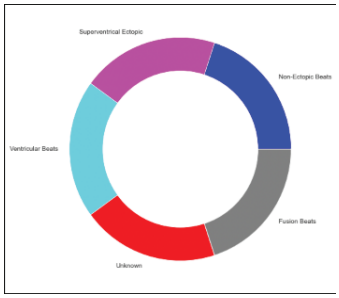
Fig. 4 ECG of five different beats

The Fig. 5 below is a plot of a pie chart showing (a) the bias or unbalance between the five classes of heartbeats and (b) how these classes have been affected after the resampling and balancing of each of them.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 01,2023 at 07:23:01 UTC from IEEE Xplore. Restrictions apply.



(a) Unbalanced Classes MIT_train



(b) Balanced Classes MIT_train

Fig. 5 Classes before and After Balancing

C. *Splitting the dataset into training and testing*

For supervised machine learning techniques, two different sets, one is used to train the model and the second for testing it. These are referred to as training and testing sets. Thus, we are first required to split the dataset into training and testing sets. But this process is already completed in the dataset provided which has a MIT-train dataset of 80% train split and MIT-test dataset with 20% test split. Same has been done for the PTB diagnostic dataset.

Thus, there is no need to further classify the models into splitting ratios and directly move forward towards the application of the models.

D. *Applying the Machine Learning models*

After splitting our dataset into training and testing sets, we now apply our models for classification. The random forest classification model is fitted on the training sequences using the model.fit() of the RandomForest classifier. After fitting, we tested the model by letting it predict on the test set as model.predict (on both MIT_test and PTB_test splits).

After application of the classifier the model is tested for accuracy using the performance classifier, which is accuracy and then plot a classification report of both our tested sets. The accuracy of a model is simply the number of correctly identified cases by the model over the total number of cases in the dataset. We also record f1 scores for different algorithms.

Next, we move onto the Convolutional Neural network model. Here, we define a model using the parameters provided in the table at the end of the paragraph. We ran this model for training upto 50 epochs with relu and softmax as the activation functions.

Again, we used the accuracy performance classifier to determine the performance of the model and plotted a classification report.

```
Metal device set to: Apple M1
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv1d (Conv1D)              (None, 187, 64)           448
_________________________________________________________________
batch_normalization (BatchNo (None, 187, 64)           256
_________________________________________________________________
max_pooling1d (MaxPooling1D) (None, 94, 64)            0
_________________________________________________________________
conv1d_1 (Conv1D)            (None, 94, 64)            24640
_________________________________________________________________
batch_normalization_1 (Batch (None, 94, 64)            256
_________________________________________________________________
max_pooling1d_1 (MaxPooling1 (None, 47, 64)            0
_________________________________________________________________
conv1d_2 (Conv1D)            (None, 47, 64)            24640
_________________________________________________________________
batch_normalization_2 (Batch (None, 47, 64)            256
_________________________________________________________________
max_pooling1d_2 (MaxPooling1 (None, 24, 64)            0
_________________________________________________________________
flatten (Flatten)            (None, 1536)              0
_________________________________________________________________
dense (Dense)                (None, 64)                98368
_________________________________________________________________
dense_1 (Dense)              (None, 64)                4160
_________________________________________________________________
dense_2 (Dense)              (None, 5)                 325
=================================================================
Total params: 153,349
Trainable params: 152,965
Non-trainable params: 384
```

Table showing the cnn model parameters.

E. *Checking the Performance*

The results of these models were now checked for their performance. A confusion matrix drawn helped us to determine whether or not there is a difference between random forest classifier, SVC and CNN, and if there is a difference, which of the models is superior in our task for classifying heartbeats.

VI. Results

For conducting the tests, we have taken the dataset in which the ECG signals of the preprocessed datasets of MIT-BIH Arrhythmia and the PTB diagnostic were chosen. The task we did was to remove the vias between the classes of the MIT database and apply different machine learning algorithms and classifiers on the organized dataset. The splitting was done using the traditional train-test split method in a 20-80 ratio.

Upon conducting the two models, we have obtained the following results:

A. *Random Forest*

After running the code on the Jupyter Notebook tool using the Python programming language, we were successfully able to classify the ECG signals in the dataset for the random forest model. Results of the model are as follows:

(i) The model has an accuracy on the MIT train dataset was 1.0 and on the MIT test dataset it was 0.9669.
(ii) According to the classification report obtained by the model the weighted average of precision score is 0.98, recall is 0.97, & the f1-score is 0.97.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 01,2023 at 07:23:01 UTC from IEEE Xplore. Restrictions apply.

(iii) The accuracy of the model on the PTB train dataset was 1.0 and on the PTB test dataset was also 1.0.
(iv) According to the classification report obtained by the model the weighted average of precision score is 1.0, recall is 1.0, & the f1-score is 1.0.

B. SVM

The results obtained from the support vector machine are as follows.
(i) The accuracy for the MIT train split was 0.9457 and the accuracy for MIT test split was 0.9032.
(ii) According to the classification report obtained by the model the weighted average of precision score is 0.95, recall is 0.90 & the f1-score is 0.92.
(iii) The accuracy of the model on the PTB train dataset was 1.0 and on the PTB test dataset was also 1.0.
(iv) According to the classification report obtained by the model the weighted average of precision score is 1.0, recall is 1.0, & the f1-score is 1.0.

C. CNN

Again, after running the code on jupyter lab in python, we have successfully made a CNN model for classification of ECG signals and have come up with the following results:

(i) After running the model for 50 epochs and 3 dense layers the accuracy of the test dataset is 0.9826.
(ii) According to the classification report obtained for this multinomial dataset the weighted average of precision, recall and the f1-score is 0.98.

VII. Discussion

We have successfully implemented all the conventional machine learning and deep learning techniques on the classification and early detection of arrhythmia on the preprocessed datasets of MIT-BIH and PTB Diagnostic heart ECG signals. These algorithms provide an accuracy acceptable to be used for research purposes. The classifiers have shown that the attributes provided to them in the dataset correspond well and seem the right amount on the basis to classify signals of different types of cardiac patients.

We shall compare the different models that were used for evaluating the dataset. The dataset was divided into 80% train data and 20% test data. The train data was further divided into 4 balanced classes in order to remove bias from the dataset. Let's go through each of the models separately and then compare the results.

A random forest classifier is a collection of numerous decision tree classifiers and an ensemble technique. As a result, a random forest classifier generates a large number of decision trees for various sorts of signals, making it a useful tool for detecting signal anomalies. Here, the accuracy of the model on MIT test data comes out to be 96.7% and f1 score is 98% while for the PTB data the accuracy score of RF is 100% and f1 score is also 100% which is very good considering the bias which was obtained and the diversified data.

Next model we have used is the Support Vector machine. This classification technique divides the n-dimensional space into separate classes based on a hyperplane. For this we got the accuracy for MIT test data as 94.6% and f1 score as 92% while for the PTB diagnostic data we got both the accuracy and the f1 score of 100%.

A convolutional neural network is presumed as a neural network intended specifically to process image and pixel input. As a result, it can successfully manage new unprocessed signals. The results obtained on the MIT data was an accuracy of 98.26%.

In order to check if there is a difference in the performance of the models, we can definitely say that the CNN model is much better than the random forest model or SVM. Further, we have also created a confusion matrix and a classification report for comparing the two models from which we can conclude that there is a significant difference between the performance of the models under different evaluation techniques. Here is the accuracy and loss graph for the CNN model in Fig. 6.

Accuracy curve

Loss curve

Fig. 6 Accuracy and loss graph for the CNN model

To interpret the results better we have also plotted the heatmap of the confusion matrix in Fig. 7.

Fig. 7 Heatmap of the confusion matrix

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 01,2023 at 07:23:01 UTC from IEEE Xplore. Restrictions apply.

VIII. Conclusion and Future Work

So, after the application of modern machine learning and data analysis techniques, we were able to successfully classify ECG signals on the basis of different types of heartbeats available in the dataset. Here, we concluded that the overall performance of CNN model is better than the performance of the Random Forest Model or SVC.

Further developments would include building deep learning-based models by using techniques BERT or SCI-BERT, etc. Moreover, we would try linking the brain functions with the ECG signals such that we can gain early diagnosis of any underlying diseases at the early stages of behavioral anomaly. A further development would include the extraction of these features from live images collected from the ECG monitors in a live working environment of a hospital, or ECG center. Then that model can be further improved to classify which are safe signals and which are an alarm to the human race.

IX. Acknowledgements

We would like to express our gratitude to our professor, Mr. Sanjay Patidar, for having provided us with this incredible opportunity to work on this project and for guiding us during the course of this semester without which we could not have completed this project. We have learnt a lot from this project about the vast fields of data analytics and machine learning and hope to continue to learn more.

X. References

[1] M. Kachuee, S. Fazeli and M. Sarrafzadeh, "ECG Heartbeat Classification: A Deep Transferable Representation," 2018 IEEE International Conference on Healthcare Informatics (ICHI), New York, NY, USA, 2018, pp. 443-444, doi: 10.1109/ICHI.2018.00092..

[2] Liu Yang, Li Qince, He Runnan, Wang Kuanquan, Liu Jun, Yuan Yongfeng, Xia Yong, Zhang Henggui, "Generalizable Beat-by-Beat Arrhythmia Detection by Using Weakly Supervised Deep Learning", Frontiers in Physiology, Vol: 13, 2022, Doi: 0.3389/fphys.2022.850951, ISSN: 1664-042X.

[3] Wu Mengze, Lu Yongdi, Yang Wenli, Wong Shen Yuong, "A Study on Arrhythmia via ECG Signal Classification Using the Convolutional Neural Network", Frontiers in Computational Neuroscience, Vol: 14, 2021, Doi: 10.3389/fncom.2020.564015, ISSN: 1662-5188.

[4] Sultan Qurraie, S., Ghorbani Afkhami, R. ECG arrhythmia classification using time frequency distribution techniques. *Biomed. Eng. Lett.* 7, 325–332 (2017). https://doi.org/10.1007/s13534-017-0043-2

[5] Jing Cai, Ge Zhou, Mengkun Dong, Xinlei Hu, Guangda Liu, Weiguang Ni, "Real-Time Arrhythmia Classification Algorithm Using Time-Domain ECG Feature Based on FFNN and CNN", Mathematical Problems in Engineering, Vol 2021, Jilin University, China, May 2021, Article ID 6648432, 10.1155/2021/6648432.

[6] Sandeep Chandra Bollepalli, Rahul K. Sevakula, Wan Tai M. Au Yeung, Mohamad B. Kassab, Faisal M. Merchant, George Bazoukis, et al., "Real Time Arrhythmia Detection Using Hybrid Convolutional Neural Networks", Journal of the American Heart Association, 2021; vol: 10, No. 23, 10:e023222, 2 Dec 2021.

[7] Vessela Krasteva, Irena Jekova, Remo Leber, Ramun Schmid and Roger Abächerli, "Real-time arrhythmia detection with supplementary ECG quality and pulse wave monitoring for the reduction of false alarms in ICUs", Physiological Measurement, vol. 37, Number 8, 2016 Institute of Physics and Engineering in Medicine, Physiol. Meas. 37 1273, DOI: 10.1088/0967-3334/37/8/1273, 25 July 2016.

[8] T. Petty et al., "Evaluating Deep Learning Algorithms for Real-Time Arrhythmia Detection," 2020 IEEE/ACM International Conference on Big Data Computing, Applications and Technologies (BDCAT), 2020, pp. 19-26, doi: 10.1109/BDCAT50828.2020.00022.

[9] Pranav Rajpurkar, Awni Y. Hannun, Masoumeh Haghpanahi, Codie Bourn and Andrew Y. Ng, "Cardiologist-Level Arrhythmia Detection with Convolutional Neural Networks", Computer Vision and Pattern Recognition, ver: 1, Thu, 6 July 2017, 1707.01836.

[10] Ali Isin and Selen Ozdalili, "Cardiac arrhythmia detection using deep learning", Procedia Computer Science, Vol 120, 2017, Pages 268-275, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2017.11.238.

[11] M. I. Owis, A. H. Abou-Zied, A. . -B. M. Youssef and Y. M. Kadah, "Study of features based on nonlinear dynamical modeling in ECG arrhythmia detection and classification", IEEE Transactions on Biomedical Engineering, vol. 49, no. 7, pp. 733-736, July 2002, doi: 10.1109/TBME.2002.1010858.

[12] Zahra Ebrahimi, Mohammad Loni, Masoud Daneshtalab, Arash Gharehbaghi, "A review on deep learning methods for ECG arrhythmia classification", Expert Systems with Applications: X, Volume 7, 2020, 100033, ISSN 2590-1885, https://doi.org/10.1016/j.eswax.2020.100033.

[13] Meng-Hsi Wu and Edward Y Chang, "DeepQ Arrhythmia Database: A Large-Scale Dataset for Arrhythmia Detector Evaluation", MMHealth '17: Proceedings of the 2nd International Workshop on Multimedia for Personal Health and Health Care, October 2017, Pages 77–80, https://doi.org/10.1145/3132635.3132647

[14] Cuihua Tian, Yiping Zhang, Jingmin Gao, Zhigang Hu, "Arrhythmia Classification Using 2D-CNN Models", CCEAI 2022: The 6th International Conference on Control Engineering and Artificial Intelligence, March 2022, Pages 34–37, https://doi.org/10.1145/3522749.3523080.

[15] "Heartbeat", The Texas Heart Institute.

[16] H. Society, "Heart diseases and disorders," 2018.

[17] World Health Organization, "Cardiovascular Diseases (CVDs)", 11 June 2021.

[18] Smith, Chris; Koning, Mark; "Decision Trees and Random Forests: A Visual Introduction For Beginners" ISBN 13: 9781549893759.

[19] Wikipedia, "Random Forest", 24 December 2022.

[20] O. T. Inan, L. Giovangrandi, and G. T. Kovacs, "Robust neural network-based classification of premature ventricular contractions using wavelet transform and timing interval features," IEEE Transactions on Biomedical Engineering, vol. 53, no. 12, pp. 2507–2515, 2006

[21] L. v. d. Maaten and G. Hinton, "Visualizing data using t-sne," Journal of Machine Learning Research, vol. 9, no. Nov, pp. 2579–2605, 2008.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 01,2023 at 07:23:01 UTC from IEEE Xplore. Restrictions apply.

Materials Research Express

ACCEPTED MANUSCRIPT • OPEN ACCESS

Effect of fiber synergy on strength and durability of high-performance concrete

To cite this article before publication: Mohit Gupta *et al* 2023 *Mater. Res. Express* in press https://doi.org/10.1088/2053-1591/acb75a

Manuscript version: Accepted Manuscript

Accepted Manuscript is "the version of the article accepted for publication including all changes made as a result of the peer review process, and which may also include the addition to the article by IOP Publishing of a header, an article ID, a cover sheet and/or an 'Accepted Manuscript' watermark, but excluding any other editing, typesetting or other changes made by IOP Publishing and/or its licensors"

This Accepted Manuscript is **© 2023 The Author(s). Published by IOP Publishing Ltd**.

As the Version of Record of this article is going to be / has been published on a gold open access basis under a CC BY 3.0 licence, this Accepted Manuscript is available for reuse under a CC BY 3.0 licence immediately.

Everyone is permitted to use all or part of the original content in this article, provided that they adhere to all the terms of the licence https://creativecommons.org/licences/by/3.0

Although reasonable endeavours have been taken to obtain all necessary permissions from third parties to include their copyrighted content within this article, their full citation and copyright line may not be present in this Accepted Manuscript version. Before using any content from this article, please refer to the Version of Record on IOPscience once published for full citation and copyright details, as permissions may be required. All third party content is fully copyright protected and is not published on a gold open access basis under a CC BY licence, unless that is specifically stated in the figure caption in the Version of Record.

View the article online for updates and enhancements.

This content was downloaded from IP address 14.139.251.98 on 06/02/2023 at 05:18

Effect of fiber synergy on strength & durability of High-Performance Concrete.

Received xxxxxx
Accepted for publication xxxxxx
Published xxxxxx

Mohit Gupta [1], Dr. Ritu Raj [2], Dr. Anil Kumar Sahu [3], Dr. Sapna Yadav[4*]

[1] Research Scholar, Delhi Technological University, Shahbad Daulatpur, Main Bawana Road, Delhi-110042. India (mohit_2k17phdcept029@dtu.ac.in)

[2] Assistant Professor, Delhi Technological University, Shahbad Daulatpur, Main Bawana Road, Delhi-110042. India (rituraj@dtu.ac.in)

[3] Professor, Delhi Technological University, Shahbad Daulatpur, Main Bawana Road, Delhi-110042. India (sahuanilkr@yahoo.co.in)

[4*] Assistant Professor, KIET Group of Institutions, Delhi- NCR, Ghaziabad, India (sapna.yadav@kiet.edu)

Correspondence should be addressed to Dr. Sapna Yadav; sapna.yadav@kiet.edu

Abstract

The distinct fiber addition to concrete can improve weak performance in tension. Due to its higher performance, fiber hybridization is becoming more and more competitive. In this study, the hardened and durability characteristics of performance-based concrete reinforced with polypropylene, basalt and hooked end steel fibers are investigated. Total 12 mixes are prepared to evaluate the different characteristics of performance-based concrete. The influence of basalt fibers (0.2, 0.25, 0.5, 0.75 & 1 %), polypropylene fibers (0.025, 0.03, 0.035, 0.05 & 1%), hooked end steel fibers (0.225, 0.475, 0.65, 0.75, 0.80 & 1%) to explore the impact of fiber synergy, various hybrid fibers with maximum volume fraction of 1% were used out of total concrete volume with and without copper slag (20% & 40%) replacement with fine aggregate are investigated on the hardened characteristics including; compressive strength, flexural strength, splitting tensile strength. Moreover, durability properties including; water absorption, electrical Resistivity, for performance-based concrete are also conducted. Out of the 12 mixes considered for this study, PP0.05B0.2S0.75CS40 performs better and exhibit excellent performance in terms of strength and durability.

Keywords: *Performance based concrete, fiber synergy, mechanical property, durability.*

1. Introduction

Construction projects frequently involve the usage of concrete. Further improvements to concrete's durability, flexibility, impact resistance, and energy absorption are advised given the current building standards [1,2]. To make concrete stronger and more durable, fibers are added to the mix. Its key benefits include preventing macro crack formation and micro crack expansion to microscopic and macroscopic levels. Low tensile strength and tensile strain capacity are combined with significant brittleness in concrete. Every day, there is a greater requirement for concrete to be more resilient and long-lasting in challenging situations. The material qualities and fiber geometry of a specific type of fiber will determine how well it can accomplish this task. A technique for maximizing the enhancements provided by fiber reinforcement is to incorporate two or more types of fiber that contribute in complementary and additive ways to performance in a concrete mix [3,4,5,6]. The greatest difficulty resulting from the industrial revolution is environmental conservation. More energy and money are spent on waste disposal than on waste generation. Additionally, copper slag created by copper companies has grain size comparable to sand particles and doesn't need any additional alterations.

Large amounts of Portland cement, silica fume (SF), and small aggregates with steel fibers for reinforcement are typical characteristics of Ultra high- performance concrete (UHPC) [11,12,13]. UHPC also exhibits exceptional flowability and a close-packed condition while having a very less w/b ratio. By adding fiber, cracks can be bridged and prevented from spreading, which considerably

Accepted Manuscript

lessens the characteristics of dry shrinkage. When fibers are added, the stress-strain relationship of cured concrete performs better [14]. To improve performance and reduce the life cycle costs of concrete structures, engineers must continue to face a formidable challenge: obtaining acceptable concrete durability levels. The advantages of utilizing additional binder components for concrete durability are well recognized [15,16,17].

One of the main problems is that regular concrete doesn't vibrate enough, which seriously impairs mechanical performance. In order to complete intricate formwork and create high-quality buildings, new cementitious composites are required, as shown by this occurrence [18,19,20]. Achieving the correct level of fiber efficiency can increase composite toughness, which is important for energy absorption, crack prevention, and durability [21]. Explosive spalling in high strength concrete (HSC) was reduced by adding 0.5 to 3 kg/m3 of PP fibers [22,23,24,25,26].

According to research by Khaloo et al. [27] on the mechanical and rheological characteristics of steel fiber reinforced concrete (SFRC), the addition of silica fume decreased the compressive strengths and workability of SCC. Otherwise, it might improve the energy absorption capability, rupture modulus, and splitting tensile strength. Afroughsabet and Ozbakkaloglu [28] revealed that the inclusion of steel and polypropylene fibers increased the concrete's compressive, splitting tensile, and flexural strengths. The electrical resistivity did, however, drop when steel fibers were present, as opposed to rising when polypropylene fibers were included.

There is ample scope to work with different fiber hybridization such as, PP, B & HES fibers in different combinations on Performance based concrete. Incorporation of Copper slag in Fiber reinforced performance-based concrete is limited to a large extent. Lack of Mechanical and Durability properties with PP, B & HES fibers with substitution of copper slag was seen throughout the literature review. No work was reflected from literature in terms of PP, B & hooked end steel fiber combination for High performance concrete with copper slag. As the work is only limited to other different fibers without substitution of Fine aggregate. Therefore, there is ample scope to introduce all these fiber in combination (PP-B-HES) in concrete and it is required to investigate further important properties for the development of Performance based Concrete of M65 which can be used in actual practice. Based on the fiber constitutive response, fiber dimensions, fiber functions and fiber performances hooked end steel fiber, polypropylene fiber and basalt fiber are selected for this research. Silica fume, fly ash are the essential constituents for making high strength concrete and to prevent depletion of fine aggregate, other waste material such as copper slag is used in this research to make the environment ecological sustainable [31-37].

Ye Li et al. (2019) investigated PP fiber at a rate of 4 kg/m3 addition inhibited explosive spalling of UHPC, Steel fiber (SF) incorporation alone offered no protection against it. Due to a large increase in permeability, mixed PP and SF fibers fully avoided explosive spalling even at low fiber dose. Spalling-induced weight loss decreases proportionately as UHPC permeability increases.

According to Estela O. Garcez et al. (2019), all of the mixtures had an ideal flow of 560 mm in 4 seconds. In terms of elastic modulus and compressive strength during a 28-day period, (Forta ferro- low density) FF-LD had the highest values. Among all the combinations, (Forta ferro- medium density) FF-MD fiber displayed the highest splitting tensile strength, measuring 5 MPa. The fact that residual flexural strength improved dramatically from 36% to 79% is undoubtedly evidence of the contribution of fibers to concrete's post-cracking behavior. The mix design with the lowest Forta Ferro and RheoShore fiber dosages appears to be the most appropriate for the intended application of the four.

In their research, Emad A.H. Alwesabi et al. (2019) shown that hybrid fibers (0.1% PP + 0.9% MS) had greater elastic modulus, CS, STS, and other mechanical properties. The FS of concrete was greatly increased by the mixes containing 1.0% MS. The mechanical characteristics of concrete were negatively impacted by the 1% PP mixes.

According to research by Bilal Masood et al. (2019), the workability of recycled aggregate concrete (RAC) is higher than that of the natural aggregate concrete (NAC) because RAC contains more free water. Comparatively speaking, RAC mixes have more air content than NAC mixes did. The fresh density of concrete mixes was decreased by substituting NCA for RCA. Comparing RAC to its NAC mix competitors, the compressive strength of RAC dropped.

PP and BF both attain the best mechanical performance at a content of 0.1%, according to research by Xinyu Hu and Yihong Guo (2019). The workability of concrete is not significantly affected by PP and BF hybridization when compared to single fibers at the same concentration [38-42].

Dehong Wang et al. (2019) discovered that the highest synergy effect of hybrid fibres occurred when the volume percentage of basalt fibre was 0.15% and the volume fraction of polypropylene fibre was 0.033%.

Iman Sadrinejad et al. (2018) investigated how hybridizing poly-olefin (PO) with PP fibers might increase mixtures' compressive and splitting tensile strengths by up to 7.5% and 23%, respectively, in comparison to control mixtures. These fibers' hybridization did not have a beneficial impact on the post-cracking behavior during the flexural test. large volume fiber content, the detrimental effects of PO and PP fiber hybridization were evident. For

Accepted Manuscript

actual use, it is advised to use PO and PP fibers at volume fractions of 0.9% and 0.1%, respectively.

In their study, Doo-Yeol Yoo et al. (2017) discovered that the flexural performance suffered when short fibres length of fiber to diameter of fiber (l_f/d_f of 13/0.2) were used in place of long fibres (l_f/d_f of 30/0.3). Long fibers were used, which significantly enhanced complimentary energy.

2. Experimental program

2.1 Materials used in mix design

As per ACI Committee 363R -92, the main raw ingredients used to develop the concrete mix design for HSC were cement, fly ash, silica fume, coarse aggregates, fine aggregates, Polypropylene fibers, basalt fibers, hooked end steel fibers, polycarboxylate ether (PCE) based superplasticizer, and water.

2.1.1 Cement

OPC of 53 grade cement has a minimum compressive strength of 53 MPa. The cement met the requirements of IS 8112-2013. Table 2 displays the chemical make-up of the 53-grade cement used in OPC. As shown in Table 1, all required cement tests have been completed before casting and compared to OPC 53 grade standards.

2.1.2 Admixtures

A pozzolan substance called fly ash admixture was acquired from Trans Delhi Signature City, Ghaziabad. Fly ash utilized in the mix design was Grade 1 with requirements that complied with IS 3812 codes. According to IS 456-2000, FA is typically used as a partial substitution for cement. According to Table 2, this grade of fly ash has been identified as having the following chemical makeup. The widely used mineral additive silica fume, according to IS 5388: 2003, was employed in the mix design to create high-strength concrete. According to the specifications supplied by ASTM C-1240, AASHTO M307, and IS Code, the desired results were obtained for the silica fume admixture as indicated in Table 3. An excellent water content reduction, slump retention admixture, and poly carboxylate ether (PCE) chemical admixture were utilized. According to marsh cone test, the ideal concentration of the superplasticizer (SP) was 1.12%. The parameters were in accordance with IS 9103: 1999, and Table 4 lists the superplasticizer's physical characteristics.

1.1 Problem Statement

Concrete faults become more noticeable as the service life lengthens, which causes numerous issues in engineering applications. As concrete ages, local failure mechanisms that regulate its ability to bridge are reduced bonding capacity, fibre pull-out, and fibre sliding. The linear elastic fracture mechanics form the foundation of the fibre spacing theory. It is believed that the concrete has flaws and microscopic fissures. The crack enlarges as a result of the stress concentration created under the influence of the external force. By using fibre, the crack can be successfully stopped from spreading. This study represents the utilization of hybrid fibres which greatly reduces the effect due to shrinkage in construction practices.

2.1.3 Aggregates

A neighboring site in New Delhi, India provided Badarpur sand (fine dust) of zone II, which was used as fine aggregates in accordance with IS 383: 1970. The tests' fine dust was produced to fit through a sieve with a 4.75mm opening. Based on the results of the sieve analysis test, the fineness modulus of sand utilized in the experiment was determined to be 2.80. The values obtained using the pycnometer approach, which took into account water absorption and the specific gravity of the fine dust, were 1.2% and 2.60, respectively. The fine aggregate (FA) was subjected to sieve analysis to produce the particle size distribution curve depicted in Figure 1. For the design mix, coarse aggregates (CA) with a nominal size of 10 mm were employed, and their fineness modulus was once more assessed using a sieve analysis test. It was determined to be 7.2 The results obtained using the wire basket method were 0.2% and 2.8, respectively, for water absorption and the specific gravity of the coarse aggregates.

2.1.4 PP, Basalt and Hooked End Steel fibers

Polypropylene fiber, Basalt Fiber and Hooked End Steel fibers were taken from Dilshad Composites, Ghaziabad and is utilized throughout the research work. Hooked end steel fiber is having 60mm length and 0.75mm diameter, polypropylene fiber is having 12 mm length and 0.022mm diameter, basalt fiber is having 12 mm length and 12 micrometer diameter.

2.2 Mix Design Proportion of Concrete

To comprehend the mechanical behavior of high-strength concrete and the viability of hybrid fibers, twelve mix proportion compositions were

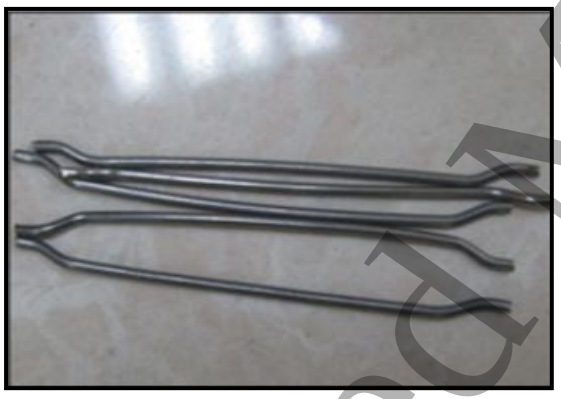
Accepted Manuscript

established. Table 6 displays the numerous mixed design compositions that were created. Basalt fiber, polypropylene fiber, and hooked end steel fiber were used as hybrid fibers.

Table 2 focusses on various mix proportion. PP0B0C0CS0 shows the control mix. Mix PP0B0S0CS20 and PP0B0S0CS40 represents 0% Polypropylene, 0% Basalt, 0% Hooked End Steel fibers and Copper Slag of 20% and 40% respectively. Similarly, PP0.025B0.5S0.475CS40 represents 0.025% Polypropylene, 0.5% Basalt, 0.475% Hooked End Steel fibers and Copper Slag of 40%. In a similar manner, all the remaining mix designations were nomenclature.

Initially, CA and FA were introduced to the concrete mixer for batching, and the mixer received automated revolutions for roughly a minute. At this point, mineral additive silica fume, fly ash, and cement were also added. The determined amount of water and SP was added, and the mixture was properly homogenized before being slowly added to the mixer. Finally, the appropriate dose of hybrid fibers was then equally distributed throughout the dry mixer, and the mixer once more received automatic revolutions for another minute. For around three minutes, all the mix components received enough revolutions to produce a satisfactory blend.

The mixing process was now halted, and the prepared concrete mixture was poured into molds with the dimensions 150*150*150mm and 500*100*100mm, respectively, for the manufacture of cubic and prismatic beam samples for subsequently conducting compressive and flexure tests. Using a shaking table to create a dense mixture for generating a high strength concrete mix, the prepared concrete mix was poured in three layers to achieve enough compaction. The set samples were also demolded after 24 hours, and they were then cured in water for 28 days. The samples were ready for testing for ultrasonic pulse velocity tests after 28 days of water curing, prior to flexure and compressive strength tests in accordance with IS 516-1959 and IS 456-2000, these tests were conducted.



(a)

(b)

(c)

Fig.1 (a) HESF, (b) PP Fibers (c) Basalt Fibers

Accepted Manuscript

Table 1. Properties of Hybrid fibers used

Properties of Fibers	Polypropylene Fiber	Basalt Fiber	Hooked End Steel Fiber
Density (g/cc)	0.91	2.65	7.85
Tensile Strength (MPa)	310-540	2800-3100	1140
Elastic Modulus (GPa)	3.72	85-87	210
Melting Point (ºC)	160-170	-200 to +900(Operative Temp.)	400-500
Failure Strain (%)	19-20	3.1	3-4

Table 2. Concrete mix design for HFRC (kg m^{-3})

S.No.	Mix Designation	Cement	Fine Aggregate	Coarse aggregate	Water	Chemical admixture	Fly ash	SF	PP (Vf%)	B (Vf%)	HES (Vf%)	CS
1	Control Mix PP0B0C0CS0	441	578	1169	142	1.12	105	43	-	-	-	-
2	PP0B0S0CS20	441	578	1169	142	1.12	105	43	-	-	-	20%
3	PP0B0S0CS40	441	578	1169	142	1.12	105	43	-	-	-	40%
4	PP0.025B0.5S0.475CS40	441	578	1169	142	1.12	105	43	0.025	0.5	0.475	40%
5	PP0.025B0.25S0.725CS40	441	578	1169	142	1.12	105	43	0.025	0.25	0.725	40%
6	PP0.025B0.75S0.225CS40	441	578	1169	142	1.12	105	43	0.025	0.75	0.225	40%
7	PP0.05B0.2S0.75CS40	441	578	1169	142	1.12	105	43	0.05	0.2	0.75	40%
8	PP0.15B0.2S0.65CS40	441	578	1169	142	1.12	105	43	0.15	0.2	0.65	40%
9	PP0.1B0.2S0.7CS40	441	578	1169	142	1.12	105	43	0.1	0.2	0.7	40%
10	PP0.01B0.24S0.75CS40	441	578	1169	142	1.12	105	43	0.01	0.24	0.75	40%
11	PP0.03B0.22S0.75CS40	441	578	1169	142	1.12	105	43	0.03	0.22	0.75	40%
12	PP0.035B0.215S0.75CS40	441	578	1169	142	1.12	105	43	0.035	0.215	0.75	40%

3 Experimental Program

3.1 Compressive Strength

To determine the CS of various mix compositions, 36 cubes as samples with dimensions of 150* 150* 150 mm casted, cured, and tested at 28 days in accordance with the IS Code [IS-456:2000]. Standard equipment with a capacity of 2000 kN for strength estimation was used. Loading rate was 5.18 kN s^{-1} for the samples. 3 cubes average strength of every mix was considered as compressive strength.

3.2 Flexural Strength

36 beam samples having 150*150*700 mm in size, with 12 various mix proportions, were cast and subjected to flexure testing. According to ASTM C1609 requirements, four-point loading was used to calculate flexural strength incorporated hybrid fibers in the specimen. A load of 0.035 to 0.1 mm per minute was applied to the specimen. Under displacement control, a 0.6 mm min^{-1} deflection rate was performed. This test method and residual loads that take into consideration the L/600 and L/150 deflection were used to calculate the peak load (P) because aggregate can only be 20 mm in size, the size of the beam specimen that was taken is 150 mm x 150 mm x 700 mm. This flexural strength test was conducted in UTM & a loading rate of 400 kg/min till failure. The modulus of rupture used in Equation 2 to calculate strength.

$$f = PL/Bd\text{^}2 \quad (1)$$

Where "L" stands for the beam's length, "d" for its depth, "P" for its failure load, and "f" for its flexural strength.

3.3 Split Tensile Strength

Three cylinders for each type of mix were subjected to STS testing in accordance with ASTM C39 specifications. Every minute, 12 MPa of loading

Accepted Manuscript

occurs. For testing, 36 cylinders of hybrid fibers in varied proportions were put to each sample. The specimen is 150mm x 300mm in size. From 0.7 to 1.4 MPa/min of load was applied to the ST test until failure. STS Determined using the formula provided in the Indian standard, in accordance with ASTM C496 standards.

$$fst = 2P/\pi DL \quad (2)$$

In this formula, L denotes the cylinder's length, D its diameter, P its failure load, and fst its split tensile strength.

3.4 Specific Electrical Resistivity and Water absorption.

With an AC-Impedance spectroscopic technique, the particular electrical resistivity was determined using a 1.0 kHz frequency and 1.0 Ω-m ultimate capacity. According to BS 1881-72, cubic specimens were subjected to water absorption tests. In water, samples were cured for 27 days at 22 degrees centigrade. After that, the samples were dried for at least 14 days in an oven that was set at 45 degrees Celsius. If the values obtained from two successive measures of mass differed by more than 0.5% of the smaller value, the specimens were put back in the oven for an additional 24-hour drying period. The weights of the samples were then measured over a longer period of time until there was a 0.5% variation between any two successive measurements. The cured specimens were then wiped with a dry paper towel, submerged in tank filled with water for half an hour and 168 hours, and weighed on a 0.01 g balance. The early and final water absorptions are referred to in the current study as occurring at 0.5 hours and 7 days, respectively. The remaining tests were conducted at 7, 28, and 91 days.

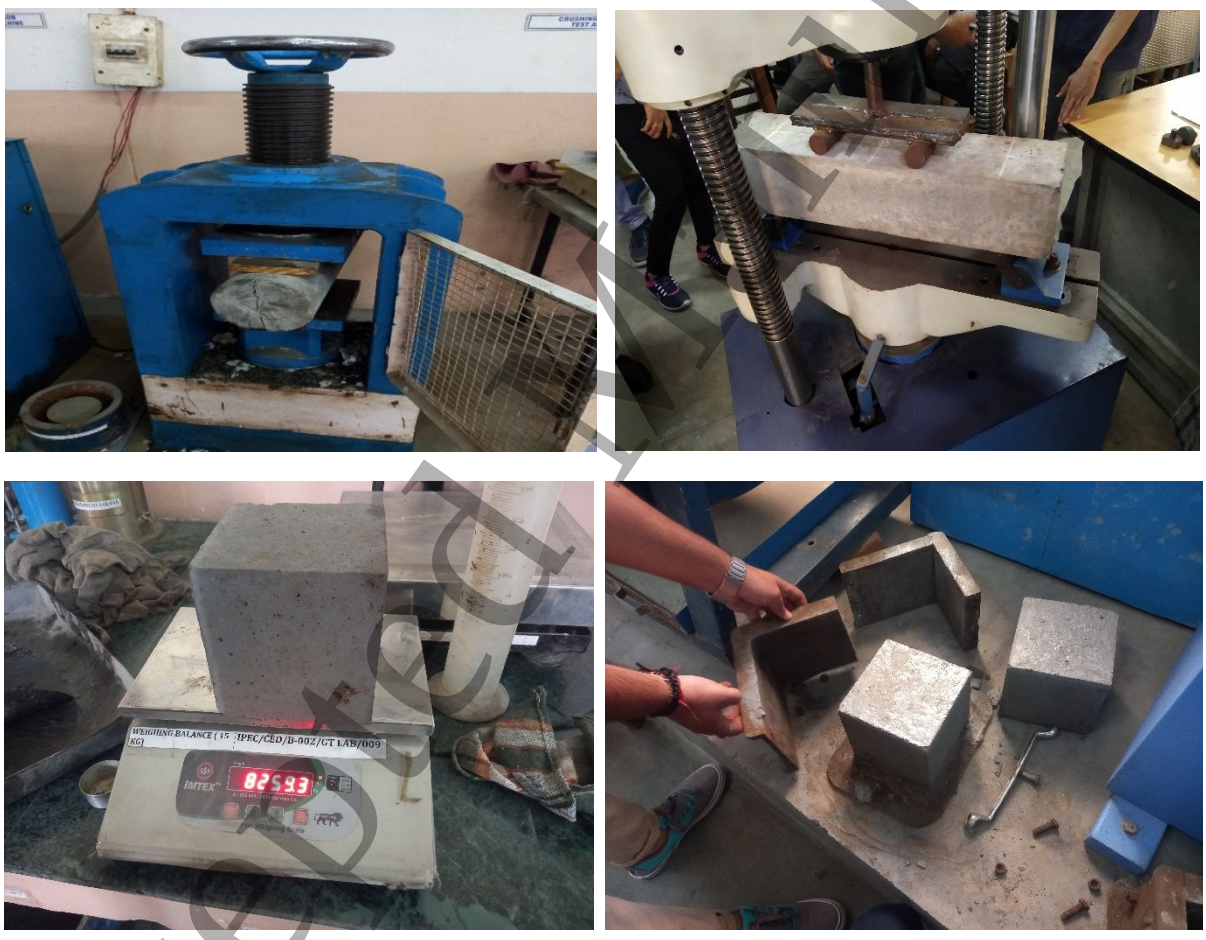


Fig. 2 Casted specimens for testing of compressive strength, flexural strength and split tensile strength

4 Results and Discussion

4.1 Effect of hybrid fibers on compressive strength of concrete

As, it is very clear from figure 3 and table 3 that compressive strength of hybrid fibers of 12 mixes have different strength at 7 days and 28 days of curing. The findings after minute monitoring of tested cubes in CTM reveals that control mix PP0B0S0CS0 has compressive strength of 76.5 MPa. In addition to this on increasing the dosage of hybrid fibers the behavior of tested samples changed and maximum compressive strength is reported for mix PP0.05B0.2S0.75CS40 i.e., 78 MPa which is 1.96% greater than control mix PP0B0S0CS0. After that, for all the mixes strength starts decreasing with a similar trend. The mix PP0.035B0.215S0.75CS40 shows lowest compressive strength out of 12 mixes

Accepted Manuscript

prepared. Because the failure is in the compressive axis and not due to fracture, the concentration of fibers in compression does not significantly increase. Although compressive strength in terms of capacity to carry loads does not satisfy the criteria, the increased fiber dosage in cubes indicates that the fracture width has decreased. Observed comparable results on varied fiber percentages and stated that steel, polypropylene fiber and basalt fiber influence on compressive strength is quite minimal. This might be caused by an increase in the amount of fiber in the mixture, which creates voids and lowers the Intermediate transition zone (ITZ) between the fibers and other elements. Table 3 also demonstrates that, depending on the replacement amount of steel, polypropylene fiber and basalt fiber PP0.05B0.2S0.75CS40 mix shows maximum strength. The mix of 0.05% PP, 0.2 % Basalt fiber and 0.75% steel fibers produced the maximum compressive strength out of all the steel, PP fiber and basalt fiber combinations taken into consideration in this investigation.

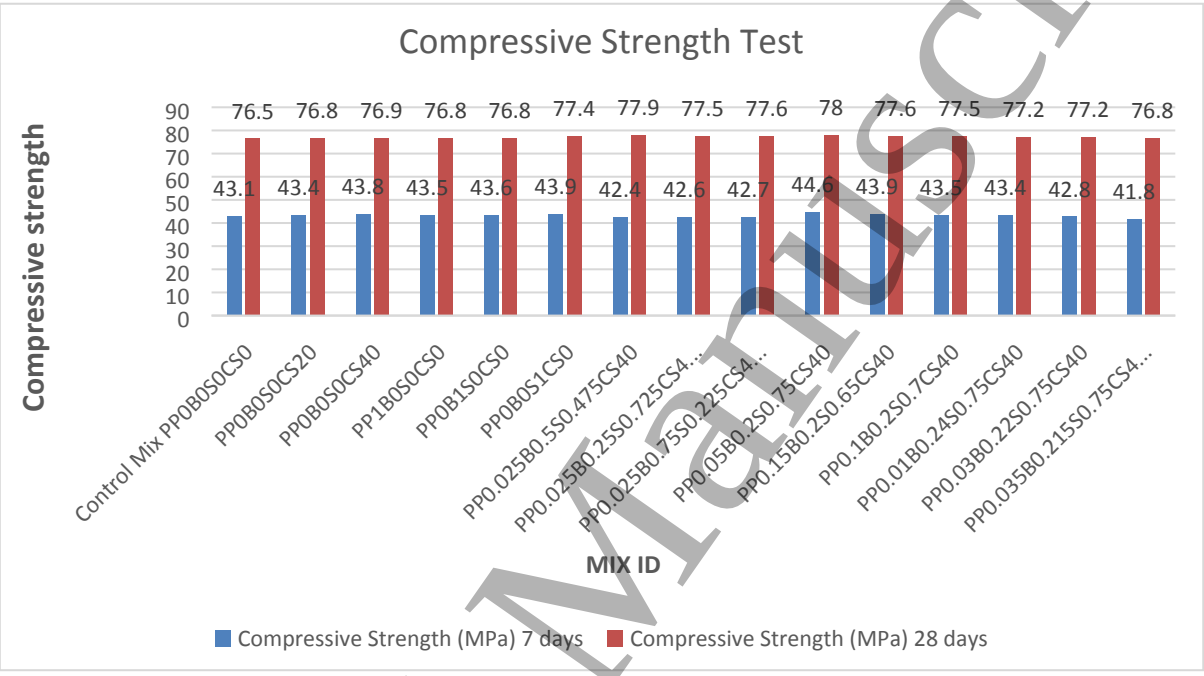


Figure 3 Comparison of compressive strength of hybrid fiber HPC with controlled concrete.

Table 3 Compressive strength of hybrid fiber HPC.

Mix Designation	Compressive Strength (MPa)	
	7 days	28 days
Control Mix PP0B0S0CS0	43.1	76.5
PP0B0S0CS20	43.4	76.8
PP0B0S0CS40	43.8	76.9
PP1B0S0CS0	43.5	76.8
PP0B1S0CS0	43.6	76.8
PP0B0S1CS0	43.9	77.4
PP0.025B0.5S0.475CS40	42.4	77.9
PP0.025B0.25S0.725CS40	42.6	77.5
PP0.025B0.75S0.225CS40	42.7	77.6
PP0.05B0.2S0.75CS40	44.6	78
PP0.15B0.2S0.65CS40	43.9	77.6
PP0.1B0.2S0.7CS40	43.5	77.5
PP0.01B0.24S0.75CS40	43.4	77.2
PP0.03B0.22S0.75CS40	42.8	77.2
PP0.035B0.215S0.75CS40	41.8	76.8

PP0B0S0CS0 signifies 0% polypropylene fiber, 0% basalt fiber, 0% hooked end steel fiber, 0% copper slag and same nomenclature with different percentage of all other mixes with fibers are addressed.

Accepted Manuscript

4.2 Effect of hybrid fibers on flexural strength of concrete

As, it is very clear from figure 4 and table 4 that flexural strength of hybrid fibers of 12 mixes has different strength at 28 days of curing. The findings after minute monitoring of tested beams in UTM reveals that control mix PP0B0S0CS0 has flexural strength of 5.96 MPa. In addition to this on increasing the dosage of hybrid fibers the behavior of tested beam samples changed and maximum flexural strength is reported for mix PP0.05B0.2S0.75CS40 i.e., 10.3 MPa which is 72.81 % greater than control mix PP0B0S0CS0. After that, for all the mixes strength starts decreasing with a similar trend. The mix PP0.035B0.215S0.75CS40 shows lowest flexural strength out of 12 mixes prepared. The improvement in strength brought on by the addition of supplementary cementitious materials including silica fume and is due to the stronger connection between the wet cement matrix and the aggregate. Due to the conversion of calcium hydroxide, which typically occurs on the surface of aggregate particles, into CSH in the presence of reactive silica, the connection is reinforced. The reason of increased flexural strength incorporating hybrid fibers is the densest inter- particle packing in the concrete matrix and strong bond. Table 4 also demonstrates that, depending on the replacement amount of steel, polypropylene fiber and basalt fiber PP0.05B0.2S0.75CS40 mix shows maximum strength. The mix of 0.05% PP, 0.2 % Basalt fiber and 0.75% steel fibers produced the maximum flexural strength at 28 days of curing out of all the steel, PP fiber and basalt fiber combinations taken into consideration in this investigation. As a result, hybrid fiber-reinforced concretes lost some of their flexural strength when PP fibers were 0.15%, basalt fiber was 0.2 % and steel fiber was 0.65% and starts decreasing after all the mixes.

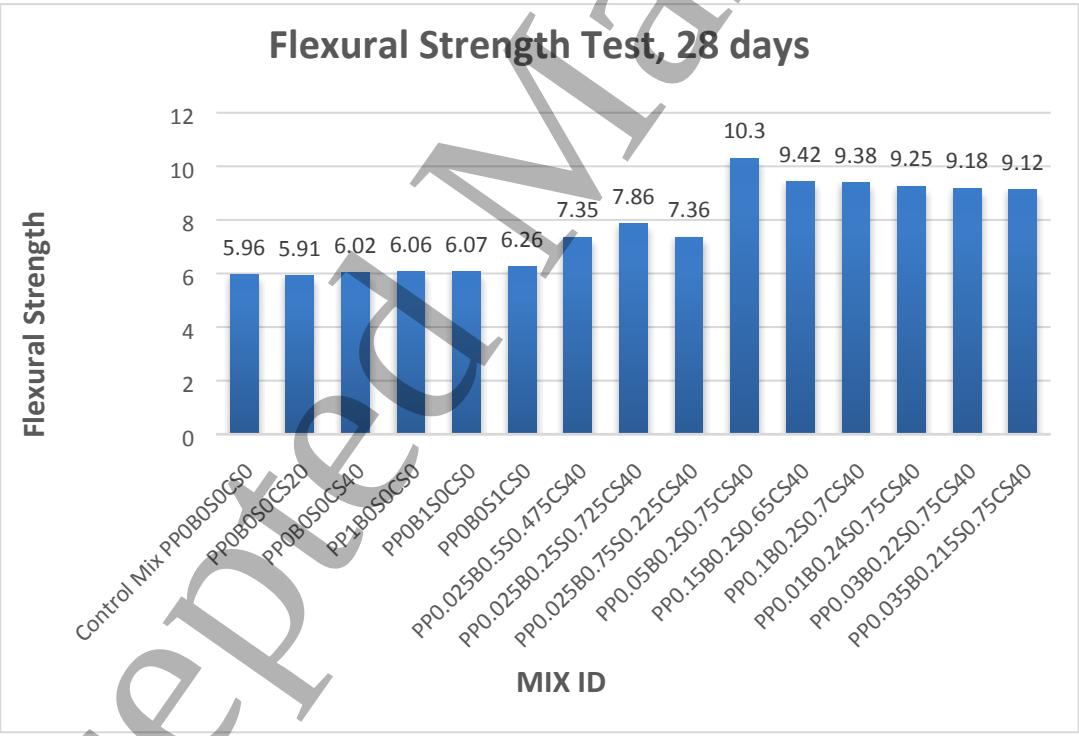


Figure 4 Comparison of flexural strength of hybrid fiber HPC with controlled concrete.

Mix Designation	Flexural Strength 28 days
Control Mix PP0B0S0CS0	5.96
PP0B0S0CS20	5.91
PP0B0S0CS40	6.02
PP1B0S0CS0	6.06
PP0B1S0CS0	6.07
PP0B0S1CS0	6.26
PP0.025B0.5S0.475CS40	7.35
PP0.025B0.25S0.725CS40	7.86

Accepted Manuscript

PP0.025B0.75S0.225CS40	7.36
PP0.05B0.2S0.75CS40	10.3
PP0.15B0.2S0.65CS40	9.42
PP0.1B0.2S0.7CS40	9.38
PP0.01B0.24S0.75CS40	9.25
PP0.03B0.22S0.75CS40	9.18
PP0.035B0.215S0.75CS40	9.12

PP0B0S0CS0 signifies 0% polypropylene fiber, 0% basalt fiber, 0% hooked end steel fiber, 0% copper slag and same nomenclature with different percentage of all other mixes with fibers are addressed.

4.3 Effect of hybrid fibers on Split tensile strength of concrete

As, it is very clear from figure 5 and table 5 that split tensile strength of hybrid fibers of 12 mixes has different strength at 28 days of curing. The findings after minute monitoring of tested beams in CTM reveals that control mix PP0B0S0CS0 has split tensile strength of 5.04 MPa. In addition to this on increasing the dosage of hybrid fibers the behavior of tested cylinder samples changed and maximum split tensile strength is reported for mix PP0.05B0.2S0.75CS40 i.e., 5.47 MPa which is 8.53 % greater than control mix PP0B0S0CS0. After that, for all the mixes strength starts decreasing with a similar trend. The mix PP0.035B0.215S0.75CS40 shows lowest split tensile strength of 4.98 MPa out of 12 mixes prepared. The other researchers also support the similar trend in their research.

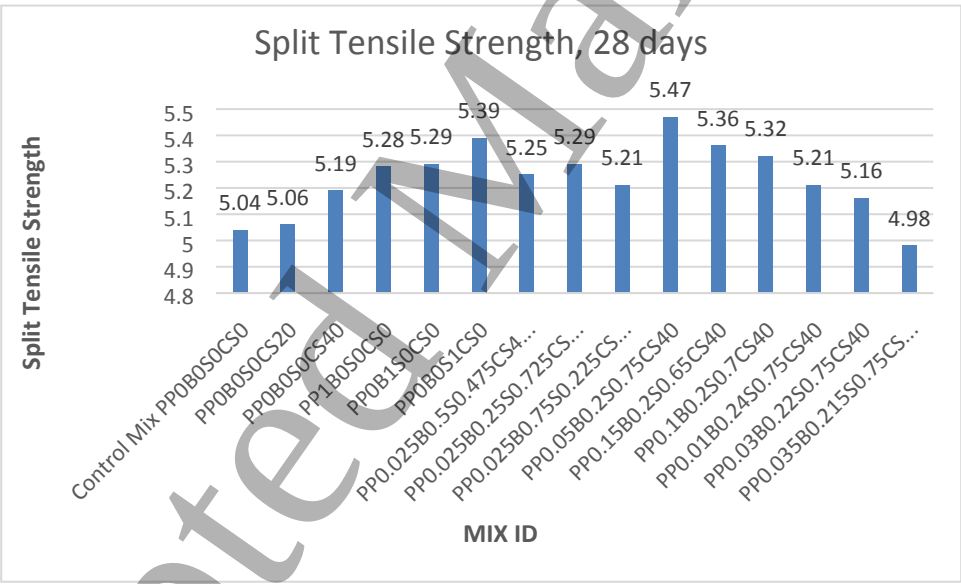


Figure 5 Split tensile strength of hybrid fiber reinforced high performance concrete with controlled concrete.

Table 5 Split Tensile strength of hybrid fiber reinforced high performance concrete.

Mix Designation	Split Tensile Strength 28 days
Control Mix PP0B0S0CS0	5.04
PP0B0S0CS20	5.06
PP0B0S0CS40	5.19
PP1B0S0CS0	5.28
PP0B1S0CS0	5.29
PP0B0S1CS0	5.39
PP0.025B0.5S0.475CS40	5.25
PP0.025B0.25S0.725CS40	5.29
PP0.025B0.75S0.225CS40	5.21
PP0.05B0.2S0.75CS40	5.47

Accepted Manuscript

PP0.15B0.2S0.65CS40	5.36
PP0.1B0.2S0.7CS40	5.32
PP0.01B0.24S0.75CS40	5.21
PP0.03B0.22S0.75CS40	5.16
PP0.035B0.215S0.75CS40	4.98

PP0B0S0CS0 signifies 0% polypropylene fiber, 0% basalt fiber, 0% hooked end steel fiber, 0% copper slag and same nomenclature with different percentage of all other mixes with fibers are addressed.

4.4 Comparison of optimum dosage of hybrid fibers with controlled concrete

Table 6 Optimum dosage of hybrid fiber reinforced high performance concrete.

S.No.	Mechanical Property	Strength (MPa) of Controlled Mix	Strength (MPa) of PP0.05B0.2S0.75CS40 at 28 days	% Increase with Controlled Concrete
1	Compressive Strength	76.5	78	1.96
2	Flexural Strength	5.96	10.3	72.81
3	Split tensile strength	5.04	5.47	8.53

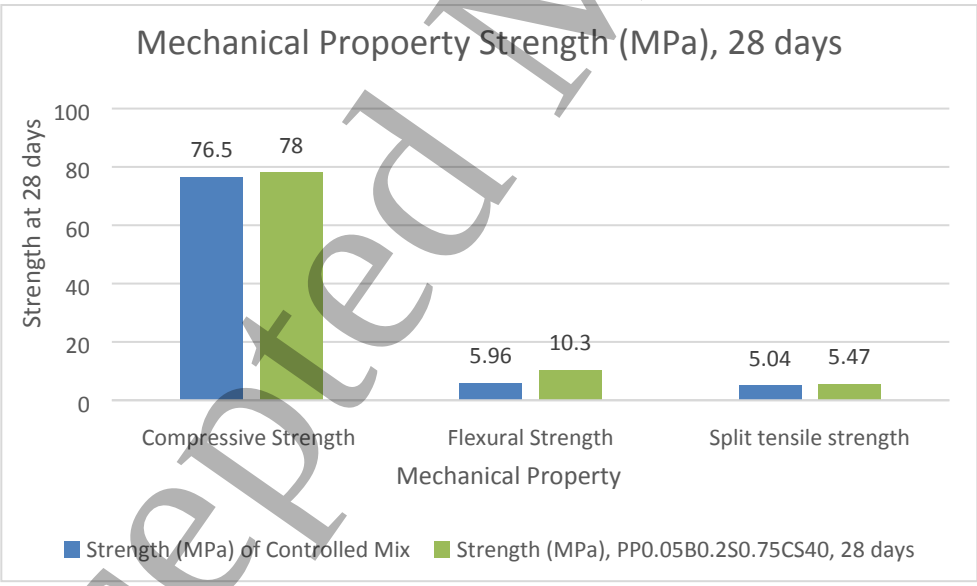


Figure 6 Optimum dosage of hybrid fiber reinforced high performance concrete and comparison with controlled concrete.

4.5 Effect of water absorption on hybrid fibers.

Another significant factor that affects concrete longevity is its ability to resist the introduction of hazardous ions. The absorption characteristics of concrete provide crucial information about the quantity of permeable pores inside the concrete and the connections between those pores in addition to acting as an indirect indicator of its porosity. Figure 7 & Table 7 provide a visual representation of the water absorption of the numerous hybrid FRC mixes used in the current research. According to the findings, PP0.05B0.2S0.75CS40 showing 0.41% water absorption at 0.5 hours (early) and 0.61% water absorption at 168 hours (ultimate) i.e., 7 days and significantly reduce concrete's capacity to absorb water than controlled concrete mix PP0B0C0CS0 which is having 1.53 % (ultimate) water absorption.

Accepted Manuscript

As a result, various mixes examined in this research displayed water absorption value low, which suggested "excellent" concrete quality, in accordance with these classifications. Those with poor water absorption results are a result of the concrete mixtures' low porosity and restricted pore connectivity.

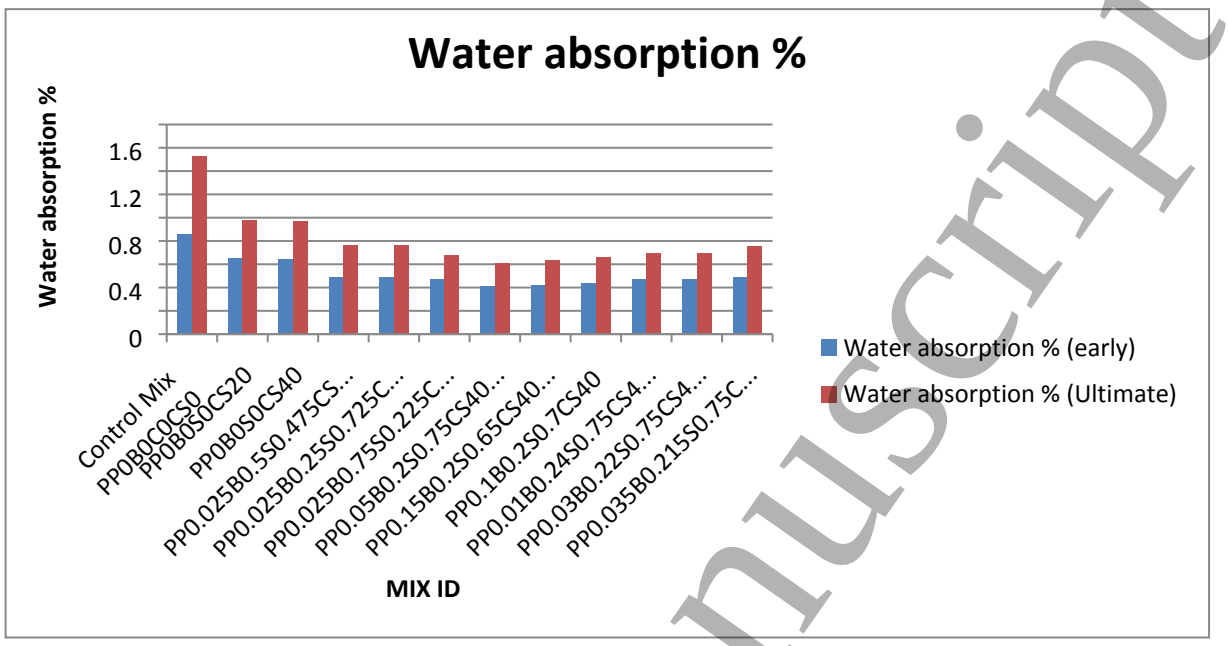


Figure 7 Early and Ultimate water absorption of hybrid fiber reinforced high performance concrete.

Table 7 Early and Ultimate water absorption of hybrid fiber reinforced high performance concrete.

Mix No.	Mix ID	Water absorption % (early)	Water absorption % (Ultimate)
1	Control Mix PP0B0C0CS0	0.86	1.53
2	PP0B0S0CS20	0.65	0.98
3	PP0B0S0CS40	0.64	0.97
4	PP0.025B0.5S0.475CS40	0.49	0.76
5	PP0.025B0.25S0.725CS40	0.49	0.76
6	PP0.025B0.75S0.225CS40	0.47	0.68
7	PP0.05B0.2S0.75CS40	0.41	0.61
8	PP0.15B0.2S0.65CS40	0.42	0.63
9	PP0.1B0.2S0.7CS40	0.44	0.66
10	PP0.01B0.24S0.75CS40	0.47	0.69
11	PP0.03B0.22S0.75CS40	0.47	0.69
12	PP0.035B0.215S0.75CS40	0.49	0.75

PP0B0S0CS0 signifies 0% polypropylene fiber, 0% basalt fiber, 0% hooked end steel fiber, 0% copper slag and same nomenclature with different percentage of fibers are addressed.

4.6 Effect of hybrid fibers on Specific Electrical Resistivity.

Given that corrosion in reinforced concrete is significantly influenced by electrical resistivity, it is one of the most crucial characteristics of concrete durability. It has been demonstrated in the past that corrosion of concrete reinforcement is not likely to occur above an electrical resistance of 120 Ω-m, which is the limit for corrosion propagation in internal steel reinforcing bars. The microstructure of concrete becomes denser due to the presence of supplementary cementitious materials and hybrid

Accepted Manuscript

fibers, and secondary calcium silicate hydrate is created as a result of its pozzolanic reaction (C–S–H). The C-S-H gel, which is a well-known source of concrete strength, increases the volume of solid phases and inhibits the development of capillary pore systems. Concrete resistivity, one of the endurance characteristics of concrete, is improved as a result of these events. Electrolytes in the pores of composite materials and the presence of conductive materials 112.4 (Ω m) at 28 days and 215.6 (Ω m) at 91 days. Further, specific electrical resistivity of PP0.035B0.215S0.75CS40 is 217.2 (Ω m) at 91 days. Varying types of fibers added to concrete have different effects on how electrically resistive the concrete is. Polypropylene fiber and basalt fiber like steel fibers in fiber-reinforced specimens both have a significant impact on the electrical resistance of concrete. The findings show that the electrical resistance of concrete was greatly decreased by addition of hybrid fibers. From table 8 it is very clear that Control Mix PP0B0C0CS0 specific electrical resistivity (Ω m) was 67.8 at 28 days while 104.4 at 91 days. Moreover, PP0.05B0.2S0.75CS40 shows

addition makes mixes more porous and slightly lowers the electrical resistance of concrete. The electrical resistivity of concrete is greatly reduced when all hybrid fibers are added because of the fibers' high conductivity.

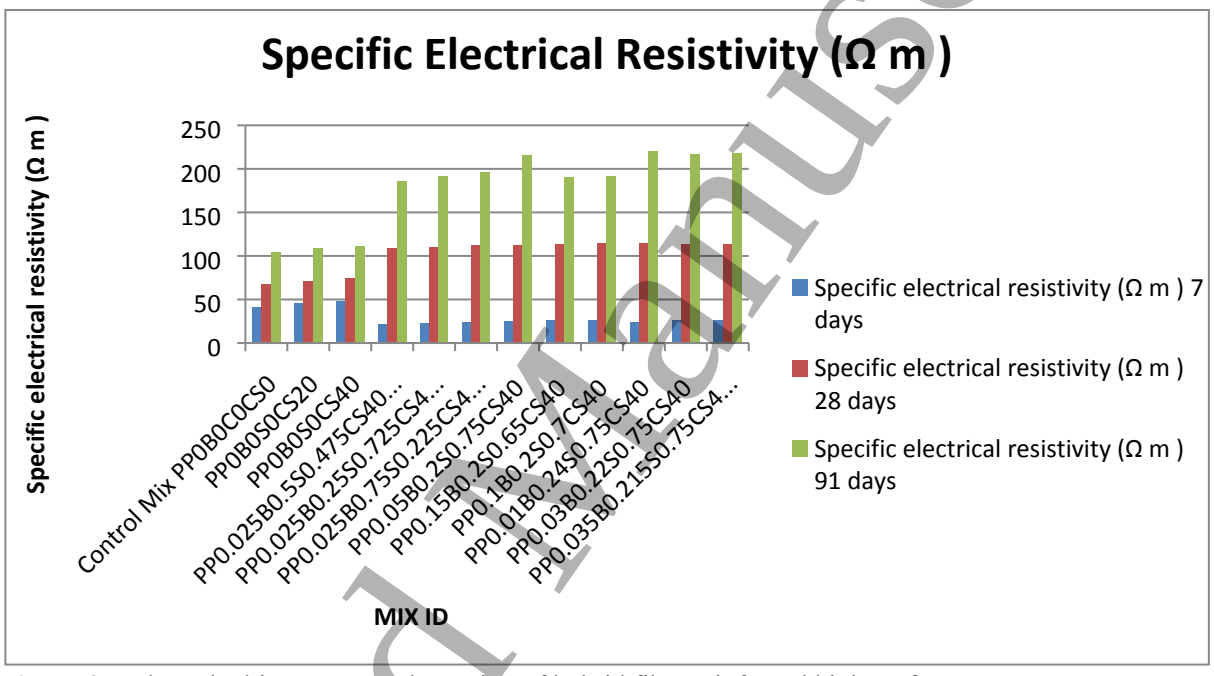


Figure 8 Early and Ultimate water absorption of hybrid fiber reinforced high performance concrete.

Table 8 Specific electrical resistivity (Ω m) at 7 days, 28 days and 91 days of hybrid fiber reinforced high performance concrete.

Mix No.	Mix ID	Specific electrical resistivity (Ω m)		
		7 days	28 days	91 days
1	Control Mix PP0B0C0CS0	41.2	67.8	104.4
2	PP0B0S0CS20	45.5	71.2	108.3
3	PP0B0S0CS40	47.8	74.5	111.2
4	PP0.025B0.5S0.475CS40	21.5	108.5	185.6
5	PP0.025B0.25S0.725CS40	22.6	110.3	191.5
6	PP0.025B0.75S0.225CS40	23.6	111.8	195.6
7	PP0.05B0.2S0.75CS40	24.5	112.4	215.6
8	PP0.15B0.2S0.65CS40	25.6	113.5	190.6
9	PP0.1B0.2S0.7CS40	25.9	114.8	191.8
10	PP0.01B0.24S0.75CS40	23.5	114.9	219.8
11	PP0.03B0.22S0.75CS40	25.8	112.8	216.8
12	PP0.035B0.215S0.75CS40	26.2	113.5	217.2

Accepted Manuscript

PP0B0S0CS0 signifies 0% polypropylene fiber, 0% basalt fiber, 0% hooked end steel fiber, 0% copper slag and same nomenclature with different percentage of fibers are addressed.

4. Conclusion

The research's findings and analysis can be used to draw the following conclusions:

(1) All of the mechanical qualities of concrete are enhanced by the addition of hybrid fibers. The copper slag & silica fume particles serve as a microfiller, densifying the transition zone and strengthening the link between the matrix and the aggregate. Additionally, it has been discovered that silica fume works better in the later phases of curing to increase concrete strength at increasing fiber dosage.

(2) The mechanical properties of all fibers including basalt, PP and hooked end steel fiber are improved as the fiber content is raised. This is because fibers can prevent cracks from spreading, lessen the amount of stress concentrated at the tips of fissures, and slow down the rate at which cracks expand.

(3) The two key elements that contribute to the enhanced performance of PP0.05B0.2S0.75CS40 are the higher tensile strength and elastic modulus of steel fibers with compatible bonding with other two fibers.

(4) The addition of hybrid fibers with supplementry cementitious materials increases the features of the transition zone and has a major impact on the durability properties of concrete. The early and final water absorption were reduced by 25% and 36%, respectively, when fiber hybridization takes place.

(5) The electrical resistivity of concrete is impacted differentially by the addition of various types of fibers. The electrical resistivity of concrete for the optimum mix i.e., PP0.05B0.2S0.75CS40 is 112.4 (Ω m) at 28 days and 215.6 (Ω m) at 91 days.

(6) Concrete's ability to absorb water is significantly reduced when copper slag at 40% and hybrid fibers are combined. The mixture containing 0.05% PP, 0.2 % Basalt fibers and 0.75 % steel fibers has been discovered to have the lowest water absorption of all fiber-reinforced concretes.

List of Acronyms

Hybrid fiber reinforced concrete**:** HFRC
Water to binder ratio: w/b ratio
Steel fiber reinforced concrete: SFRC
Hooked end Steel fibers: HES
Compressive strength: CS
Split tensile strength: STS
Flexural Strength: FS
Ordinary Portland Cement: OPC
Fly ash: FA
Polypropylene fibre: PP
Basalt fiber: BF
Split tensile: ST
Micro steel: MS
Self compacting concrete: SCC
American concrete Institute: ACI
Indian statndard: IS
Compression testing machine: CTM
American society of testings and Materials: ASTM

Data Availability: The data used to support the findings is itself available in this research article and it is author's own experimental work.

Conflicts of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding Information: The research is supported by Delhi Technological University.

Ethical Statement: This paper is not submitted anywhere nor consider anywhere for publication.

References

1. R. Yu , P. Spiesz , H.J.H. Brouwers 2015 Development of Ultra-High Performance Fiber Reinforced Concrete (UHPFRC): Towards an efficient utilization of binders and fibers Construction and Building Materials **79** 273–282
2. Doo-Yeol Yoo , Sung-Wook Kim , Jung-Jun Park 2017 Comparative flexural behavior of ultra-high-performance concrete reinforced with hybrid straight steel fibers Construction and Building Materials **132** 219–229
3. Lihua Xu, Fanghong Wu, Yin Chi , Peng Cheng, Yanqin Zeng, Qian Chen 2019 Effects of coarse aggregate and steel fiber contents on mechanical properties of high performance concrete Construction and Building Materials **206** 97–110
4. Dehong Wang, Yanzhong Ju, Hao Shen, Libin Xu 2019 Mechanical properties of high performance concrete reinforced with basalt fiber and polypropylene fiber Construction and Building Materials **197** 464–473.
5. Ngoc Thanh Tran , Dong Joo Kim 2017 Synergistic response of blending fibers in ultra-high-performance concrete under high rate tensile loads Cement and Concrete Composites **78** 132-145.
6. Iman Sadrinejad, Rahmat Madandoust , Malek Mohammad Ranjbar 2018 The mechanical and durability properties of concrete containing hybrid synthetic fibers Construction and Building Materials **178** 72–82.
7. Vahid Afroughsabet , Togay Ozbakkaloglu 2015 Mechanical and durability properties of high-strength concrete containing steel and polypropylene

Accepted Manuscript

fibers Construction and Building Materials 73–82.
8. H.R. Pakravan , M. Latifi , M. Jamshidi 2017 Hybrid short fiber reinforcement system in concrete: A review Construction and Building Materials **142** 280–294.
9. Guncheol Lee et al. 2012 Combining polypropylene and nylon fibers to optimize fiber addition for spalling protection of high-strength concrete Construction and Building Materials **34** 313–320.
10. Dong Joo Kim et al. 2011 Comparative flexural behavior of Hybrid Ultra High Performance Fiber Reinforced Concrete with different macro fibers Construction and Building Materials **25** 4144–4155.
11. Su-Tae Kang et al. 2016 Hybrid effects of steel fiber and microfiber on the tensile behavior of ultra-high performance concrete Composite Structures **145** 37–42.
12. Vikram Jothi Jayakumar and Sivakumar Anandan Hindawi 2014 Publishing Corporation Advances in Civil Engineering Volume Article ID 363649 9 pages http://dx.doi.org/10.1155/2014/363649
13. Valeria Corinaldesi , Alessandro Nardinocchi 2016 Influence of type of fibers on the properties of high performance cement-based composites Construction and Building Materials **107** 321–331.
14. N. Banthia et al. 2014 Fiber synergy in Hybrid Fiber Reinforced Concrete (HyFRC) in flexure and direct shear. Cement & Concrete Composites **48** 91–97.
15. Xinyu Hu, Yihong Guo et al. 2019 The Mechanical Properties and Chloride Resistance of Concrete Reinforced with Hybrid Polypropylene and Basalt Fibers. Materials Construction and Building Materials **12** 2371 doi:10.3390/ma12152371
16. M.Velumani et al. 2014 Durability And Characteristics of Copper Slag as Fine Aggregate and Fly Ash as Cement in Concrete 2nd International Conference on Current Trends in Engineering and Technology ICCTET'14.
17. Mohamed A.E.M. Ali 2016 Mechanical Performance Of Hybrid Fiber-Reinforced Engineered Cementitious Composite Incorporating Niti-Sma Short Fibers Resilient Infrastructure June 1–4.
18. M.A.E.M. Ali, M.L. Nehdi 2017 Innovative crack-healing hybrid fiber reinforced engineered Cementitious composite Construction and Building Materials **150** 689–702.
19. Bilal Masood et al. 2019 Mechanical and durability performance of recycled aggregate concrete incorporating low calcium bentonite Construction and Building Materials **237** 117760.
20. Mo Alkaysi et al. 2016 Effects of silica powder and cement type on durability of ultra high performance concrete (UHPC) Cement and Concrete Composites 66 47-56.
21. Yu Wang et al. 2020 The effect of tensile and bond characteristics of NiTi shape memory alloy,steel and polypropylene fibers on FRSCC beams under three-point flexural test. Construction and Building Materials **233** 117333
22. John S. Lawler et al. 2005 Microfiber and Macrofiber Hybrid Fiber-Reinforced Concrete J. Mater. Civ. Eng. **17** 595-604.
23. Emad A.H. Alwesabi et al. 2020 Experimental investigation on mechanical properties of plain and rubberised concretes with steel–polypropylene hybrid fiber Construction and Building Materials **233** 117194.
24. Huanghuang Huang et al. 2017 Influence of rice husk ash on strength and permeability of ultra-high performance concrete Construction and Building Materials **149** 621–628.
25. Estela O. Garcez et al. 2011 Development of high strength self-compacting fiber reinforced concrete for prefabricated concrete industry , 0 (201 MATEC Web of Conferences 275 9) ACEM2018 and SBMS1 2011 https://doi.org/10.1051/matecconf/201927502011
26. Safeer Abbas et al. 2015 Exploring mechanical and durability properties of ultra-high performance concrete incorporating various steel fiber lengths and dosages. Construction and Building Materials **75** 429–441.
27. Ye Li et al. 2019 Synergistic effects of hybrid polypropylene and steel fibers on explosive spalling prevention of ultra-high performance concrete at elevated temperature Cement and Concrete Composites https://doi.org/10.1016/j.cemconcomp.2018.11.009
28. Rahesh Hari, K.M. Mini 2019 Mechanical and durability properties of sisal-Nylon 6 hybrid fiber reinforced high strength SCC Construction and Building Materials **204** 479–491.
29. Farhad Aslani et al. The effect of NiTi shape memory alloy, polypropylene and steel fibers on the fresh and mechanical properties of self-compacting concrete.
30. Mojtaba Tabatabaeian et al. 2017 Experimental investigation on effects of hybrid fibers on rheological, mechanical, and durability properties of high-strength SCC Construction and Building Materials **147** 497–509.
31. Hussain, A., Ding, Y., Liu, G., & Naqi, A. 2019. Study on self-monitoring of multiple cracked concrete beams with multiphase conductive materials subjected to bending. Smart Materials and Structures, **28**(9), 095003.
32. Khan, M., Cao, M., Chu, S. H., & Ali, M. 2022. Properties of hybrid steel-basalt fiber reinforced concrete exposed to different surrounding conditions. Construction and Building Materials, **322**, 126340.
33. Ding, Y., Liu, G., Hussain, A., Pacheco-Torgal, F., & Zhang, Y. 2019. Effect of steel fiber and carbon black on the self-sensing ability of concrete cracks under bending. Construction and Building Materials, **207**, 630-639.
34. Hussain, Z., Pu, Z., Hussain, A., Ahmed, S., Shah, A. U., Ali, A., & Ali, A. 2022. Effect of fiber dosage on water permeability using a newly designed apparatus and crack monitoring of steel fiber–reinforced concrete under direct tensile loading. Structural Health Monitoring, **21**(5), 2083-2096.
35. Wang, Q., Ding, Y., Zhang, Y., & Castro, C. 2021. Effect of macro polypropylene fiber and basalt fiber on impact resistance of basalt fiber-reinforced polymer-reinforced concrete. Structural Concrete, **22**(1), 503-515.
36. Ahmed, S., Hussain, A., Hussain, Z., Pu, Z.,

Accepted Manuscript

Ostrowski, K. A., & Walczak, R. 2021. Effect of Carbon Black and Hybrid Steel-Polypropylene Fiber on the Mechanical and Self-Sensing Characteristics of Concrete Considering Different Coarse Aggregates' Sizes. Materials, **14**(23), 7455.

37. Khan, M., Cao, M., Hussain, A., & Chu, S. H. 2021. Effect of silica-fume content on performance of CaCO3 whisker and basalt fiber at matrix interface in cement-based composites. Construction and Building Materials, **300**, 124046.

38. Sharma, S. K. 2013. Zero energy building envelope components: a review. Int J Eng Res Appl, **3**(2), 662-675.

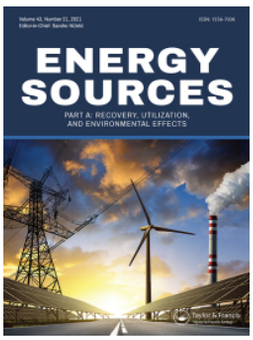
39. Gupta, M., Raj, R., & Sahu, A. K. 2022. Effect of Rice Husk Ash, silica fume & GGBFS on compressive strength of performance based concrete. *Materials Today: Proceedings*, ***55***, 234-239.

40. Gupta, M., Raj, R., & Sahu, A. K. 2022. Mechanical properties of high strength concrete incorporating chopped basalt fibers: experimental and analytical study. *Materials Research Express*, ***9***(12), 125305.

41. Gupta, M., Raj, R., & Sahu, A. K. 2022. EDNN based prediction of strength and durability properties of HPC fibres &. Advances in concrete construction, **14**(3), 185-194.

42. Sharma, S. K., & Lee, J. 2020. Finite element analysis of a fishplate rail joint in extreme environment condition. International Journal of Vehicle Structures & Systems, **12**(5), 503-506.

Accepted Manuscript



Energy Sources, Part A: Recovery, Utilization, and Environmental Effects

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/ueso20

Effect of orange peel oil on the performance and emission characteristics of diesel engine fueled with quaternary blends

Dushyant Mishra, Naveen Kumar & Rajiv Chaudhary

To cite this article: Dushyant Mishra, Naveen Kumar & Rajiv Chaudhary (2023) Effect of orange peel oil on the performance and emission characteristics of diesel engine fueled with quaternary blends, Energy Sources, Part A: Recovery, Utilization, and Environmental Effects, 45:1, 650-660, DOI: 10.1080/15567036.2023.2171510

To link to this article: https://doi.org/10.1080/15567036.2023.2171510

Published online: 31 Jan 2023.

Submit your article to this journal

Article views: 3

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
2023, VOL. 45, NO. 1, 650–660
https://doi.org/10.1080/15567036.2023.2171510

Taylor & Francis
Taylor & Francis Group

Check for updates

Effect of orange peel oil on the performance and emission characteristics of diesel engine fueled with quaternary blends

Dushyant Mishra, Naveen Kumar, and Rajiv Chaudhary

Centre for Advanced Studies and Research in Automotive Engineering, Department of Mechanical Engineering, Delhi Technological University, Delhi, India

ABSTRACT

The present research work investigates the effect of orange peel oil on the ternary blend of diesel–biodiesel of waste cooking oil and n-butanol. The waste cooking oil methyl ester(WCOME) is produced by the transesterification process followed by one-step esterification. Four blends with varying orange peel oil concentrations (3%–21% v/v) with a fixed concentration (5% v/v) of WCOME were used in the present work. Evaluation of physicochemical properties provides sufficient validation for the utilization of fuel blends in a single-cylinder four-stroke air-cooled diesel engine. All the investigations are as per the standard of ASTM. The results show that the addition of orange peel has better results for the brake thermal efficiency and brake-specific energy consumption (BSEC) characteristics of the diesel engine. In comparison to neat diesel, the highest brake thermal efficiency is 22.49%, for the blend has a 21% (v/v) concentration of orange peel oil, while the BSEC is 18.44% for the same blend. The hydrocarbon and carbon monoxide emission was observed to be lowest for O21B15, while 0.46% nitrogen oxide reduction was observed in the same blend as compared to neat diesel.

ARTICLE HISTORY
Received 25 October 2021
Revised 10 November 2022
Accepted 11 November 2022

KEYWORDS
Straight vegetable oil; quaternary blend; orange peel oil; waste cooking oil; engine characteristic

Introduction

The terminal decline and depletion of fossil fuels were the major concerns regarding the current times. However, the consumption of petro-fuels was found to be increasing globally. A total of 1.1% rise in average energy usage has been observed in the transportation sector (Othman et al. 2017). More than 6.0% of the total growth in world consumption of fossil-based fuels from 2010 to 2040 comes from the transportation sector only (Mofijur et al. 2016). For a growing economy like India, the agriculture and transportation sectors may provide major support in the development and welfare of the nation.

The reliance of the agriculture and transportation sector on fossil-based fuels required rigorous innovations in the field of alternative fuels during the emerging energy crisis worldwide. However, the use of fossil fuels for fulfilling the energy requirements has led to the cause of undesirable environmental effects. This involves environmental degradation and pollutant emissions, such as particulate matter (PM), smoke, nitrogen oxide (NOx), carbon monoxide (CO), and carbon dioxide (CO_2). This makes the replacement of conventional fuels like gasoline, diesel, and crude oil, etc. with non-conventional fuels either partially or completely. Previous works in the literature have revealed that the production of biofuels and their application in CI engines cause a substantial downfall of up to 40% in greenhouse gas emissions (Kim and Dale 2005). The suitability of alternative fuel majorly depends on its physicochemical properties when compared to petro-fuel while used in compression ignition (CI) engines (Tesfa et al. 2014). However, some extensive physicochemical properties, such as kinematic viscosity, calorific value, cetane index, boiling point, and cold flow properties, must be in the suitable range of fuel (Sakthivel et al. 2018).

CONTACT Dushyant Mishra iitddushyantmishra@gmail.com Centre for Advanced Studies and Research in Automotive Engineering, Department of Mechanical Engineering, Delhi Technological University, Main Bawana Road, Delhi 110042, India
© 2023 Taylor & Francis Group, LLC

The application of methyl ester in the existing CI engine is not that efficient when compared with the diesel-fuel-operated engine. The use of waste cooking oil (WCO) has been considered as a desirable feedstock for producing biodiesel which can be a relevant fuel for CI engines (Nayak et al. 2021). However, its use in a CI engine has too started in up-rise of harmful gaseous emissions, mainly of NOx emissions (Mishra, Chauhan, and Mishra 2020). Hoang and Pham (2018) stated that long-term use of vegetable oils as an engine-fuel for diesel engines may result in a significant rise in toxic emissions, pollutants along with the reduced engine power, and degrades the quality of lubrication oil (LO) because of deposit development, which damaged the engine significantly. To solve this issue, a variety of peroxides and alcohol-based additives used as cetane improvers are blended with biodiesel to increase engine performance and emission standards (Roslan, Veza, and Said 2020). A review study conducted by Hoang (2021) concludes that there is a reduction in emission characteristics of diesel engine using nanoparticles. The author also concluded that NOx, CO, and hydrocarbon (HC) emissions were reduced by using CeO_2 nanoparticles in the CI engine.

Sometimes, vegetable oils (bio-oils) and their biodiesel have been tested as prime alternative fuels for their utilization in an internal combustion engine (Yilmaz, Atmanli, and Vigil 2018). However, there are certain technical limitations for using straight vegetable oil as a vehicle fuel. A study was conducted by Hoang, Le, and Pham (2019) to investigate the spray effect of preheated and unpreheated straight jatropha oil with diesel in CI engine. It was observed that, after a long-term test of about 300 h on a CI engine, the cone angle is reduced by 3° for straight jatropha oil, both thermal efficiency and heat release rate also getting reduced significantly. It was also observed that straight vegetable oils are preheated before being injected into the engine or are blended with diesel (Yilmaz and Morton 2011). These techniques may have been able to resolve the problems to some extent but were not able to eliminate them.

Atmanli et al. (2015) has conducted an experimental investigation using ternary blends of diesel-vegetable oil-alcohol in CI engine. The authors have concluded that the brake thermal efficiency (BTE), brake mean effective pressure (BMEP), and exhaust gas temperature (EGT) have decreased with the utilization of the above-mentioned blends in the CI engine. A similar investigation conducted by Atmanlı, İleri, and Yüksel (2014), the emissions of CO and NO increased while decreasing trends in HC and CO_2 emissions when compared to diesel. To overcome this, Yesilyurt et al. (2020) have conducted an experimental study using vegetable oil with diesel–biodiesel–alcohol blends to form the quaternary blends that uplift the flow properties (specific gravity and kinematic viscosity) of the overall blends. The authors havefigured out that diesel–biodiesel–vegetable oil–alcohol blends can prove to be a favorable fuel mixture to ameliorate combustion, a performance characteristic with a reduction in exhaust emissions.

Appavu, Ramanan M, and Venu (2019) conducted an investigation on CI engine operated with a quaternary blend of diesel, biodiesel, vegetable oil, and alcohol. The authors observed that BTE increased for all the quaternary blends when adding up the concentration of alcohol. The study was conducted by Yilmaz, Atmanli, and Vigil (2018) using vegetable oil in a quaternary blend. The authors compared the effect of long carbon-chain alcohols (propanol and pentanol) on the engine characteristics of CI engine and concluded that basic fuel properties were found to be enhanced upon adding up the alcohol. Moreover, as compared to propanol, pentanol test fuel blends showed a reduced trend for BSFC.

The goal of the current experimental investigation is to determine how orange peel oil, a vegetable oil, affects the performance and emission characteristics of diesel engines that have not been modified by adding it to ternary mixes of diesel, biodiesel, and alcohol (to create quaternary blends). Upon the blending of orange peel oil into the ternary blend, the blend becomes quaternary, so orange peel oil is used as fuel extender Extraction of orange peel oil from the orange peel (agro-waste) is done by the process of cold pressing of peels. The presence of a major fraction of D-limonene (approximately 91%) components among all the volatile components in orange peel oil made it an attractive alternative fuel for utilizing it in the CI engines, either completely or in blends. From the previous literature (Jorge, Da Silva, and Aranha 2017), it can be seen that D-Limonene has an antioxidant nature that resulted in improvement of the

combustion rate of the fuel (Jha et al. 2019). In the absence of unsaturated fats, the orange peel oil is considered a less-viscous low-cetane fuel because of its shorter carbon-chain length, which leads to a lower cetane index and less viscosity (Vallinayagam et al. 2015).

The novelty of the experimental investigation

The literature survey shows several studies have been conducted on fuel blends of biodiesel, alcohol, and diesel in the form of binary and ternary blends. However, very such types of investigation were conducted on CI engines by using quaternary fuel blends. In the present investigation, the use of straight vegetable oil (orange peel oil) as a fuel extender was done in ternary blends of biodiesel (waste cooking oil methyl ester [WCOME]), alcohol (n-butanol), and pure diesel. However, the concentration of WCO biodiesel is kept fixed at 5% throughout the experimentation, according to the current biofuel policy of India. The results from the present investigation reveal that the addition of orange peel oil improves the BTE and brake-specific energy consumption (BSEC) of CI engine. Along with NOx, it helps to bring down the emissions of hydrocarbon (HC), CO, and smoke opacity.

Materials and methods

Preparation of waste cooking oil (WCO) and fuel blends

WCO of soybean refined oil was collected from a nearby restaurant. The WCO collected was filtered through cellulose filter paper (grade 1) and heated at 70 °C to 100 °C so that the water content was removed completely. The oil was then stored in an airtight bottle. The free fatty acid (FFA) content of the WCO was found to be 3.5% that implies the requirement of a two-stage esterification cum transesterification process for the production of biodiesel (methyl ester of WCO). The schematic arrangement of methyl ester production is depicted in Figure 1. The four fuel blends were prepared by volume-wise substitution of WCO, n-butanol, and orange peel oil in neat diesel.

Experimental setup

A naturally aspirated, 1-cylinder, 4-stroke Direct Injection compression ignition engine was used for the experimental investigation. Studies were conducted on CI engine set up as depicted in Figure 2. The specification of the engine setup is also illustrated in Table 1. In this study, the experiments were successfully performed with 200 bar injection pressure using four different test fuel blends at 1500 rpm engine speed and seven varying loads (0.05, 0.1, 0.15, 0.20, 0.25, 0.3, and 0.35 MPa). All engine testing is performed as per IS 10,000 (Part viii) standard. The test fuel blends were prepared using the splash blending method. The concentration of WCOME is kept constant at 5% (v/v), while that of n-butanol (2–14% v/v) and orange peel oil (3–21% v/v) is varied in the diesel. The various physicochemical properties of the prepared fuel blends are provided in Table 2 and measuring instruments are in accordance with ASTM D-6751 standards in Table 3. First, reference data are obtained by performing the experimental run on a diesel engine fueled by neat diesel. The prepared test fuel blends were tested with the same operating conditions. An AVL smoke meter is used to measure the smoke opacity. However, the AVL di-gas analyzer was used to measure the NOx, CO, and HC emissions. The specifications of both the equipment are shown in Table 4. Information related to equipment used in the investigation and their working range, accuracy, and uncertainties is given in Table 5.

Results

This section gives the outcomes of engine characteristics (performance and emission) for neat diesel and different fuel blends utilized in this study.

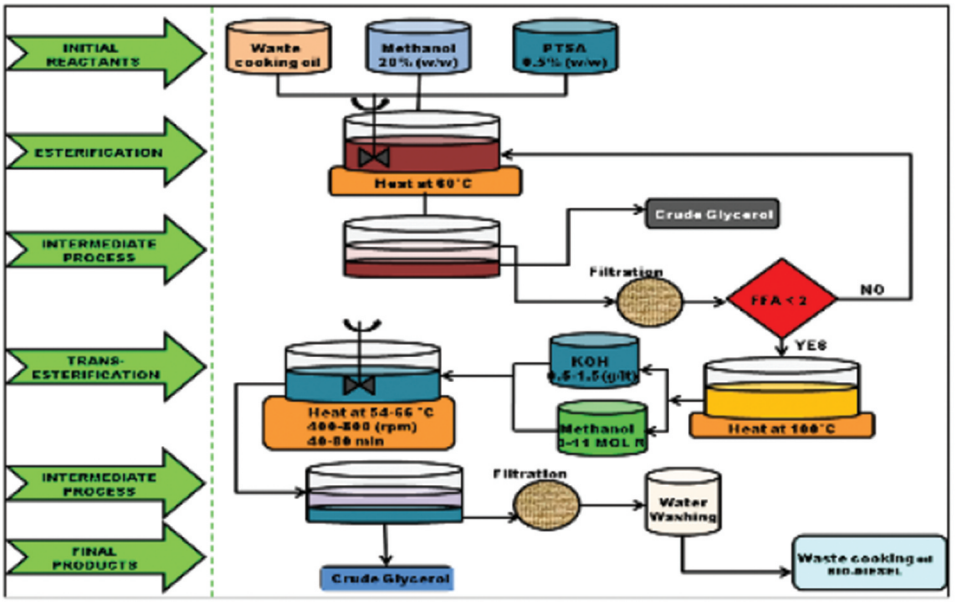


Figure 1. Waste cooking oil methyl ester production.

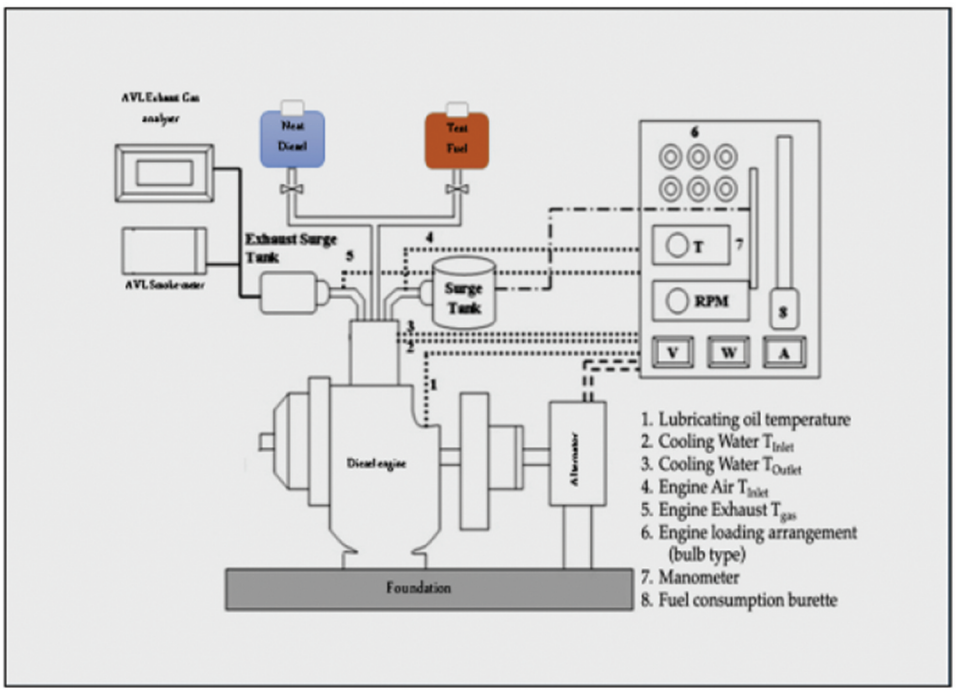


Figure 2. Engine test rig.

Table 1. Technical description of the engine and alternator.

Specification of Diesel Engine			
Make	Kirloskar	Cooling Type	Air cooled (radial cooled)
Model	DAF 8	Fuel injection timing	23 degree BTDC
Brake power (kW)	3.5	No. of injector holes	3
Cylinders	1	Nozzle diameter (mm)	9
Rated speed (rpm)	1500	DI type	Mechanical pump
Injection Ppressure (bar)	200	Compression ratio	17.5:1
Stroke x ore (mm)	110 × 87.5	Lubricating system	Forced feed
Connecting rod length (mm)	234	Cubic capacty	0.78 L
Displacement	0.661	Starting	Hand start with cranking handle
Dynamometer specification			
Manufacturer	Kirloskar Electric Company. Limited. India	Rated voltage (V)	230
Rated output (KVA)	5 @1500 rpm		
Dynamometer type	1 phase, 50 Hz, AC alternator	Rated current (A)	32.6

Table 2. Physicochemical properties of fuel blends.

Property	Unit	n-Butanol	OPO	Diesel	WCOME	O3B2	O9B6	O15B10	O21B14
Kinematic viscosity @40°C	mm^2/s	3.6	1.71	3.1	6.21	3.17	2.65	2.39	2.20
Density @15°C	kg/m^3	810	851	833	892	822.6	823.6	826.7	829.1
Flash point	^0C	37	75	74	154	85.8	85.2	81.6	76.7
Pour point	^0C	−45	-	−15 to 5	−11	−12.6	13.8	14.6	−15.6
Cloud point	^0C	−115.5	-	−13 to15	−6.2	−5.7	−6.8	−8.5	−10.3
Calorific value	MJ/kg	33.20	41.72	45.8	43.2	44.6	46.2	48.1	45.5
Cetane number	-	13.21	39	59	53	51.3	49.6	47.1	45.6
Oxygen (%)									
C/H ratio			7.51						

Table 3. Equipment used for property evaluation.

Property	Equipment & Standard Method	Accuracy
Kinematic viscosity @ 40°C (mm^2/s)	Visco bath, Petrotest ASTM D445	±0.01 mm^2/s
Density @ 15°C (kg/m^3)	Anton Parr DMA 4500 ASTM D4052	±0.1 kg/m^3
Flash point (^0C)	Pensky-Martens (AD0093–600) ASTM D93	±0.1°C
Cloud Point (^0C)	-	-
Pour Point (^0C)	-	-
Calorific value (MJ/kg)	Parr 6100 Oxygen Bomb Calorimeter ASTM D240	±0.001 MJ/kg
Sulfur content (ppm)	XOS Sindie Sulphur analyzer ASTM D7039	0.5 (repeatability)
Cetane number	ASTM D4737	±0.1

Table 4. Emission Test rig specification.

S.N	Gas	Measurement technique	Measurement range (by volume)	Resolution	% Uncertainity
Instrument Name:- AVL DI Gas Analyzer					
1	Carbon monoxides	Non dispersive infra-red sensor	0–10%	0.01% volume	0.20%
2	Hydrocarbon	Fame ionization detector (FID)	0–20,000 ppm	1 ppm	0.20%
3	Oxides of nitrogen	Chemi luminescence principle, electrochemical sensor	0–5,000 ppm	1 ppm	0.20%
Instrument Name:-AVL Somke meter					
1	Smoke opacity	Hatridge Principle	0–100%	±1% volume	0.10%

Table 5. Equipment used for properties assessment.

Property	Equipment and Standardize Method	Accuracy
Kinematic viscosity @ 40°C (mm^2/s)	Visco bath, Petrotest ASTM D445	±0.01 mm^2/s
Density @ 15°C (kg/m^3)	Density-meter Anton Parr DMA 4500 ASTM D4052	±0.1 kg/m^3
Flash point (^0C)	Pensky-Martens (AD0093–600) ASTM D93	±0.1°C
Sulfur content (ppm)	XOS Sindie Sulphur analyzer ASTM D7039	0.5 (repeatability)
Cloud Point (^0C)	-	-
Pour Point (^0C)	-	-
Calorific value (MJ/kg)	Parr 6100 Oxygen Bomb Calorimeter ASTM D240	±0.001 MJ/kg
Cetane number	ASTM D4737	±0.1

Performance characteristic

Brake thermal efficiency (BTE)

BTE signifies the engine's ability to change over the fuel power into useful mechanical work. The performance characteristics of the used engine are generally evaluated by using the BTE. Higher BTE indicates higher engine performance. Figure 3(a) shows the BTE vs BMEP graph for different fuel blends. The trends in Figure 3(a) show an up-rise in the BTE (%) with the increase in BMEP (MPa) values. In addition, an increase in the % volume fraction of orange peel oil in fuel blends leads to an increase in BTE at a particular BMEP. This is caused to a higher volumetric composition of orange peel oil in fuel blends, which improves combustion characteristics due to its lower viscosity and boiling point. However, the calorific value of orange peel oil is slightly lower, but according to Vallinayagam et al. (2013), combustion process can be improved by mixing of low-viscous and low-boiling point that improve the mixing process. Moreover, the vaporization process will also be improved. The BTE obtained are 21.99, 22.09%, 22.31%, and 22.49%, for O3B2, O9B6, O15B10, and O21B14, respectively, at higher loads, as compared to neat diesel. The results are satisfied with the study by Deep et al. (2015).

Brake-specific energy consumption (BSEC)

The energy is obtained by fuel combustion per hour to produce one kW brake power known as BSEC. Figure 3(b) shows the BSEC vs BMEP graph for different fuel blends. The graph indicates that when the BMEP increases, the BSEC drops that owing to efficient burning of the fuel mixture. Also, with the increase in the concentration of orange peel oil in the fuel blends, a decreasing trend of BSEC is observed at a particular BMEP. This is due to the mixing of low viscous orange peel oil in the fuel blends. The combustion inside the cylinder depends on the Sauter mean diameter. The injection velocity is higher for the fuel having lower viscosity as observed by (Kumar and Tomar 2019). Lower Sauter mean diameter leads to better combustion which significantly reduced the BSEC. At higher loads, the BSEC obtained are 4.10%, 8.16%, 12.27%, and 18.44% for O3B2, O9B6, O15B10, and O21B14, respectively, as compared to neat diesel. According to Anandavelu, Alagumurthi, and Saravannan (2011), the author has found that the fuels that have similar calorific values show an increase in the BSEC for the fuel having higher viscous properties.

Emission characteristic

Hydrocarbon (HC) emission

HC formation during engine combustion depends upon various factors, such as poor air entrainment, improper fuel-blends, and air mixing, and formation of fuel-rich zone regions (Baskar and Senthilkumar 2016). Figure 4(a) shows the HC vs BMEP graph for different fuels blends. The trend in the figure shows that the increase in BMEP resulted that the HC emissions also increase. Also, with increasing the dosage of orange peel oil in the fuel blends, the decreasing trend of HC is observed at a particular BMEP. As compared to neat diesel at maximum load the HC emission has been reduced by 9.72% for O3B2, 19.44% for O9B6, 34.72% for O15B10, and 40.27% for O21B15, respectively, as

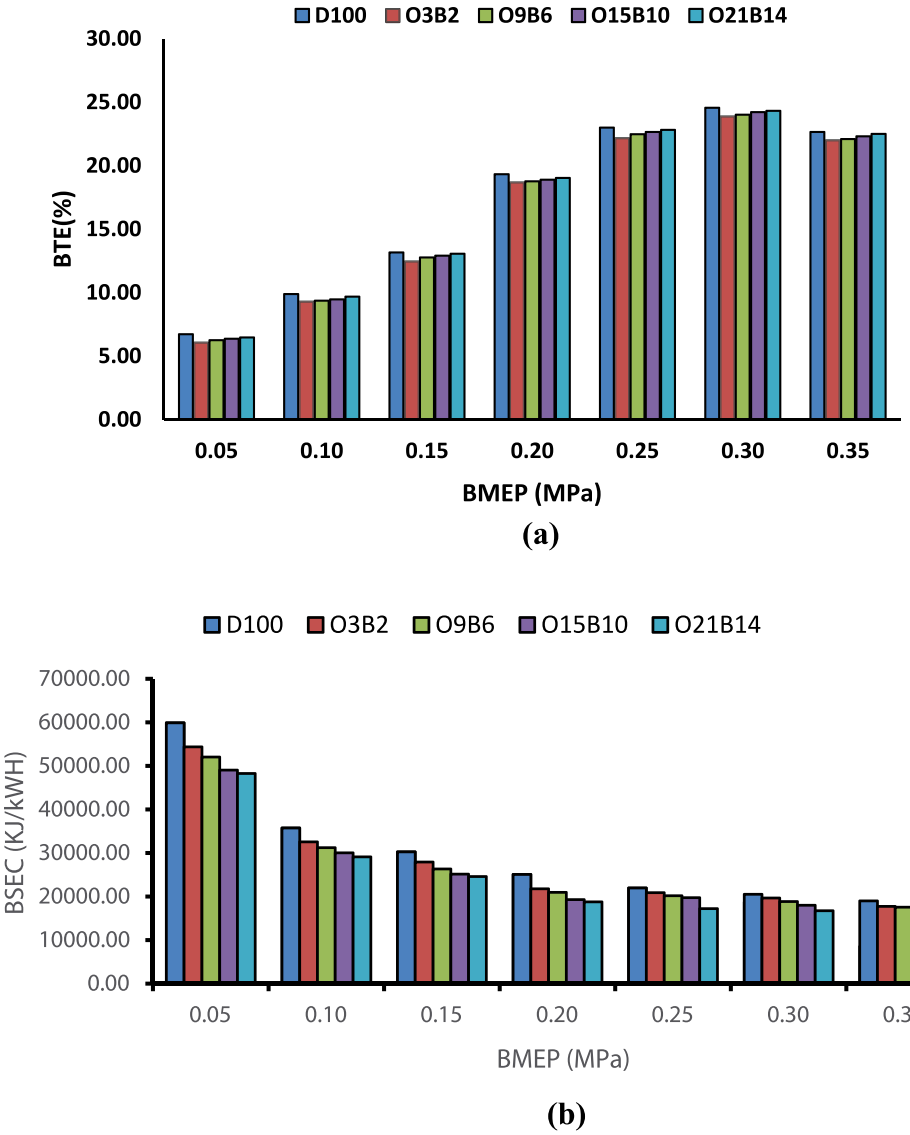


Figure 3. Effect of orange peel oil addition on (a) BTE and (b) BSEC on different engine loads at 1500 rpm.

seen from the figure. The trend observed in the study conducted by Barnawal and Kumar (2022) was quite similar. Park, Suh, and Lee (2010) state that the atomization process inside the cylinder is affected by the surface tension of the fuel and its viscosity, so that affects the mixing process. Due to the less viscous orange peel oil, its increasing concentration in the fuel blends improves the atomization of the fuel blends that leads to the formation of finer droplets that enhance the evaporation process inside the cylinder.

Nitrogen oxides (NOx) emission

NOx formation during engine combustion depends upon the oxygen content and mass of fuel burned. Generally, there are three types of NOx formation, i.e.,, thermal, fuel, and prompt NOx. For all of them, thermal NOx is significant for improving engine efficiency (Omari et al. 2017). Figure 4(b) shows NOx vs BMEP graph for different fuel blends. The trend in the figure shows that emission of

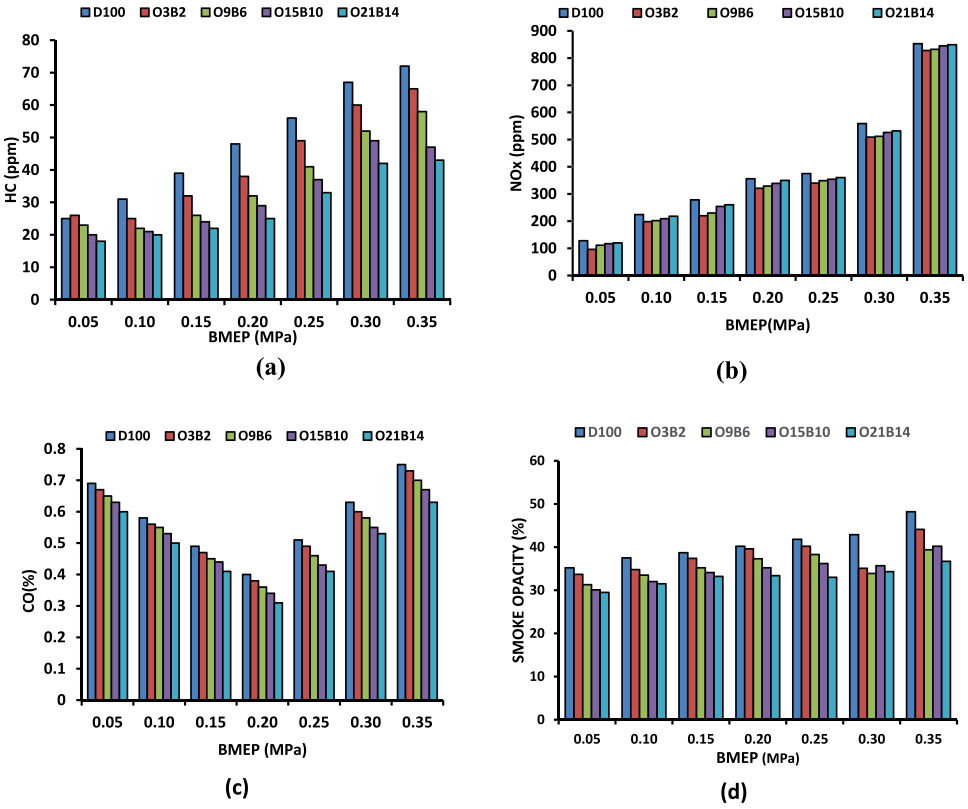


Figure 4. Variation of emission parameter with BMEP.

NOx increases with the increase in BMEP. Also, with the increase in the concentration of orange peel oil in the fuel blends, an increasing trend of NOx is observed at a particular BMEP. The trend was also similar to that found by Ganesan and Masimalai (2019). As seen from the figure, the NOx emission has been reduced by 2.93% for O3B2, 2.49% for O9B6, 0.93% for O15B10, and 0.46% for O21B15, respectively, at full load with reference to neat diesel. This is because of the add up in the concentration of orange peel oil that lowers the overall cetane index of the fuel blends. As the cetane index is a significant parameter for defining the self-ignition property of the fuel, the reduction in the cetane index increases the period of ignition delay under all loading conditions. This longer ignition delay period causes the accumulation of a larger amount of air-fuel mixture that results in a higher in-flame temperature during the premixed combustion stage. Lee et al. 2011; Zhang et al. (2012) showed that mixing the fuel having a lower cetane number could result in fuel deposition during a longer ignition delay period that leads to an increase in the combustion temperature and heat release rate.

Carbon monoxide (CO) emissions

The formation of CO inside the engine is due to partial or incomplete combustion due to the relatively lower entrained oxygen or improper air–fuel ratio during combustion reactions. CO emission formation is generally affected by the equivalence ratio, quality of fuel burning, injection parameters and timing, combustion chamber design, and many others (Annamalai et al. 2016; Nantha Gopal and Thundil Karupparaj 2015). Figure 4(c) shows the CO vs BMEP graph for different fuel blends. The trend in the figure shows that with the increase in BMEP, the CO emissions first decrease and then increase. Also, with the surge in the dosage of orange peel oil in the fuel blends, a decreasing trend of CO is observed at a particular BMEP. As seen from the Figure, the CO emission has reduced by 2.66%

for O3B2, 6.66% for O9B6, 10.66% for O15B10, and 16.00% for O21B15, respectively, as compared with pure diesel at full load condition. The availability of orange peel oil, an antioxidant component, is accountable for the decreasing trend in CO emissions for various fuel blends (Jorge, Da Silva, and Aranha 2017).

Smoke opacity

Smoke formation, in general, can be correlated with the formation of CO and HC emissions for all biofuels. Figure 4 (d) portrays the smoke vs BMEP graph for different fuel blends. The trend in the figure shows that with the increase in BMEP, there is a slight increase in smoke emissions. Also, with the increase in the concentration of orange peel oil in the fuel blends, the decreasing trend of smoke is observed at a particular BMEP. As seen from the figure, the smoke emission has reduced by 9.72% for O3B2, 19.44% for O9B6, 34.72% for O15B10, and 40.27% for O21B15, respectively, in comparison to pure diesel when the engine operated at full load. The decreasing trend is due to the presence of orange peel oil in the fuel blends that is an antioxidant additive that improves the combustion process and quite similar trend observed by (Jorge, Da Silva, and Aranha 2017). Also, because of the lower viscosity of the orange peel oil, its increasing concentration in the fuel blends enhances the atomization that leads to the smaller size of droplets and improves the evaporation process inside the cylinder, thus improving the combustion process. This, overall, resulted in lower smoke formation.

Conclusion

The effect of varying the amount of orange peel oil in quaternary blends of diesel-biodiesel (WCO)-alcohols (n-butanol)-vegetable oil (orange peel) at 3%, 9%,15$, and 21% were investigated in this study. The investigation has been done at a constant engine speed without any major modifications. The physiochemical properties of the quaternary blends are found to be in accordance with ASTM standards. Finally, the effect of the addition of orange peel oil in the diesel-biodiesel-alcohol blends was investigated on the engine characteristics. The main conclusions from this study are as follows:

- The BTE of all the quaternary blends at higher loads is above 20% higher as compared to neat diesel and it increases with an addition of orange peel oil concentration in the blends. This is because orange peel has better combustion properties due to its lower viscosity and lower boiling point.
- The BSEC of all quaternary mixes is found to be lower than that of pure diesel, and at higher loads it declines from 4% to 19% as the fraction of orange peel oil in the blend increases. This is due to the presence of lower viscosity oil in fuel mixtures, which reduces BSEC.
- HC, CO, and smoke opacity of all the quaternary fuel mixtures are achieved to be lowered, as compared to pure diesel. Up to 40% reduction in HC and 16% reduction in CO emissions are found. Moreover, these exhaust emissions were decreasing by adding the orange peel oil in the test fuel blends.
- NOx emissions of all the quaternary blends are observed to be less than pure diesel and it increases with the rise in orange peel oil concentration in the blends. The maximum NOx reduction of about 3% at minimum fraction of orange peel and upon minimum fraction of orange peel it was up to 0.5%. This is because the addition of orange peel oil lowers the overall cetane index of the fuel blends that leads to an increase in the ignition delay period.

Finally, it can be concluded that the addition of orange peel oil in the diesel-biodiesel-alcohol blends can replace diesel fuel with improved engine performance and emission characteristics. The main motivation of this study is to identify orange peel oil as a suitable vegetable oil that can be utilized in quaternary blends.

More studies are required in the field of stability and production aspects before using orange peel oil as an additive in the blends. Moreover, a detailed study of spray formation inside the engine cylinder is required to further investigate the combustion characteristics, namely inside cylinder temperature and pressure, heat release rate.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Dushyant Mishra http://orcid.org/0000-0002-0946-7673
Naveen Kumar http://orcid.org/0000-0002-4408-9997

References

Anandavelu, K., N. Alagumurthi, and C. G. Saravannan. 2011. Experimental investigation of using eucalyptus oil and diesel fuel blends in Kirloskar TV1 direct injection diesel engine. *Journal of Sustainable Energy & Environment* 2 (March):93–97.

Annamalai, M., B. Dhinesh, K. Nanthagopal, P. SivaramaKrishnan, J. Isaac JoshuaRamesh Lalvani, M. Parthasarathy, and K. Annamalai. 2016. An assessment on performance, combustion and emission behavior of a diesel engine powered by Ceria nanoparticle blended emulsified biofuel. *Energy Conversion and Management* 123:372–80. Elsevier Ltd. doi:10.1016/j.enconman.2016.06.062.

Appavu, P., V. Ramanan M, and H. Venu. 2019. Quaternary blends of diesel/biodiesel/vegetable oil/pentanol as a potential alternative feedstock for existing unmodified diesel engine: Performance, combustion and emission characteristics. *Energy* 186. 115856. Elsevier Ltd. doi:10.1016/j.energy.2019.115856.

Atmanlı, A., E. İleri, and B. Yüksel. 2014. Experimental investigation of engine performance and exhaust emissions of a diesel engine fueled with Diesel–n-Butanol–Vegetable oil blends. *Energy Conversion and Management* 81:312–21. doi:10.1016/j.enconman.2014.02.049.

Atmanli, A., E. Ileri, B. Yuksel, and N. Yilmaz. 2015. Extensive analyses of Diesel–Vegetable Oil–n-Butanol Ternary Blends in a diesel engine. *Applied Energy* 145:155–62. doi:10.1016/j.apenergy.2015.01.071.

Barnawal, S. K., and N. Kumar. 2022. Environmental effects experimental investigation and artificial neural network modeling of performance and emission of a CI engine using orange peel oil- diesel blends. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects* 44 (1):1–15. Taylor & Francis. doi:10.1080/15567036.2021.1967520.

Baskar, P., and A. Senthilkumar. 2016. Effects of oxygen enriched combustion on pollution and performance characteristics of a diesel engine. *Engineering Science and Technology, an International Journal* 19 (1):438–43. Elsevier B.V. doi:10.1016/j.jestch.2015.08.011.

Deep, A., N. Kumar, M. Kumar, A. Singh, D. Gupta, and J. Singh Patel. 2015. "Performance and emission studies of diesel engine fuelled with orange peel oil and N-Butanol alcohol blends." *SAE Technical Papers*. doi:10.4271/2015-26-0049.

Ganesan, N., and S. Masimalai. 2019. Experimental investigation on a performance and emission characteristics of single cylinder diesel engine powered by waste orange peel oil biodiesel blended with antioxidant additive. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects* 0 (0):1–12. Taylor & Francis. doi:10.1080/15567036.2019.1604856.

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. Elsevier B.V. doi:10.1016/j.fuproc.2021.106840.

Hoang, A. T., A. T. Le, and V. V. Pham. 2019. A core correlation of spray characteristics, deposit formation, and combustion of a high-speed diesel engine fueled with Jatropha oil and diesel fuel. *Fuel* 244 (February):159–75. Elsevier. doi:10.1016/j.fuel.2019.02.009.

Hoang, A. T., and V. V. Pham. 2018. Impact of Jatropha oil on engine performance, emission characteristics, deposit formation, and lubricating oil degradation impact of Jatropha oil on engine performance, emission characteristics, deposit formation, and lubricating oil. *Combustion Science and Technology* 00 (00):1–16. Taylor & Francis. doi:10.1080/00102202.2018.1504292.

Jha, P., S. Singh, M. Raghuram, G. Nair, R. Jobby, A. Gupta, and N. Desai. 2019. Valorisation of orange peel: Supplement in fermentation media for ethanol production and source of limonene. *Environmental Sustainability* 2 (1):33–41. Springer Singapore. 10.1007/s42398-019-00048-2

Jorge, N., A. C. Da Silva, and C. P. M. Aranha. 2017. Antioxidant activity of oils extracted from Orange (Citrus Sinensis) seeds. *International News on Fats, Oils and Related Materials* 28 (6):23–26.

Kim, S., and B. E. Dale. 2005. Life cycle assessment of various cropping systems utilized for producing biofuels: Bioethanol and biodiesel. *Biomass & bioenergy* 29 (6):426–39. doi:10.1016/j.biombioe.2005.06.004.

Kumar, N., and M. Tomar. 2019. Influence of nanoadditives on ignition characteristics of Kusum (Schleichera Oleosa) biodiesel. *International Journal of Energy Research* 43 (8):3223–36. doi:10.1002/er.4446.

Lee, W. J., Y. C. Liu, F. K. Mwangi, W. H. Chen, S. L. Lin, Y. Fukushima, C. N. Liao, and L. C. Wang. 2011. Assessment of energy performance and air pollutant emissions in a diesel engine generator fueled with water-containing ethanol-biodiesel-diesel blend of fuels. *Energy* 36 (9):5591–99. Elsevier Ltd. doi:10.1016/j.energy.2011.07.012.

Mishra, S., A. Chauhan, and K. B. Mishra. 2020. Role of binary and ternary blends of WCO biodiesel on emission reduction in diesel engine. *Fuel* 262:116604. doi:10.1016/j.fuel.2019.116604.

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.

Nantha Gopal, K., and R. Thundil Karupparaj. 2015. Effect of Pongamia Biodiesel on emission and combustion characteristics of Di compression ignition engine. *Ain Shams Engineering Journal* 6 (1):297–305. Faculty of Engineering, Ain Shams University. doi:10.1016/j.asej.2014.10.001.

Nayak, S. K., A. T. Hoang, B. Nayak, and P. C. Mishra. 2021. Influence of fish oil and waste cooking oil as post mixed binary biodiesel blends on performance improvement and emission reduction in diesel engine. *Fuel* 289 (October 2020):119948. Elsevier Ltd. doi:10.1016/j.fuel.2020.119948.

Omari, A., S. Pischinger, O. P. Bhardwaj, B. Holderbaum, J. Nuottimäki, and M. Honkanen. 2017. Improving engine efficiency and emission reduction potential of HVO by fuel-specific engine calibration in modern passenger car diesel applications. *SAE International Journal of Fuels and Lubricants* 10 (3):756–67. doi:10.4271/2017-01-2295.

Othman, M. F., A. Adam, G. Najafi, and R. Mamat. 2017. Green fuel as alternative fuel for diesel engine: A review. *Renewable and Sustainable Energy Reviews* 80 (August 2016):694–709. doi:10.1016/j.rser.2017.05.140.

Park, S. H., H. K. Suh, and C. S. Lee. 2010. Nozzle flow and atomization characteristics of ethanol blended biodiesel fuel. *Renewable Energy* 35 (1):144–50. doi:10.1016/j.renene.2009.06.012.

Roslan, M. F., I. Veza, and M. F. M. Said. 2020. "Predictive simulation of single cylinder N-Butanol HCCI engine." *IOP Conference Series: Materials Science and Engineering* 884 (1). doi:10.1088/1757-899X/884/1/012099.

Sakthivel, R., K. Ramesh, R. Purnachandran, and P. Mohamed Shameer. 2018. A review on the properties, performance and emission aspects of the third generation biodiesels. *Renewable and Sustainable Energy Reviews* 82 (November):2970–92. doi:10.1016/j.rser.2017.10.037.

Tesfa, B., F. Gu, R. Mishra, and A. Ball. 2014. Emission characteristics of a CI engine running with a range of biodiesel feedstocks. *Energies* 7 (1):334–50. Multidisciplinary Digital Publishing Institute. doi:10.3390/en7010334.

Vallinayagam, R., S. Vedharaj, W. M. Yang, P. S. Lee, K. J. E. Chua, and S. K. Chou. 2013. Combustion performance and emission characteristics study of pine oil in a diesel engine. *Energy* 57:344–51. doi:10.1016/j.energy.2013.05.061.

Vallinayagam, R., S. Vedharaj, W. M. Yang, W. L. Roberts, and R. W. Dibble. 2015. Feasibility of using Less Viscous and Lower Cetane (LVLC) fuels in a diesel engine: A review. *Renewable and Sustainable Energy Reviews* 51:1166–90. doi:10.1016/j.rser.2015.07.042.

Yesilyurt, M. K., M. Aydin, Z. Yilbasi, and M. Arslan. 2020. Investigation on the structural effects of the addition of alcohols having various chain lengths into the vegetable oil-biodiesel-diesel fuel blends: An attempt for improving the performance, combustion, and exhaust emission characteristics of a Compressi. *Fuel* 269:117455. doi:10.1016/j.fuel.2020.117455.

Yilmaz, N., A. Atmanli, and F. M. Vigil. 2018. Quaternary blends of diesel, biodiesel, higher alcohols and vegetable oil in a compression ignition engine. *Fuel* 212:462–69. doi:10.1016/j.fuel.2017.10.050.

Yilmaz, N., and B. Morton. 2011. Effects of preheating vegetable oils on performance and emission characteristics of two diesel engines. *Biomass & bioenergy* 35 (May):2028–33. doi:10.1016/j.biombioe.2011.01.052.

Zhang, Q., M. Yao, Z. Zheng, H. Liu, and J. Xu. 2012. Experimental study of N-Butanol addition on performance and emissions with diesel low temperature combustion. *Energy* 47 (1):515–21. doi:10.1016/j.energy.2012.09.020.

Emotion Recognition in Speech by Multimodal Analysis of Audio and Text

Siddhant Bikram Shah
Computer Science and Engineering
Delhi Technological University
New Delhi, India
siddhantshah3000@gmail.com

Shubham Garg
Computer Science and Engineering
Amity University
Noida Uttar Pradesh, India
shubgarg17@gmail.com

Aikaterini Bourazeri
Computer Science and Electronic Engineering
University of Essex
Colchester, United Kingdom
a.bourazeri@essex.ac.uk

2023 13th International Conference on Cloud Computing, Data Science & Engineering (Confluence) | 978-1-6654-6263-1/23/$31.00 ©2023 IEEE | DOI: 10.1109/Confluence56041.2023.10048872

***Abstract*—Emotion recognition remains a very challenging task in research because of its sensitive and multifaceted nature. Recently, emotion recognition has garnered a lot of attention owing to its significance in psychology, human-computer interaction, and healthcare, where people's facial expressions, voice qualities, and spoken words are used to better understand it. While emotion recognition holds the power to facilitate various health problems, the main challenge emotion recognition systems face is to accurately identify hidden nuances in expressions and thus, the underlying emotions conveyed by them. The true emotions of a person may remain concealed or not properly identified when only one mode of input is analyzed, therefore, multimodal streams of inputs are used to provide a more holistic view of a person's emotions. In this paper, a novel framework that fuses the results of two uni-modal methods of emotion recognition, audio, and text, to develop a robust and versatile emotion recognition system is proposed. The results show that signal processing and language processing can be utilized to reliably detect emotion from audio and text, with an accuracy of 96% and 94.1% respectively. Further, the approach presented in this paper can be used as a depression detection and monitoring tool to further enable mental healthcare professionals accurately detect symptoms of depression.**

***Index Terms*—Affective Computing, BERT, Deep Learning, Emotion Recognition, Multimodality**

I. Introduction

Humans have the innate ability to recognize the emotions conveyed by other humans. Emotion recognition drives interaction between people and is crucial in social environments for mutual understanding, gauging affability, and avoiding conflict. Humans are very revealing by nature, as they disclose emotions in many ways [1]. For example, even something as seemingly uninvolved as eyebrows can convey a multitude of emotions [2]. While being taken for granted due to its intrinsic nature, the potential of emotion recognition in a wide variety of applications cannot be understated.

Emotion recognition is a challenge being undertaken by a plethora of multidisciplinary researchers owing to its relevance in psychology, human-computer interaction, affective computing, and healthcare [3]. Recent developments in computer vision, natural language processing, and signal processing have brought new life to these efforts, which can be ultimately attributed to the rise of deep learning. Many researchers exploring emotion recognition have been trying to answer the same question: can machines achieve or even surpass human-level emotion recognition?

Human speech is a complex bundle of data comprising multiple modalities, each having its own features, making it a very widely used data medium in the deep learning community, especially in emotion recognition [4]. From speech data, the words spoken by a person and the innate features of their voice can be extracted. These modalities prove to be descriptive and reliable indicators of emotions when processed by deep learning techniques, especially when used in tandem.

The words spoken by a person can be stored as text. Thereafter, analyzing the emotion in the text becomes a language-processing problem. This text, however, is not very realistic when it comes to the recognition of a person's emotion in real-life situations [5]. For instance, when a person uses sarcasm, they may use words that convey a completely different picture when another context is not considered. Therefore, while text data is a powerful medium for predicting emotions, it suffers from many problems that leave much to be desired for accurate emotion recognition. To move towards more realistic and accurate emotion recognition, relevant additional context must be considered by intelligent systems.

The audio signals within speech data can be analyzed to recognize the speaker's emotions. While spoken words are the most direct indicators of emotions, there is more to the picture. This can be clearly seen in the case of affective animals like dogs, who consider the underlying tone of human speech and put less weight on the words spoken [6]. This is a testament to the natural instincts of creatures when it comes to sound. However, acoustic data does not come with its shortcomings, as it has a high probability of noisy data and is easily skewed by the innate voice qualities of people.

The inadequacies of the aforementioned modalities make it apparent that one single modality is not enough to represent this complex real-world problem. The combination of multiple modalities, however, may allow them to account for each other's limitations. Furthermore, it can be used to obtain a better view of the bigger picture, which may lead to a better solution. In the context of emotion recognition, multimodality is a powerful tool that allows intelligent systems to recognize the emotions of a person more accurately.

In this study, a multimodal approach for emotion recognition

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 01,2023 at 07:29:39 UTC from IEEE Xplore. Restrictions apply.

is proposed that predicts emotion in speech data through unimodal analysis of text and audio modalities and fusion of the results obtained. A detailed review of the recent developments in emotion recognition is presented in Section II. Section III describes our methodology and all its constituents in detail. In Section IV, the results of our experiments are presented. A discussion along with the future directions of our work is presented in Section V. Finally, Section VI concludes our study.

II. Related Work

A. Emotion Recognition

Until recently, most of the work regarding emotion recognition was based on unimodal data. This section presents a review of various multimodal approaches for emotion recognition.

A self-assessment tool was proposed by Prabhu et al. [7] that uses audio, video, and text input for emotion recognition and ultimately, depression screening. The tool uses a one-of-three scheme, and therefore if one modality indicates that the person is depressed, the person is classified as depressed.

Wu et al. [8] proposed a framework to detect depression based on recognizing emotions conveyed through audio and text inputs. This work explored the association between depression and emotion recognition, and how they influence each other. However, they did not subcategorize the severity of depression of the individual.

A novel framework, the cLSTM-MMA, was proposed by Pan et al. [9] to aid speech emotion recognition using video and text. Multimodal information from visual and textual cues helps to alleviate the ambiguity in emotion recognition via speech.

The work by Luna-Jimenez et al. [10] showed that audio and video carry enough separate but relevant details that when combined, the accuracy of emotion recognition systems is greatly improved. This could also be applicable to the text modality, however, this study did not explore it.

Peng et al. [11] proposed MSCNN-SPU-ATT for emotion recognition by extracting, attending, and fusing audio and text features. This simple system forwent the implementation of LSTM networks for computationally simpler CNNs and still achieved competitive results.

Chamishka et al. [12] proposed a novel method for extracting Bag-of-Audio-Words (BoAW) based features for six basic emotions from the IEMOCAP dataset. They proposed an RNN (Recurrent neural network) based architecture for real-time conversation emotion recognition. This architecture was able to outperform several previous state-of-the-art models and achieved a weighted accuracy of 60.87% and an unweighted accuracy of 60.97%.

Sowmya et al. [13] experimented with several machine learning models like SVM, RF and Decision Tree and an MLP (Multi-Layer-Perceptron) for distinguishing between 4 basic emotion types: happy, neutral, sad, and angry. They were able to achieve an accuracy of 85% on their dataset using these traditional approaches. However, the advantages of this model could not be validated using only these 4 classes proposed by the authors. Better results could be obtained using newer methods such as the BERT transformer.

Krishna et al. [14] experimented with various traditional machine learning models over 8 classes: happy, sad, angry, fearful, disgusted, calm, neutral, and surprised. They achieved an accuracy of 86.5% which is slightly better than the results obtained by the previous study.

Finally, Susithra et al. [15] proposed an architecture to recognize the speaker's emotions and gender. For emotion detection, the model was trained to recognize 4 basic emotion types. A simple Feed-Forward Neural Network (FNN) was used for the gender recognition block and a Convolutional Neural Network (CNN) with a good dataset was used for the emotion detection block. The proposed model obtained an accuracy of 91.46% on the gender detection part and 86% on the emotion recognition block.

B. Our Contribution

While a lot of research has been conducted on emotion recognition using audio-visual modalities, the multimodal analysis of audio and text towards the same goal remains relatively unexplored, despite the combination of these two modalities being more seamless. Therefore, this work explores a novel multimodal fusion framework that aims to account for some of the shortcomings of prior work on this topic by using a combination of custom and state-of-the-art detection architectures, using more emotion categories and equal consideration of both modalities.

III. Methodology

A. Text

For the classification of textual data, feature vectors are extracted from a sequence of words and fed into a classifier. This framework is to be used on common phrases often encountered in regular conversations between humans. Therefore, to train our model, an NLP dataset [16] containing sentence-emotion pairs was used. Training the model on such a dataset enabled us to predict the emotions conveyed by a person through their words. This dataset has six emotion categories: joy, sadness, anger, fear, love, and surprise.

For data augmentation, all the sentences in our dataset were parsed through a translation service, converting them to one of 5 random languages, and then back to English. Owing to these new data samples, the number of samples in our dataset increased from 16,000 to 96,000, a 500% increase. In addition, the samples that were translated back to English contained an assortment of synonyms in the place of the original words, enabling our model to associate a wide variety of words with their corresponding emotions, just by increasing its vocabulary. This also allowed our model to become more robust to possible syntactical and semantical errors in the people's sentences.

The invention of transformers has given new life to language processing research as they possess advanced language understanding capabilities. Bidirectional Encoder Representations (BERT) [17] is a family of transformers that have consistently

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 01,2023 at 07:29:39 UTC from IEEE Xplore. Restrictions apply.

achieved state-of-the-art performance in text classification problems since its invention. Models trained on the BERT paradigm possess a general understanding of the language, making fine-tuning much faster. For the text processing modality, the proposed framework uses the pre-trained BERT-base-uncased model from the Huggingface library. Transfer learning was implemented by fine-tuning it in our augmented text dataset. The model takes a sentence as input and recognizes the emotion conveyed by it.

B. Audio

The classification of audio data is mainly done by converting analog data into a digital format such as a mel spectrogram and then feeding the binary representation of that data into a deep learning classifier. This data consists of a multitude of valuable features that the classifier can correlate with emotion. For this study, five audio features from the audio samples were extracted:

- Chroma Short-Time Fourier Transform (Chroma STFT) [18]: Chroma features contain the harmonic and tonal information of an audio signal.
- Mel Frequency Cepstral Coefficient (MFCC) [19]: MFCCs are a set of features within an audio wave that describes the structure of a spectral envelope.
- Mel Spectrogram [20]: A mel spectrogram represents the different frequencies of an audio wave in the mel scale.
- Root Mean Square Value (RMS Value) [21]: The RMS value of a signal indicates the steady state of energy possessed by that signal. It scales with the amplitude of the signal.
- Zero Crossing Rate (ZCR) [22]: The ZCR of a signal is used as an indicator of the smoothness of the audio signal.

To generalize our model by using a wide variety of data, five different speech datasets were combined to train our audio classification model. The combined dataset has seven emotion categories: angry, disgust, fear, happy, neutral, sad, and surprise. Table I presents a summary of the datasets used.

TABLE I
COMPARATIVE ANALYSIS OF THE METHODS IN TERMS OF ACCURACY

Dataset	No. of actors	No. of utterances
Crema-D [23]	91	7442
EmoDB[24]	10	535
RAVDESS[25]	24	1440
SAVEE[26]	4	480
TESS[27]	2	2800

In real-world applications, background noise or disturbances accompany audio recordings. To account for realistic acoustic conditions, the noise was added to the original audio sample, and the pitch and speed of the audio recording were further modified. Additionally, time shift was performed on the audio wave. These modifications also accomplished data augmentation, producing additional data samples to train the model. In total, 48,649 samples were used to train the model.

Convolutional Neural Networks (CNN) are a subclass of deep learning architectures that can be applied to a wide variety of data. CNNs are used to recognize the underlying patterns in sequential data, and thus are very useful in computer vision and signal processing. For the classification of audio, our framework uses a 1-D CNN to process the features present in audio data. The model takes in audio features and predicts the emotion conveyed by them.

The proposed audio emotion detection model is 21 layers deep and consists of a total of 6 types of layers namely: Convolution1d layers for feature extraction, batch normalization layers to normalize the outputs before passing to the next layer, MaxPooling1D for downsampling the feature maps, dropout layer for preventing overfitting of the model, flatten layer for flattening the final output and the dense layer for getting the final output. The model is a comparatively lightweight model when compared to other state-of-the-art architectures and has only 1,674,375 trainable parameters. The inference time of the model is also almost real-time, making it perfect for the task. The Adam optimizer with a learning rate of 0.001 was used to train the model and the loss function used was categorical cross-entropy loss. Figure 1 describes the architecture of our audio classification system.

While text and audio are powerful modalities that convey emotions in various ways, they can individually fall short of capturing the entire picture. This is especially true when dealing with emotion recognition, which is a complex and multifaceted real-world problem. One solution to overcome this problem is multimodality: the combination of multiple streams of data. Multimodality provides rich training signals that help to make models more robust.

To overcome the respective limitations of the text and audio modalities, a multimodal emotion recognition framework was developed. The principle behind this idea is that the audio and text modalities would complement each other by providing different data that accounts for each other's shortcomings. Our framework processes audio and text by using separate classification models and then fuses the results obtained from them in a weighted average model to obtain the final emotion class.

The working process of our framework is described below:

- Speech data: Speech data is fed to our framework as input. The speech data is stored in the form of .wav files and uploaded to the AWS cloud for storage in an AWS S3 bucket.
- Preprocessing: For the text classification system, AWS transcribe, a speech-to-text service, is used to process speech data. Transcribe takes the speech data file as input and reproduces the words spoken in the speech file in the form of text sentences. Thereafter, the feature vectors from the text sentence(s) are extracted by the BERT tokenizer.

  The speech file is directly fed to the audio feature extractor that uses the Librosa library to extract various audio features (Chroma STFT, MFCC, Mel Spectrogram, RMS Value, ZCR) from the speech data.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 01,2023 at 07:29:39 UTC from IEEE Xplore. Restrictions apply.

- Processing: The text feature vectors are fed into our BERT classifier that predicts the underlying emotion in the given text. Similarly, a 1-D CNN is used to process the audio features and produces the emotion conveyed by them. The outputs from both of these systems are obtained individually in the form of probability distributions that represent the probability of each emotion.

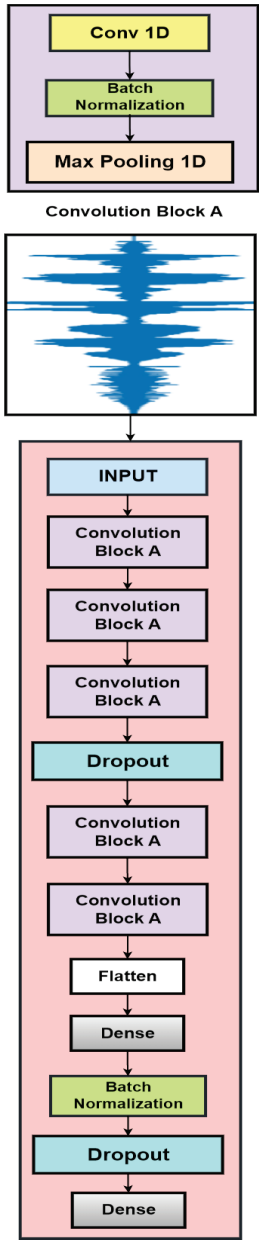


Fig. 1. Architecture of our audio classification CNN.

- Fusion: The outputs obtained from the classification models are fused by using late fusion. A weighted average model aggregates both inputs and produces one final output – the emotion most probably conveyed in the initial speech data.

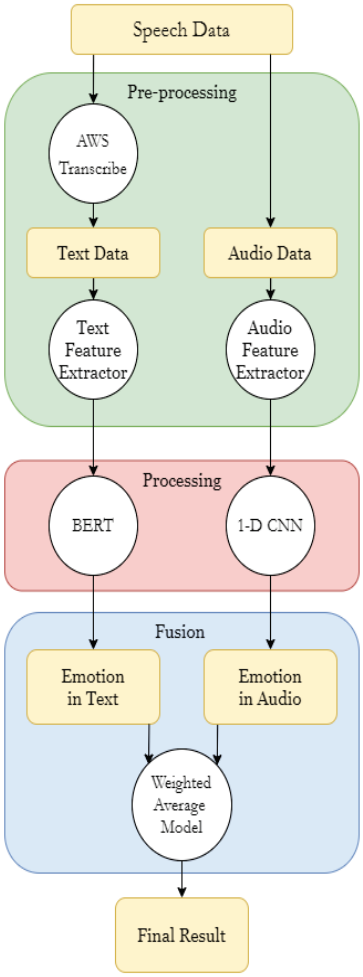


Fig. 2. Flow diagram of our proposed framework.

Figure 2 represents the flow diagram of our proposed framework which combines the text and the audio modalities using a weighted average metric.

IV. Results

A. Text

The text classification system, which uses a BERT transformer, succeeded at classifying emotions with an accuracy of 94.1% and an F-1 score of 0.94. The augmentation and preprocessing of the text dataset improved on the previously obtained accuracy of 91.8%. The model was trained with a learning rate of 0.00002 and a weight decay of 0.01. The batch size for the training was chosen as 8 based on several experiments. The proposed model outperformed most previous state-of-the-art models by a significant margin and employs a fairly novel architecture.

According to our experiments, the pre-trained BERT-base-uncased model performed better than the smaller language models like the Naive Bayes Classifier. BERT transformer is a very powerful and large-weight model, therefore it works

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 01,2023 at 07:29:39 UTC from IEEE Xplore. Restrictions apply.

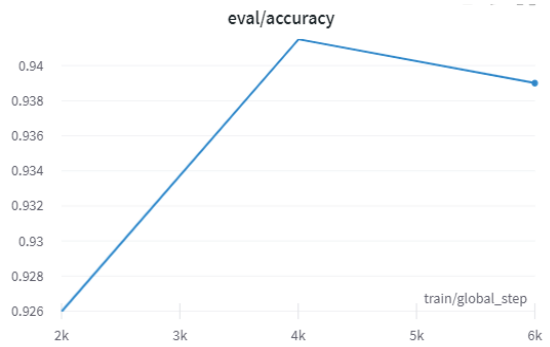


Fig. 3. BERT training curve

great in almost all NLP scenarios but suffers from a huge training time. As it can be inferred from Figure 3, two epochs were enough to make the model converge, owing to the large number of samples in our dataset. Any further epochs caused overfitting, resulting in lowered evaluation accuracy. Figure 4 presents the normalized confusion matrix for the text classification categories.

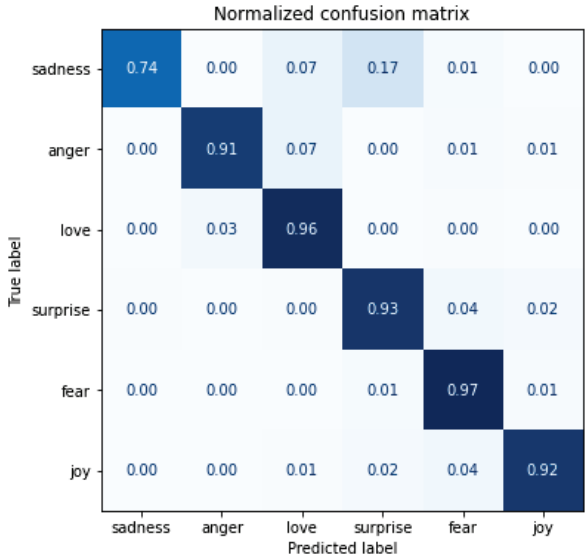


Fig. 4. Normalized confusion matrix of the text modality.

B. Audio

The proposed audio classification model (1D CNN) achieved a validation accuracy of 96% and a validation loss of 0.1095 with an F-1 score of 0.964. Also, the total precision and recall for the model were 0.965 and 0.964 respectively. The model was trained for a total of 250 epochs and reached its top accuracy within the first 180 epochs. As shown by Figure 5, the model did not have any overfitting, and it converged quite smoothly with minimal fluctuations.

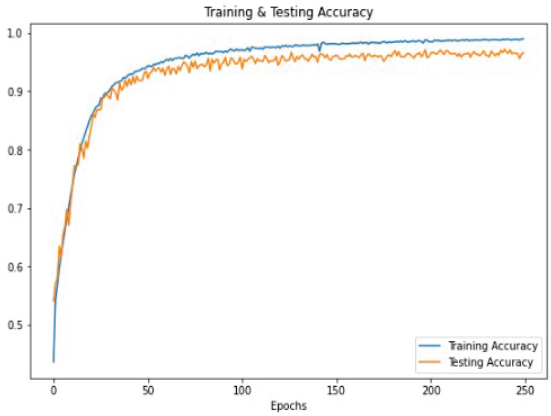


Fig. 5. CNN training curve

TABLE II
CLASSIFICATION REPORT FOR AUDIO (1D CNN)

Class Name	Precision	Recall	F1 Score
Angry	0.96	0.98	0.97
Disgust	0.96	0.95	0.95
Fear	0.94	0.96	0.95
Happy	0.97	0.95	0.96
Neutral	0.98	0.96	0.97
Sad	0.96	0.97	0.96
Surprise	0.99	0.98	0.99
Total	0.965	0.964	0.964

Figure 6 and Table II present the confusion matrix and the classification report respectively for the audio classification categories. It can be inferred from the table that the model performed well over all classes with a maximum F-1 score for surprise (0.99) and a minimum F-1 score for disgust and fear (0.95). The proposed model classified all 7 classes accurately with minimal errors.

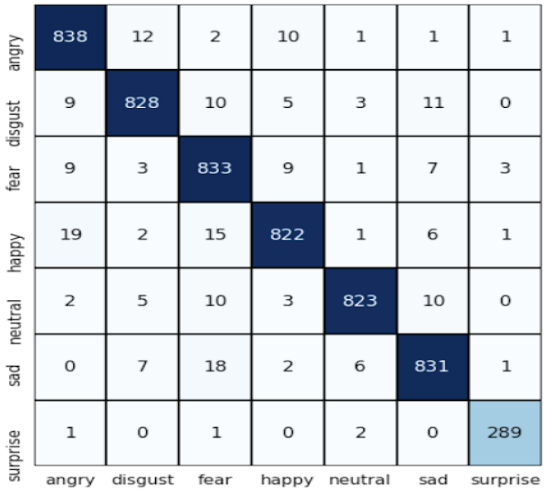


Fig. 6. Confusion matrix of the audio modality.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 01,2023 at 07:29:39 UTC from IEEE Xplore. Restrictions apply.

Moreover, the training time for the 1D CNN model was under 85 seconds for training 41350 audio samples which is minimal when compared to other state-of-the-art architectures. This makes it a reliable, lightweight and low latency model choice for the task of audio emotion classification, and perfect choice to be used with BERT as its text-emotion detection counterpart is a heavyweight model and likely to take up more computation power and processing time.

V. Discussion and Future Work

There is a plethora of emotion recognition datasets available for researchers. The proposed framework is meant to be used on real-life speech dialogues, therefore, it is trained using the best-suited datasets, for the training of the audio classification system as well as the text classification system.

In the text dataset, the number of samples for the emotion 'joy' outnumbers the other emotion categories by a significant margin. To address this problem, multiple datasets can be combined to balance the number of samples in each emotion class. Furthermore, while BERT is very good for text classification tasks, newer text classification models, some from the BERT family, have been engineered to excel at sentiment analysis. Experimentation with these specific models may lead to even better results.

The audio classification part of our framework uses a simple CNN, as in our experiments, heavier models tended to overgeneralize data points and lead to less accuracy. However, due to the large number of data samples used in our experiments, a simple CNN might not be able to effectively capture the scope of the entire data. Therefore, further experimentation might reveal an optimal CNN size that addresses both problems outlined above. In addition, some audio features may increase the complexity of the data but may lead to better results for the audio classification task.

The proposed framework uses more than one modality to achieve realistic emotion recognition, a combination of text and audio modalities. However, introducing more modalities to our framework might produce an even more accurate depiction of a person's true feelings. A vision-based emotion recognition system that operates by using facial movement markers to detect emotions, will make this framework even better.

Persistent pessimistic emotions are considered a symptom of depression. As the next step, this emotion recognition framework can be utilized as an early diagnosis method for depression. Expert advice from mental health professionals, linking the relationship between emotions and depression, will be consulted in order to facilitate this transition. The proposed diagnosis system would be later integrated with a counselling chatbot, in order to achieve our goal of building a private, convenient and non-invasive method for depression detection and therapy.

VI. Conclusion

Emotion recognition is a very delicate yet complex field. While it is useful in many fields such as healthcare, psychology, and human-computer interaction, most existing work on this subject is straightforward and does not consider the multitude of complications and abnormalities that may prevent accurate prediction in real-life scenarios.

This study proposes a speech emotion recognition framework by combining inferences from two different modalities: audio and text. The words spoken and the audio signals within speech data are processed by the framework to obtain the underlying emotion conveyed by them. Various augmentation and preprocessing methods were applied to the datasets in order to train our framework for realistic situations. Our framework uses separate classification systems for predicting emotions in text and audio. The text classification system uses the BERT model, and the audio classification system uses a 1-D CNN model. The text and audio models achieved accuracies of 94.1% and 96%, and F-1 scores of 0.94 and 0.96 respectively. The results obtained from these models are then fed into a weighted average model that combines both of them to produce the most probable emotion conveyed in the given speech data as the result.

Finally, in future work, this emotion recognition system can be translated into a diagnosis system for depression by analyzing the connection between emotions and depression.

References

[1] Schirmer, Annett, and Ralph Adolphs. "Emotion perception from face, voice, and touch: comparisons and convergence." Trends in cognitive sciences 21.3 (2017): 216-228.

[2] Barrett, Lisa Feldman, et al. "Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements." Psychological science in the public interest 20.1 (2019): 1-68.

[3] Vinola, C., and K. Vimaladevi. "A survey on human emotion recognition approaches, databases and applications." ELCVIA: electronic letters on computer vision and image analysis (2015): 00024-44.

[4] Akçay, Mehmet Berkehan, and Kaya Oğuz. "Speech emotion recognition: Emotional models, databases, features, preprocessing methods, supporting modalities, and classifiers." Speech Communication 116 (2020): 56-76.

[5] Boukes, Mark, et al. "What's the tone? Easy doesn't do it: Analyzing performance and agreement between off-the-shelf sentiment analysis tools." Communication Methods and Measures 14.2 (2020): 83-104.

[6] Albuquerque, Natalia, et al. "Dogs recognize dog and human emotions." Biology letters 12.1 (2016): 20150883.

[7] Prabhu, Sahana, et al. "Harnessing emotions for depression detection." Pattern Analysis and Applications 25.3 (2022): 537-547.

[8] Wu, Wen, Mengyue Wu, and Kai Yu. "Climate and Weather: Inspecting Depression Detection via Emotion Recognition." ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2022.

[9] Pan, Zexu, et al. "Multi-modal attention for speech emotion recognition." arXiv preprint arXiv:2009.04107 (2020).

[10] Luna-Jiménez, Cristina, et al. "Multimodal emotion recognition on RAVDESS dataset using transfer learning." Sensors 21.22 (2021): 7665.

[11] Peng, Zixuan, et al. "Efficient speech emotion recognition using multi-scale cnn and attention." ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2021.

[12] Chamishka, Sadil, et al. "A voice-based real-time emotion detection technique using recurrent neural network empowered feature modelling." Multimedia Tools and Applications 81.24 (2022): 35173-35194.

[13] Sowmya, G., et al. "Speech2Emotion: Intensifying Emotion Detection Using MLP through RAVDESS Dataset." 2022 International Conference on Electronics and Renewable Systems (ICEARS). IEEE, 2022.

[14] Vamshi, Krishna B., Ajeet Kumar Pandey, and Kumar AP Siva. "Topic model based opinion mining and sentiment analysis." 2018 International Conference on Computer Communication and Informatics (ICCCI). IEEE, 2018.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 01,2023 at 07:29:39 UTC from IEEE Xplore. Restrictions apply.

[15] Susithra, N., et al. ”Speech based Emotion Recognition and Gender Identification using FNN and CNN Models.” 2022 3rd International Conference for Emerging Technology (INCET). IEEE, 2022.
[16] Praveen (2019). Emotions dataset for NLP, Version 1. Retrieved August 21, 2022 from https://www.kaggle.com/datasets/praveengovi/emotions-dataset-for-nlp.
[17] Devlin, Jacob, et al. ”Bert: Pre-training of deep bidirectional transformers for language understanding.” arXiv preprint arXiv:1810.04805 (2018).
[18] Müller, Meinard. ”Short-Time Fourier Transform and Chroma Features.” Lab Course, Friedrich-Alexander-Universität Erlangen-Nürnberg (2015).
[19] Likitha, M. S., et al. ”Speech based human emotion recognition using MFCC.” 2017 international conference on wireless communications, signal processing and networking (WiSPNET). IEEE, 2017.
[20] Yenigalla, Promod, et al. ”Speech Emotion Recognition Using Spectrogram Phoneme Embedding.” Interspeech. Vol. 2018. 2018.
[21] Bhat, Aathreya S., et al. ”An efficient classification algorithm for music mood detection in western and hindi music using audio feature extraction.” 2014 fifth international conference on signal and image processing. IEEE, 2014.
[22] Aouani, Hadhami, and Yassine Ben Ayed. ”Speech emotion recognition with deep learning.” Procedia Computer Science 176 (2020): 251-260.
[23] Cao, Houwei, et al. ”Crema-d: Crowd-sourced emotional multimodal actors dataset.” IEEE transactions on affective computing 5.4 (2014): 377-390.
[24] Burkhardt, Felix, et al. ”A database of German emotional speech.” Interspeech. Vol. 5. 2005.
[25] Livingstone, Steven R., and Frank A. Russo. ”The Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS): A dynamic, multimodal set of facial and vocal expressions in North American English.” PloS one 13.5 (2018): e0196391.
[26] Jackson, Philip, and SJUoSG Haq. ”Surrey audio-visual expressed emotion (savee) database.” University of Surrey: Guildford, UK (2014).
[27] K. Dupuis and M. K. Pichora-Fuller, Toronto Emotional Speech Set (TESS). University of Toronto, Psychology Department, 2010

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 01,2023 at 07:29:39 UTC from IEEE Xplore. Restrictions apply.

Check for updates

Empirical Review of Various Thermography-Based Computer-Aided Diagnostic Systems for Multiple Diseases

TRASHA GUPTA, Delhi Technological University, India
RAJNI JINDAL, Delhi Technological University, India
INDU SREEDEVI*, Delhi Technological University, India

The lifestyle led by today's generation and its negligence towards health is highly susceptible to various diseases. Developing countries are at a higher risk of mortality due to late-stage presentation, inaccessible diagnosis, and high-cost treatment. Thermography-based technology, aided with machine learning, for screening inflammation in the human body is non-invasive and cost-wise appropriate. It requires very little equipment, especially in rural areas with limited facilities. Recently, Thermography based monitoring has been deployed worldwide at various organizations and public gathering points as a first measure of screening COVID-19 patients. In this paper, we systematically compare the state-of-the-art feature extraction approaches for analyzing thermal patterns in the human body, individually and in combination, on a platform, using three publicly available Datasets of medical thermal imaging, four Feature Selection methods, and four well-known Classifiers, and analyze the results. We developed and used a 2-level sampling method for training and testing the classification model. Among all the combinations considered, the classification model with Unified Feature-Sets gave the best performance for all the datasets. Also, the experimental results show that the classification accuracy improves considerably with the use of feature selection methods. We obtained the best performance with a features subset of 45, 57, and 39 features (from Unified Feature Set) with a combination of mRMR and SVM for DB-DMR-IR and DB-FOOT-IR and a combination of ReF and RF for DB-THY-IR. Also, we found that for all the feature subsets, the features obtained are relevant, non-redundant, and distinguish normal and abnormal thermal patterns with the accuracy of 94.75% on the DB-DMR-IR dataset, 93.14% on the DB-FOOT-IR dataset, and 92.06% on the DB-THY-IR dataset.

Additional Key Words and Phrases: Breast cancer, Medical Thermography, Diabetes, Thyroid cancer, Feature Extraction, Classification, Feature Selection

1 INTRODUCTION

Today's generation is developing habits that neglect health and has become susceptible to various diseases. Late-stage presentation, inaccessible diagnosis, and treatment contribute equally to deaths, especially in developing nations where access to state-of-the-art diagnostic medical infrastructure is challenging[1]. Also, medical imaging (like Computerized Tomography (CT) scan, Ultrasound, Magnetic Resonance Imaging (MRI), etc.) in combination with pathological tests form the basis for all medical conclusions. However, they may not be affordable or accessible to many people. Also, these medical imaging modalities have some disadvantages. They use ionizing

*All the authors contributed equally to this research.

[1]https://www.who.int/news/item/13-12-2019-world-bank-and-who-half-the-world-lacks-access-to-essential-health-services-100-million-still-pushed-into-extreme-poverty-because-of-health-expenses

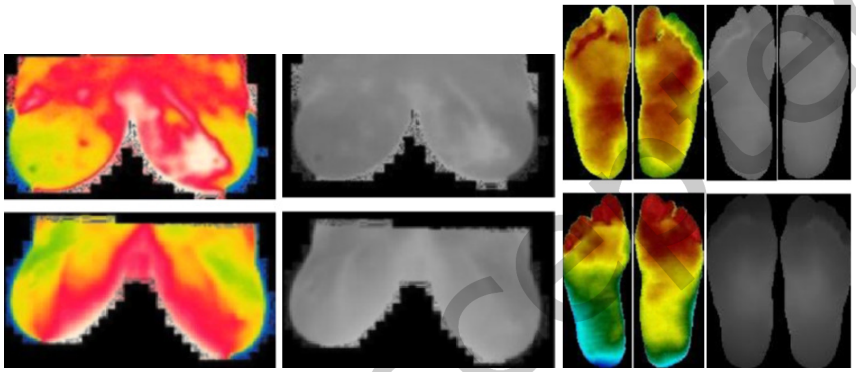
Authors' addresses: Trasha Gupta, trashagupta@dtu.ac.in, Delhi Technological University, India; Rajni Jindal, rajnijindal@dce.ac.in, Delhi Technological University, India; Indu Sreedevi, s.indu@dce.ac.in, Delhi Technological University, India.

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.
© 2023 Copyright held by the owner/author(s). Publication rights licensed to ACM.
2157-6904/2023/2-ART $15.00
https://doi.org/10.1145/3583778

Just Accepted

radiation, which is harmful to the human body[1, 46]. These devices require unique installation and a complicated process, limiting the number of possible tests per equipment.

An adjunct modality for medically screening human body[2] exists - Digital Infrared Thermal Imaging technology (DITI) (also called as Thermography-based technology). DITI states the physical information of behaviour in the human body [28]. Abnormality in the body causes inflammation (indicated by Erythrocyte Sedimentation Rate (ESR) pathological test[11, 29, 56]), causing intense thermal patterns radiated by the corresponding body part[6, 33]. Thermography-based cameras can capture these thermal patterns and use them as biomarkers for identifying the abnormal regions [5, 16, 34]. Recently, Thermography based monitoring has been deployed around the world[3] at various organizations and public gathering points as a first measure of screening COVID-19 patients [12, 31, 55]. Figure 1 illustrates the examples of normal and abnormal Thermograms and corresponding normalized grayscale images for various diseases.



(a) Breast Cancer Grayscale and Thermal Images

(b) Diabetic Feet Ulcer Grayscale and Thermal Images

Fig. 1. Example of abnormal and normal Thermograms for various diseases.

Thermography-based technology for screening/diagnosis is effortless and straightforward to use. Also, it does not use any harmful (invasive or ionizing) radiation[20, 47]. Hence, it is safe for both technician and the patient. It is cost-wise appropriate, painless, and a contactless modality. It requires very little equipment and is ideal for deployment/use in rural areas with limited facilities. The figure 2[4] illustrates the ease in setup and capturing thermal images. Also, Thermal Imaging complemented with Computer-Aided Diagnosis (CAD) has been studied thoroughly and proposed in the literature for detecting/screening various diseases like breast cancer, skin cancer, arthritis, cardiovascular diseases, ulcers, and diabetes using retinal thermal images or feet[9, 42, 57]. Smartphones combined with such technology will provide a huge potential in Telemedicine and better healthcare delivery, especially in remote/rural areas.

Many researchers have explored Thermal Imaging based CAD extensively for diagnosing various diseases in the last two decades [9, 16, 17]. Different research works have proposed their features extraction methods for identifying abnormal thermal patterns [1, 5, 15, 25, 30, 49–52]. Also, researchers have used different classifiers

[2]https://www.fda.gov/consumers/consumer-updates/breast-cancer-screening-thermogram-no-substitute-mammogram

[3]https://www.fda.gov/medical-devices/general-hospital-devices-and-supplies/thermal-imaging-systems-infrared-thermographic-systems-thermal-imaging-cameras

[4]https://www.fda.gov/medical-devices/general-hospital-devices-and-supplies/thermal-imaging-systems-infrared-thermographic-systems-thermal-imaging-cameras

Just Accepted

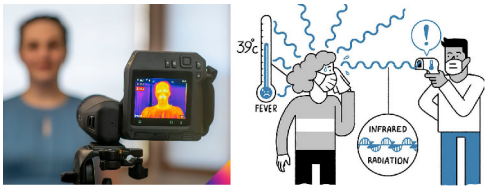


Fig. 2. Illustration of setup and capturing of thermal images.

and Cross-Validation strategies to evaluate the proposed model. A few research work [15, 25, 51] have evaluated their method on a publicly available dataset. However, most of them [1, 5, 30, 49, 50, 52] evaluated their proposed features on the respective/private datasets, which is very small in size and not available for further evaluation by the research community. To the best of our knowledge, no experimental comparison of the state-of-the-art methods exists on a platform using a common dataset, classifier, or Cross-Validation strategy. Hence, it is not appropriate to comment on the suitability of a feature extraction method for diagnosing abnormality/inflammation in the human body using Thermography. Also, it has been observed that feature selection methods are recommended when the dataset size (i.e. the number of instances) is smaller than the number of features. Evaluation of a small dataset with a large number of features can lead to degradation in the model's performance. To our knowledge, limited work is done towards exploring relevant features using feature selection techniques [1, 5, 51]. Most of the research works [3, 4, 15, 25, 41, 43, 49, 50, 52] have not utilized feature selection methods to find a subset of relevant features from thermography images. Therefore, there is a need to investigate feature selection methods to overcome the problem of a small sample size dataset. With this, we will compare whether the model's performance improves.

Due to the limitation of unavailability of large medical thermal imaging-based datasets, we have not explored Deep Learning-based models. The key contributions of the paper are listed below as:

- We performed a careful and systematic review of the various state-of-the-art feature extraction methods proposed in the literature for analyzing medical thermal images and highlighted their limitations.
- Then, we statistically compared and analyzed the performance of various feature extraction methods proposed in the literature to distinguish an abnormal thermal image from a normal one on a uniform platform.
- We investigated the above objective using four well-known feature selection methods to select relevant features from the features-set proposed in the literature. Also, we used four well-known classifiers for a fair performance comparison of the selected set of relevant features. We used three publicly available datasets of Medical Thermal Imaging for the entire investigation, having different diseases. The publicly available datasets are skewed (i.e. difference in classes frequency in the dataset is high), so we developed and used a 2-level sampling strategy to conduct the experiments (section 3.4).
- We also propose to develop a unified model that identifies the abnormal thermal patterns/inflammation (due to some disease) radiated by the human body and captured by a thermal camera. We evaluated this model on the same platform.
- We also developed a robust model ranking mechanism to suggest a suitable model for distinguishing abnormal and normal thermal biomarkers.

This paper is organized as follows. In section 2, we discussed the related work concerning the applications of DITI for abnormality/inflammation detection. The sections 3.1, 3.2, and 3.3 explain the procedure (feature extraction, feature selection, and classification) followed in the experiments. The experimental framework is described in section 4. The results are stated, analyzed, and concluded in sections 5 and 6, respectively.

Just Accepted

2 RELATED WORK

Digital Infrared technology was first used and documented in medical research for examining breast cancer in 1956 [36]. In the early '70s, a mass-ranging study, Breast Cancer Detection and Demonstration Project (BCDDP), was conducted in the United States to evaluate thermography's diagnostic ability for Breast Cancer. However, poor study design, lack of trained technicians, and unsuitable environmental controls and protocols led to the failure of the project [7, 16]. The research work [22] presented a study on 85,000 subjects with 90% sensitivity and 88% specificity. In their investigation, to attain thermal equilibrium, they included two cooling mechanisms-fan and ice water. This methodology decreased the false-positive rate by 3.5%, leading to 96.5% sensitivity. The potential of thermography has also been studied concerning inflammatory pain[18, 19].

The method proposed in the research work [52] extracted 37 asymmetry-based descriptors from a skewed and private dataset of 146 Thermograms (29 malignant and 117 benign breast thermograms) and classified them using a fuzzy classifier. A classification accuracy of 79.53% was obtained with 14 fuzzy partitions per attribute on the test dataset. The research work [48] used Hough transform to automatically segment left and right breasts and performed an asymmetry-based analysis using first-order statistics on a small dataset of 6 normal and 18 cancerous thermograms. They concluded that higher-order statistics detect asymmetry effectively. Texture-based features [1], were extracted from grey-scale breast thermogram's co-occurrence matrix and run-length matrix. These features were extracted from 50 samples and fed to the Support Vector Machine (SVM) classifier with a 3-fold Cross-Validation sampling approach. The proposed system produced an accuracy of 88.10%, sensitivity, and specificity of 85.71% and 90.48%, respectively, on their private dataset.

The research work [49] proposed 11 energy-based features for asymmetry analysis. They computed features from modified Local Binary Pattern (LBP) Colour and Edge matrices. These matrices were computed from grayscale images of left and right breasts. They used a heterogeneous dataset of 65 breast thermography images and classified the extracted features using Hidden Markov Model (HMM) Classifier. They achieved an accuracy of 93.75%. The research work [51] extracted 144 features from Wavelet Transform of normalized and segmented temperature matrices of breast thermogram. Then asymmetry-based features were computed from the Local energy matrices for 24 sub-bands. They used a publicly available DMR-IR dataset [54] with 100 frontal views of static breast thermograms. Then they used Random Subspace Feature selection (RSFS) for selecting relevant features, followed by 5-fold Cross-Validation with SVM using Gaussian Kernel. They achieved an accuracy of 91%.

The research work [23] extracted asymmetry based 24 discriminating features - 7 statistical and 17 texture features from two datasets (DBT-TU-JU- a private dataset and DB-DMR-IR dataset [54]). Then, Mann-Whitney-Wilcoxon statistical test was used to select significant features (with $p \leq 0.05$). The authors used six classifiers to detect thermography-based breast irregularity for efficient diagnosis. The results showed that Artificial Neural Network and SVM classifiers provided the best accuracy of 87.50% for the DB-DMR-IR database and 84.29% for the DBT-TU-JU dataset. The research works [21, 30] applied Curvelet Transform to segmented breast thermograms and extracted GLCM-based features. [21] obtained an accuracy rate of 90.91% with SVM classifier and Gaussian kernel with Leave-one-Out Cross-Validation (LOOCV). An accuracy of 93.33% using SVM with 20-fold cross-validation was obtained[30].

The authors [3] used thermography complemented with other imaging modalities to diagnose thyroid-related problems. The work [58] employed an Otsu-thresholding technique to segment the thyroid-based ROI from the rest of the images. The authors in the research work [5] analyzed Rheumatoid Arthritis disease. They extracted Texture based Features from the Knee thermogram and inflamed ROI (automatically segmented from Knee Thermogram). They obtained an accuracy of 73% using SVM with 3-fold Cross-validation. The authors [25] performed an asymmetry-based analysis of breasts, foot-plantar, and thyroids thermal images to identify the abnormal thermal patterns for each disease, respectively. They obtained an accuracy of 91.42%, 93.47%, and 88.57%

Just Accepted

for correctly identifying breast cancer-based, diabetes-based, and thyroid cancer-based thermal biomarkers, respectively.

The research work [53] performed experiments on average temperatures values for neuropathy patients Vs. Diabetic patients without neuropathy. They concluded that the former has a higher average temperature than the latter. A study [2, 4] also confirmed the above remark that diabetic neuropathy patients' foot temperatures were significantly higher than that of subjects without neuropathy. Infrared thermography has been successfully used to detect complications like foot ulcers related to diabetic feet. The researchers [17] segmented the abnormal regions with the help of Lazy-Snapping method. They extracted the abnormal regions for various diseases in the human body using this interactive segmentation technique. They obtained a sensitivity of 70%. The authors [50] extracted asymmetric 12 texture and 2 temperature-based features from 11 regions of interest, which are prone to ulceration in both left and right foot, thermal images, and temperature profiles of 60 subjects. They performed an asymmetry-based analysis between the ipsilateral and contralateral regions of the foot. The features are classified into normal and ulcerated using SVM classifier with an accuracy of 95.61%, sensitivity of 96.5%, and specificity of 92.41%. In most diabetic patients, symptoms of ulcers are not observed until the ulcer is infected [41, 43].

Table 1 summarizes the research work done to diagnose abnormality/inflammation due to various diseases using thermography. The table lists the datasets used, the classification model, and the best results obtained in the respective papers.

From the above discussion, we found that numerous feature extraction methods have been proposed but not compared with the existing work in the literature. Researchers have claimed the results on their private datasets. To the best of our knowledge, no experimental comparison of the state-of-the-art methods exists on a platform using a common dataset, classifier, or Cross-Validation strategy. Also, datasets used for analysis are very small or skewed in most research works. Hence, it is not appropriate to comment on the suitability of a feature extraction method for detecting abnormal thermal patterns due to various diseases. Hence, we lay one of the objectives of comparing the performance of various feature extraction methods to distinguish an abnormal sample from a normal one, individually and in combination, on the uniform platform and analyze the results statistically.

3 METHODOLOGY

In this work, we will perform a systematic and empirical review of the features proposed in the state-of-the-art approaches [1, 5, 15, 25, 30, 49–52]. We will implement and extract the features of the feature extraction methods mentioned in table 1. We will also evaluate the performance when all the features mentioned in table 1 are combined for classification. We empirically evaluate their performance on a uniform platform for a fair comparison.

We define the uniform platform as a Machine Learning (ML) pipeline of all possible combinations of four well-known feature selection strategies (Pearson Correlation (PC), Minimum Redundancy Maximum Relevance (mRMR), ReliefF (ReF), and Chi-Squared (CS)), four classifiers (K-Nearest Neighbour (k-NN), Support Vector Machine (SVM), Decision Trees (DT) and Random Forest (RF)) on three publicly available Thermal imaging based medical datasets (DB-DMR-IR, DB-FOOT-IR, and DB-THY-IR). Also, we will assess each pipeline using a 2-level sampling strategy. With this, we will evaluate the strength of feature selection methods, classifiers, and 2-level sampling strategy for the various state-of-the-art approaches for detecting abnormality characterized by inflammation in human body. This section describes and contrasts the components of the uniform platform used for comparing various feature extraction approaches.

3.1 Features

In this paper, we implemented the feature extraction methods proposed in state-of-the-art, mentioned in table 1, with the exact line of detail on three publicly available datasets. The features generated were normalized

Just Accepted

Table 1. The State-Of-The-Art research work considered for the empirical review.

Reference	Number of features	Features Description	Dataset and Disease Considered	Feature Selection	Classifier	Metrics and Result on Test data
[52] **(SCH)**	37	Features were computed from polar coordinate representation of left and bright grayscale images - Mean, Standard Deviation, Median, 90-Percentile, Moments m01 and m10, Centre of Gravity (CoG), Distance between CoG and Geometrical Centre, Cross-Correlation, Difference Histogram, Absolute Value of its Maximum, Number of Bins Exceeding 0.01, number of zero crossings, energy, difference of positive and negative (from Normalised Histograms) Homogeneity, Energy, Contrast, Symmetry (from Gray Level Cross co-occurrence matrix), Difference of absolute values, Distance of difference maximum from the centre (from Fourier Spectrum of left and right breast), Mutual Information, Cross Co-occurrence Features, Mutual Information, Fourier Descriptors from the Laplace filter applied images	Breast Cancer analysis - Private Dataset - 146 cases (29 Malignant, 117 Benign)	-	10-fold cross-validation with Fuzzy Classifier (14 fuzzy partitions) using KEEL software	Acc - 79.53, Sens - 79.86, Spec- 79.49
[1] **(ACH)**	16	Texture based analysis - Features from Normalized GLCM and Gray Level Run-length Matrix (in directions $\theta = 0°, 45°, 90°$ and 135°) were computed - Homogeneity, Energy, Entropy, Moment 1, Moment 2, Moment 3, Moment 4, Entropy, Angular Second Moment, Contrast, Mean, Short Runs Emphasis, Long Runs Emphasis, Run Percentage, Gray Level Non-Uniformity, and Run Length Non-Uniformity	Breast Cancer analysis - Private Dataset - 50 cases (25 Normal , 25 Malignant (15 patients had stage III cancer and rest had stage II cancer)	selected with P-Values (≤ 0.05)	3-fold Stratified cross-validation using SVM	Acc - 85.71, AUC-88.10, Sens - 85.71, Spec - 90.48
[49] **(RAS)**	11	Asymmetry based analysis - Modified LBP Colour and Edge matrices were computed using Clique Matrices from Grayscale images of left and right breast - Tsallis-Entropy, GLCM based Energy, Difference between Co-occurrence levels based energy, Euclidean distance of histograms, sum (from Modified LBP Colour and Edge matrices), Difference of GLCM Matrices(4 different directions 0°, 45°, 90° and 135°)	Breast Cancer analysis - Private Dataset - 65 breast thermography images collected heterogeneously (40 Malignant and 25 Normal)	-	HMM Classifier with LOO cross-validation	Acc - 93.75, Sens - 95.00, Spec - 92.00
[51] **(SAT)**	144	asymmetry-based analysis - Temperature matrices of the corresponding Breast Thermogram were normalized and segmented automatically into left and right breast. The Discrete Wavelet Transform was applied using different digital filters (Daubechies-1, Daubechies-10, Coiflet-1, Coiflet-5, Symlet-2, and Symlet-8) at different scales. Asymmetry based Features were computed from the Local energy matrices, calculated for 24 sub-bands - Energy, Mean, Variance, Skewness, Kurtosis, Entropy (from Local energy matrices, calculated for 24 sub-bands)	Breast Cancer analysis - DMR-IR Dataset -100 frontal view static breast thermogram (47 abnormal, 53 normal)	RSFS Method - 33 features were selected	5-fold Cross Validation with SVM using Gaussian Kernel	Acc - 91, Sens - 87.23, Spec - 94.34
[15] **(CRZ)**	5	Texture based analysis - 5 features - number of pixels, mean value, variance, maximum entropy and index value - are extracted from automatically segmented ROI using threshold values (up to 4) optimized by the Differential Evolutionary method (by maximizing the total fuzzy entropy measure)	Diabetic Foot Plantar analysis - DB-Foot-Plantar Database - 167 samples (122- Abnormal and 45 Normal)	-	10-fold Cross-Validation with SVM using Linear Kernel	Acc - 99.2857
[25] **(GUP)**	101	Asymmetry Based analysis - 10 Haralick's Texture based features were extracted from the submatrix of GLCM obtained in 4 directions(0°, 45°, 90° and 135°) and then their average was taken for 4 directions. The submatrix was calculated based on the probability distribution of normal and abnormal samples - Entropy, contrast, Correlation, Energy, Sum variance, Sum entropy, Difference variance, Information measure of correlation. Homogeneity, and Sum average. 5 statistical features were calculated from 18 grids (of size 50X50) of the given image's Local Binary Pattern (LBP) - mean, variance, skewness, kurtosis, and entropy. Otsu threshold was computed from the LBP transformed matrix.	Breast Cancer analysis-DMR-IR Dataset - 81 samples (41 - Abnormal and 40 Normal), Diabetic Foot analysis - DB-FOOT-PLANTAR Dataset - 95 Samples (50 - Abnormal and 45 Normal), Thyroid Cancer analysis-TH-NOD-DB Dataset - 34 samples (18 - Abnormal and 16 Normal)	T-Test with 5% level of significance - 40 features	LOO CV with SVM using Gaussian Kernel	DMR-IR Dataset - Acc - 91.42, Sens - 94.28, Spec - 88.57, DB-FOOT-PLANTAR - Acc - 93.47, Sens - 88.57, Spec - 94.28, TH-NOD-DB Dataset-Acc - 88.57, Sens - 91.42, Spec - 85.71
[50] **(SAM)**	14	Asymmetry Based analysis- 2 Statistical features were extracted from segmented temperature profiles - Mean temperature and Standard Deviation of temperature. 12 Haralick's Texture based features were extracted from normalized GLCM (in direction 0) of the same segmented grayscale images - Auto Correlation, Contrast, Correlation, Dissimilarity, Energy, Entropy, Homogeneity, Maximum Probability, Sum of Squares- Variance, Sum Average, Sum Variance, Sum Entropy	Diabetic Foot analysis - Private dataset - 660 samples (60 subjects * 11 ROI samples from each subject) (264 Normal and 396 Abnormal samples)	-	SVM with (140 samples for training and 520 samples for Testing)	Acc-95.61, Sens-96.5, Spec-92.41
[5] **(BAR)**	119	Texture Based analysis - Various Texture based Features were extracted from Knee thermogram and inflamed ROI automatically segmented from Knee Thermogram - First order statistical features (Mean, variance, median, mode, skewness, kurtosis, energy, entropy, standard deviation), 13 Haralick features from each of Mean and difference based (energy, contrast, correlation, variance, homogeneity, sum average, sum of variance, sum entropy, entropy, difference variance, difference entropy, information measures of correlation (from x axis and y axis individually), Gray level difference statistics based features (Homogeneity, contrast, energy, entropy, mean), Neighbourhood Gray-tone Difference matrix based features (Coarseness, contrast, busyness, complexity, texture strength), Statistical feature matrix based features (Coarseness, contrast, periodicity, regularity/roughness), Texture energy measure (Texture energy using LL, EE,SS, LE,ES, LS kernel), Frequency level features from Fourier power spectrum(Radial sum, angular sum) and shape based features (Area, Euler number, perimeter, convexity, eccentricity, orientation, solidity	Rheumatoid Arthritis - Private Dataset - 110 samples of Knee thermograms (60 samples with Arthritis and 50 samples from Control group)	Accuracy based selection of Feature groups - 6 feature groups (50 features)	3-fold Cross-Validation with SVM	Acc - 73
[30] **(KAR)**	20	Texture based analysis - Images are semi-automatically segmented and enhanced using morphological operators. Then three sets of features (GLCM-based, regional-based, and Statistics based) are extracted from curvelet Transformed images - Contrast, Correlation, Energy, Homogeneity, Area, Major Axis Length, Minor Axis Length, Eccentricity, Orientation, EquivDiameter, Perimeter, Min Value, Max Value, Skewness, Kurtosis Norm Value Mean, Variance, Median, Entropy	Breast Cancer analysis - Private Dataset - 60 samples (30 positive and 30 negative samples)	P-Test with 5% level of significance - 16 Features	20-fold cross-validation with SVM (cubic kernel)	Acc- 93.33

Just Accepted

in the range [0,1]. Each feature-set suggested in the literature obtained is investigated on a common platform using a combination of four feature selection, four classification approaches, and 2-level sampling strategies for comparing their performance.

A feature set was also obtained by taking the union of all the features (stated in table 1) (duplicate features were removed), named as UnionFeature_Set. UnionFeature_Set contains 467 features (in the order as stated in table 1). This was done to analyze the behaviour of the features with each other and on the proposed platform. With this, we will investigate whether combining features complements each other to give better results and find the most relevant feature set for characterizing inflammation in human body captured using a Thermal camera.

3.2 Feature Selection (FS) Methods

The feature sets proposed/extracted in the literature may contain redundant or irrelevant features for the study. In either scenario, it may degrade the predictability of the decision system. Using more/redundant features may consume a considerable amount of memory and training time. A limited number of studies [1, 5, 51] have used the feature selection method to assess the relevance of features for identifying abnormality/inflammation in human body.

We applied four well-known FS methods to rank the extracted and normalized features. A feature is included (based on its rank) incrementally to the set until the best performance is achieved. The following FS methods are used to evaluate the relevance.

- **Pearson Correlation Score (PC) -** PC Score[44], is a uni-variate approach that selects a subset of features highly correlated with the classification. The Pearson correlation coefficient is calculated as

$$PC_Score(f_i) = \frac{cov(f_i, Y)}{\sqrt{var(f_i)}\sqrt{var(Y)}} \tag{1}$$

  where f_i is the feature whose PC score is evaluated with respect to the class label Y (output), cov is the covariance, and var defines the variance. The absolute value of correlation value closer to 1 indicates a stronger correlation among the prescribed variables, while zero value implies no correlation between the variables.
- **Chi-Square Score (ChiSq) -** ChiSq Score [37] is another univariate FS method that evaluates the association of a feature with the class label. The ChiSq score for a feature f_i, with r different values and binary classification problem, is calculated as-

$$ChiSq_Score(f_i) = \sum_{j=1}^{r}\sum_{c=1}^{2}\frac{(n_{jc} - \mu_{jc})^2}{\mu_{jc}} \tag{2}$$

$$\mu_{jc} = \frac{n_{*c}n_j}{n} \tag{3}$$

  where n_{jc} defines the number of samples with j^{th} feature of class c, n_j defines the number of samples with the j^{th} feature value, and n_{*c} denotes the number of samples in class c. A higher ChiSq score indicates high relevance of a feature and its high ability to distinguish the class.
- **ReliefF Score (ReF) -** ReF Score [35] is an instance-based and sixth variation of Relief FS family [32]. It computes the discriminating strength of a sample's feature value with its K nearest neighbours. This approach awards high weightage to the features that give diverging values to their neighbours of different classes and penalizes the ones that give different values to neighbours of the same class. It basically searches for k near misses from each class and averages their contributions for updating weights, weighted with the

Just Accepted

prior probability of each class. ReF score for a feature f_i is computed as

$$ReF_Score(f_i) = 0.5 \sum_{j=1}^{l} (-\frac{1}{K} \sum_{x_r \in NH(j)} d(x_j(i) - x_r(i)) + \sum_{y \neq y_j} \frac{p(y)}{K(1-p(y))} \sum_{x_r \in NM(j,y)} d(x_j(i) - x_r(i))) \quad (4)$$

where l is the number of samples randomly selected from n samples, $NH(j)$ and $NM(j, y)$ define the nearest samples of x_j in the same class and in other class y, respectively, both of size K. $p(y)$ defines the instances ratio in class y. Higher the score, more relevant the feature is.

- **Minimum Redundancy Maximum Relevance Score (mRMR) -** mRMR Score, is multivariate FS method [45], based on Mutual Information (MI). The relevance of features with the class and the redundancy of features with each other is considered simultaneously. The mRMR algorithm measures the redundancy by taking the MI of features pairwise. Simultaneously, it quantifies relevance using the MI of attribute A with class variable Y. It selects and ranks the feature according to the Mutual Information Quotient (MIQ) value in the forward selection scheme. The MIQ value is computed as follows:

$$MIQ_A = \frac{V_A}{W_A} \quad (5)$$

where V_A and W_A are the relevance and redundancy, respectively, and computed as:

$$V_A = MI(A, Y) \quad (6)$$

and

$$W_A = \frac{1}{|S|} \sum_{Z \in S} MI(A, Z). \quad (7)$$

$MI(\cdot)$, between two variables A and B, defines the non-linear correlation between two variables and is computed as

$$MI(A, B) = \sum_{i,j} P(X = x_i, Z = z_j) \log \frac{P(X = x_i, Z = z_j)}{P(X = x_i)P(Z = z_j)} \quad (8)$$

We have implemented three univariate and one multivariate state-of-the-art FS methods. Univariate feature selection methods evaluate and rank a single feature concerning the class variable. PC is a low-cost and straightforward method. It does not consider the mean values and is invariant to linear transformations of the given feature but assumes linear dependencies between the variable and response variable, which may not always be accurate. Perfectly correlated features are genuinely redundant as no additional information is gained by adding them to the feature set. Chi-Square is another statistical FS method that is unbiased but not discriminative to an informative level of features. The Chi-Square FS method evaluates the continuous features after discretizing them. The number of bins needs to be specified. The ReliefF FS method is not dependent on any heuristics. Though ReliefF is not considered a multivariate FS method but it is noise-tolerant and robust to feature interactions. It depends on the number of nearest neighbours K.

Since the univariate FS method evaluates features individually, it may contain highly correlated features during the selection phase. mRMR provides a feature-set with maximal dependency on the target value and minimal dependency amongst each other. As mRMR considers the MI among features, better behaviour is obtained even when noise is present in data. Continuous variables require discretization pre-processing before applying mRMR.

Just Accepted

3.3 Classifiers

The problem under consideration is a binary classification problem. The features were fed iteratively to four classifiers: k-NN, SVM, DT, and RF, in an incremental manner, in the order of ranks provided by the respective feature selection methods. Next, we outline the classifiers used:

- **k-NEAREST NEIGHBOUR (k-NN) -** k-NN [14] is a famous instance-based method for classification and regression. It is a non-parametric classifier, simple to implement and interpret. In k-NN, an unlabelled instance is assigned a class belonging to most of its k neighbours. This algorithm is sensitive to the choice of k.
- **Support Vector Machine (SVM) -** SVM [13] is, originally, a binary classification approach that finds an optimal hyperplane to discriminate the instances belonging to two classes. The optimal hyperplane maximizes the margin around itself. A kernel function ($\Phi : \mathbb{R}^D \rightarrow \mathbb{R}^S$) may be used to project the data into a higher space so that the resulting data can be separated by a linear classifier (hyperplane). The optimal hyperplane $f(x) = \langle w, \Phi(x) \rangle + b$ can be obtained by solving the quadratic optimization problem

$$\text{minimize} \frac{1}{2} \parallel w \parallel_2^2 + C(\sum_{i=1}^{N} \xi_i) \tag{9}$$

$$\text{with respect to } \mathbf{w} \in \mathbb{R}^S, \xi \in \mathbb{R}_+^N, b \in \mathbb{R} \tag{10}$$

$$\text{subject to } y_i(\langle w, \Phi(x_i) \rangle + b) \geq 1 - \xi_i, \forall i \tag{11}$$

  where $\mathbf{w}$ is the weight coefficients vector, C is a positive trade-off parameter between model simplicity and classification error, ξ is the slack variable vector, and b corresponds to the bias in the hyperplane equation.
- **Decision Trees (DT) -** DT [10] summarises the data in the form of a decision tree. Nodes in the tree represent features selected for branching further on its values, and the tree's leaves indicate the class label. The decision tree employs a greedy approach to choose an attribute for the split at every node. If the purity of the class labels in the resulting sub-nodes is less than the parent node's purity, it implies that information has been gained, and the split must be maintained. The algorithm is applied recursively to form sub-trees, terminating when a given subset contains instances of only one class.
- **Random Forest (RF) -** The general framework of the RFC[26] uses an ensemble of more than one DT-based classifier. An unlabelled instance is classified to the mode of the classes (classification) or average prediction (regression) of the individual trees.

We considered four state-of-the-art classifiers for the experiments. k-NN[14] is a straightforward algorithm but has a high cost (computation and storage) of classifying new instances [8, 26]. Also, it is biased by the value of k. The smaller values of k may misclassify due to noise in the data. The larger value of k reduces the effect of noise on the classification but makes the experiments computationally expensive. It is also sensitive to the choice of similarity function due to many or irrelevant features. If the dataset is skewed, then k-NN must not be chosen [8, 26]. It may be a choice if the dataset is small and the relationship between the dependent and independent variables is unknown. SVM [8, 13, 26] is based on a robust theory and avoids estimating probabilities on the given dataset. It is highly effective in applications with dimensions more than the number of samples. Generally, it is not sensitive to over-fitting and is not biased by outliers. A non-linearly separable data can be transformed to a higher dimension that is linearly separable by applying the kernel trick. Selecting the appropriate kernel function and its hyper-parameters is a tricky task. SVM performs better when dealing with multi-dimensions and continuous features[8, 13, 26]. The major drawback of SVM is the lack of transparency and interpretability of results[8, 26].

The DT Classifier[10] is another powerful classifier that handles datasets with qualitative features[8, 26]. Numeric data has to be discretized beforehand. The learning speed concerning a high number of attributes and

Just Accepted

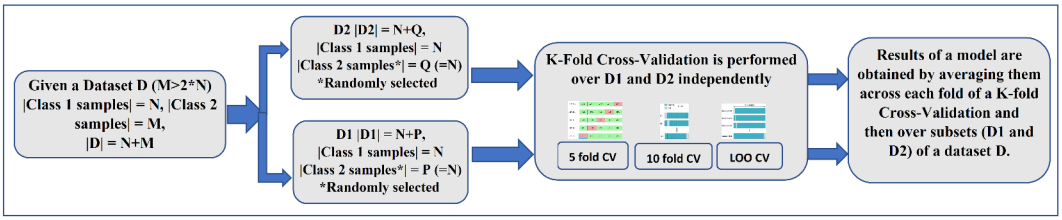


Fig. 3. 2-level sampling strategy used.

instances is fair. It is relatively tolerant of irrelevant attributes and missing values. It provides transparency in the model but may result in an over-fitted model for small datasets. One approach to tackle over-fitting is to pre-prune the decision tree. Also, a minor change in the dataset can change the tree entirely.

In contrast, RFC is robust and accurate for many problems with skewed datasets or datasets with high dimensionality and instances[8, 26]. If the data is changed a little (or noise is added), the individual trees may change, but the forest is relatively stable as it is an ensemble of many trees and gives stable results. The model generated is less over-fitted comparatively. A grid search of different parameters and random seeds is required to obtain the best results, but it is computationally expensive[8, 26]. Individual trees need to be uncorrelated/independent to get an unbiased result. RFC can easily tackle the problem of highly correlated features.

3.4 Sampling Strategies and Performance Metrics

We considered Accuracy and F-Measure as primary criteria for evaluating the performance. Accuracy defines the probability of correctly classifying an instance in the test sample. It is an optimal metric when the dataset is balanced regarding the class's instances frequency. F-Measure is calculated as the harmonic mean of sensitivity and specificity. The corresponding classifier parameters were optimal if they yielded the most considerable accuracy and F- measure.

We developed a 2-level sampling method for performing the experiments. In the first level, given a dataset D, if D is a skewed dataset (i.e. difference of class frequency in the dataset is high), two datasets - D1 and D2 are formed randomly such that the frequency of samples in both the classes in D1 and D2 is the same. In dataset D, the class with higher frequency is divided into two parts (P1 and P2) randomly such that the intersection of P1 and P2 is NULL. Finally, two subsets D1 and D2 are generated by appending minority class data to both P1 and P2, respectively.

In the second level, for each dataset (D1 and D2) formed, we independently performed a standard re-sampling method that is k-fold cross-validation for k=5, 10, and Leave-One-Out Cross-Validation (LOOCV) to obtain a more reliable estimate of the factual accuracy. The Accuracy and F-Measure values of K-fold cross-validation (K = 5, 10 and LOOCV) of a model are obtained by averaging them across each fold of k-fold cross-validation and then over subsets (D1 and D2) of dataset D. For instance, the average of D1's and D2's 5-fold cross-validation results (for instance) was taken as the final result of D's 5-fold cross-validation results for a given classifier and parameter values. This sampling strategy is illustrated in figure 3. We selected optimal model parameters based on their best performance.

Just Accepted

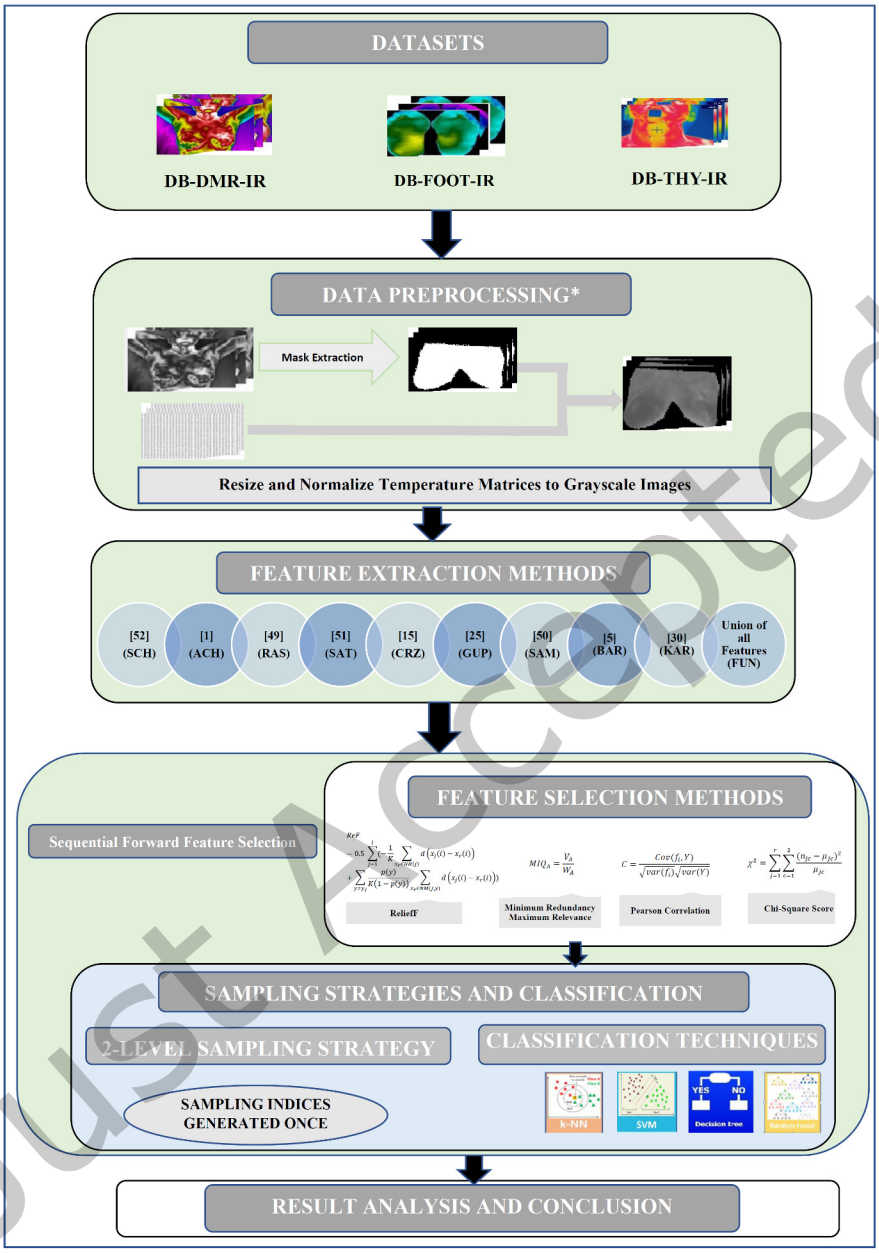


Fig. 4. The Machine Learning Framework/Uniform Platform for the Experiments conducted.

Just Accepted

4 EXPERIMENTAL DESIGN

The experiments were conducted on a computer using Windows 10 operating system with 16GB RAM and i7 (9^{th} generation) Processor. MATLAB 2020 (64-bit) was used for conducting the entire experiment. Figure 4 illustrates the ML framework followed for conducting the experiments.

Table 2. Size of the datasets used in the study

DataSets	Subsets formed	Abnormal Samples	Normal Samples	Total Cases
DB-DMR-IR Database	D1	40	40	80
	D2	40	40	80
	Total samples	40	80	120
DB-FOOT-IR Database	D1	45	45	90
	D2	45	45	90
	Total samples	90	45	135
DB-THY-IR Database	-	18	16	34

Table 3. Description of Feature Selection Method's parameters and their settings in the experiment.

Algorithm	Parameters	Parameters Values	Description of these parameters
ReF	K	$[5-15]$	The number of nearest neighbours from each class to be used.
ChiSq Score	Number of bins	$[10, 20]$	The number of bins for discretizing a numerical feature.

We considered three publicly available thermography based databases in our experiments- **DB-DMR-IR**[5] dataset [54], **DB-FOOT-IR**[27] and **DB-THY-IR**[6] [24]. The pre-processing steps of the three databases are described below:

- **DB-DMR-IR and DB-THY-IR** - The authors [24, 54] used both static and dynamic acquisition protocol. They used an environment with a temperature range of $20-24°C$ to cool the patient's relevant region, followed by image capturing using a FLIR thermal camera (model SC620). The patients were directed to relax for 15-20 minutes before capturing images to minimize the possible displacements of glands and attain thermal equilibrium. They acquired 20 Infrared images of dimensions 640 X 480 for a duration of five minutes for each patient. The thermal images taken last in the sequence are used in the experiments. The corresponding temperature matrices were available in both datasets. The medical and demographical history of the patients is also available[56]. BIRADS-related information was not available for the DB-DMR-IR database.
- **DB-FOOT-IR** - The authors [27] acquired images using a FLIR Thermal camera (model FLIR E60). During the acquisition, the camera's position was placed at a distance of one meter from the feet. The participants were directed to clean their feet with a damp towel and then relax in a room with a temperature of 20 - $22°C$ for 15 minutes before capturing images to reach a state of thermal equilibrium. The corresponding temperature matrices were also available with the images.

Relaxing for 15-30 minutes before capturing images helps achieve a state of thermal equilibrium in the body. This helps in developing a robust application and eliminates any temperature variation. Our experiments considered images acquired after achieving the body's thermal equilibrium. The acquired thermal images are pseudo-coloured/false-coloured and can not be used directly. Also, thermography-based applications use thermal

[5]http://visual.ic.uff.br/dmi/
[6]http://visual.ic.uff.br/thyroid/pt/

Just Accepted

Table 4. Description of classifier's parameters and their settings in the experiment.

Algorithm	Parameters	Parameters Values	Description of these parameters
k-NN	k	$[1-20]$	The number of neighbours to be used.
	Distance Metric	Euclidean Distance	The distance function for finding proximity with neighbours.
SVM	Cost	$[1-1024]$	The cost parameter C.
	Kernel Type	Linear, RBF	The type of kernel functions, used.
	Gamma	$[2^{-4}-2^{4}]$	The parameter of kernels functions (RBF, Polynomial, and Sigmoid).
DT	Splitting criteria	Gini's Diversity Index, Cross Entropy	Criteria used for selecting an attribute for further splitting
	Pruning criteria	Enabled, Disabled	Pruning the tree limits over-fitting of training data.
	Minimum leaf size	[4, 8, 12, 16, 20]	The minimum number of instances per leaf.
RF	Ensemble size	[3-50]	The number of weak classifiers/trees used in the ensemble.

images translated to temperature matrices. In all the databases, the temperature matrices are available along with the images. The temperature matrices are normalized, assuming the human body's minimum and maximum temperature to be 20° C and 45°C, respectively, and transformed into corresponding grey-scale images in the range [0-255] for experiments and visualization purposes using the equation -

$$I = (T - 20)\frac{255}{45 - 20} \tag{12}$$

where T denotes the temperature matrix, and I is the corresponding transformed grey-scale image. The region of interest was cropped from the transformed grayscale images, removing the background and other irrelevant body parts using MATLAB's utility *impoly*. All the images in a dataset were resized to a uniform size to eliminate any non-uniformity in images due to variations in human body sizes.

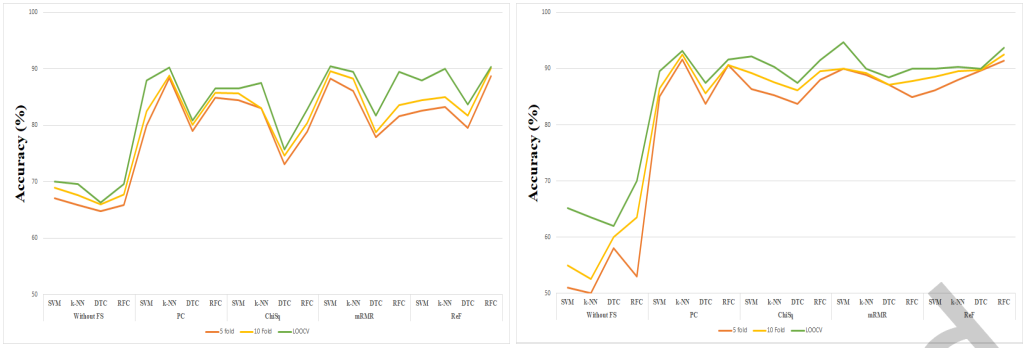
The dataset used in most of the literature is skewed [15, 49, 50, 52]. We performed 2 level sampling to get the datasets having a balanced number of samples from each class. The statistics of the datasets are stated in table 2.

With the exact line of detail, we implemented the features extractions methods mentioned in Table 1 and extracted features from transformed and segmented grey-scale images. We also created a feature set by taking a union of all the feature sets - UnionFeature_Set, for further analysis and obtained a total of 467 distinct features (UnionFeature_Set). All the feature sets generated were normalized.

We applied four well-known feature selection approaches to evaluate the relevance of all the features in feature sets, independently of the classifier. Table 3 summarizes the parameters and their values used for grid search for feature selection methods. K is a crucial tuning parameter of ReF. We ranged K in [5 - 15] with a step size of 4 and used Manhattan distance as the similarity measure. In the Chi-Square method, the number of bins for discretizing the numeric attributes ranged over [10, 20]. The features were fed iteratively to the classifier in the order of ranks provided by the respective feature selection methods (section 3.2) for each classifier parameter stated in table 4 and all the datasets. Also, we fed features incrementally as per the rank of features to the classifier and recorded the maximum performance in terms of accuracy and F-Measure.

We used four commonly used classifiers (k-NN, SVM, DT, and RF). We also performed a grid search with the respective parameters of the classifiers to achieve the best results. The classifier parameters for grid search are summarized in Table 4. k is a crucial tuning parameter of k-NN. We ranged k in [1 - 20] with a step size of 2 and used Euclidean distance as the similarity measure.

Just Accepted



(a) RAS [49] (b) UnionFeature_Set

Fig. 5. Comparison of classification accuracy values for a given combination of two feature extraction, feature selection methods, and cross-validation strategies on DB-DMR-IR dataset.

We have used LIBSVM[7] implementation of Support Vector Machine (SVM) with two kernel functions - Linear and Radial Basis function (RBF). We used RBF to handle the non-linearity between the features and class labels. We performed a grid search to optimize the hyper-parameters - Cost Parameter, C, for linear and RBF kernel functions, γ parameter for RBF kernel. To determine optimal hyper-parameters, we considered the values of C as $[1 - 1024]$ and the smoothing parameter value, γ, for RBF kernel was considered in the range $[2^{-4} - 2^{4}]$ in multiples of 2.

The parameters considered for DTC to achieve the best results are Splitting Criteria, Pruning Criteria, and Minimum Leaf Size. Gini's Diversity Index and Cross-Entropy are used as the splitting criteria. The experiments are performed with pruning enabled and disabled, and the minimum leaf size is ranged over [4, 8, 12, 16, 20]. In the Random Forest Classifier (RFC) framework, the number of weak classifiers (DTs) is varied over the range [3 - 50], with a step size of 4. A Random Forest is constructed with a template of DT with the exact parameter specifications as stated for DTC.

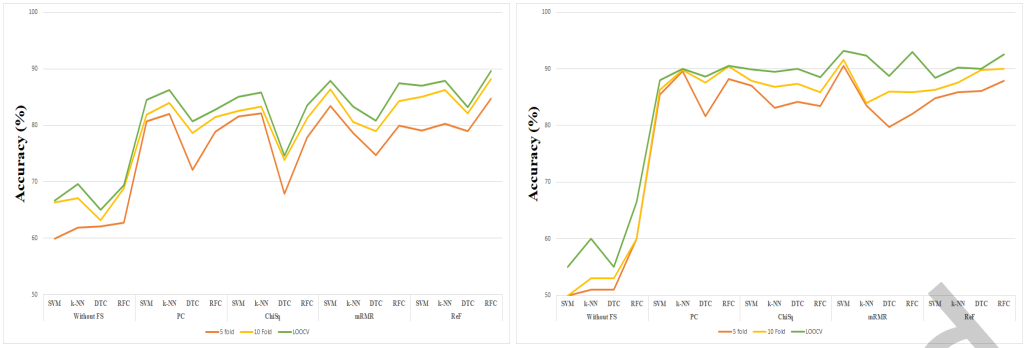
The 2-level Sampling for all the datasets was done only once for all the pipelines; that is, we divided each of DB-DMR-IR and DB-FOOT-IR datasets into two subsets, D1 and D2, and further generated 5 and 10 random stratified partitions from each subset, respectively, only once. It was done to eliminate any randomness in the experiments. That is, the partitions were kept the same for all the pipelines. Along with 5-fold and 10-fold Cross-Validation, LOOCV was also done to contrast the behaviour of data. All the results are recorded along with the parameters for which the best accuracy and F-Measure are achieved.

5 RESULTS AND DISCUSSION

The performance of experimental results is compared based on classification Accuracy and F-Measure. Also, Accuracy and F-Measure values of a model are obtained by averaging them across each fold of k-fold cross-validation and all subsets (D1 and D2) of dataset D(= DB-DMR-IR and DB-FOOT-IR). The experiments conducted on three datasets are independent and not related to each other in any respect. Figures 5, 6 and 7 illustrate the visual comparison of classification accuracy for a combination of two feature extraction methods, FS methods, classifiers, and cross-validation approaches. Graphs (figures 11, 12 and 13) showing the performance of the experiments conducted for rest of the feature extraction methods are moved to Appendix A.

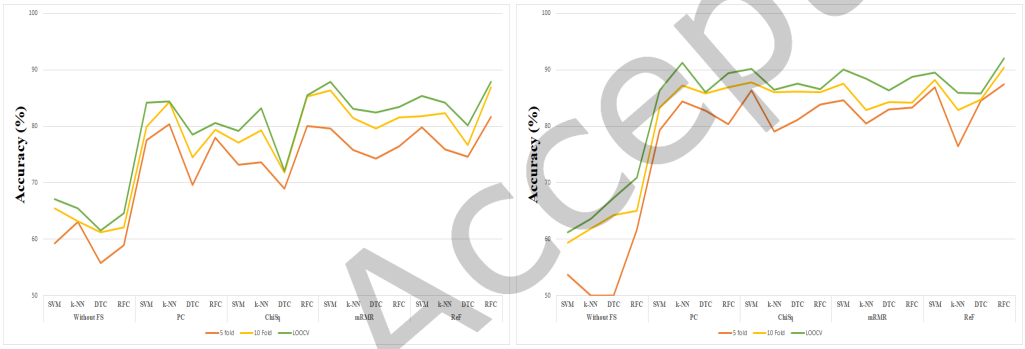
[7]https://www.csie.ntu.edu.tw/cjlin/libsvm/

Just Accepted



(a) RAS [49] (b) UnionFeature_Set

Fig. 6. Comparison of classification accuracy values for a given combination of two feature extraction, feature selection methods, and cross-validation strategies on DB-FOOT-IR dataset.



(a) RAS [49] (b) UnionFeature_Set

Fig. 7. Comparison of classification accuracy values for a given combination of two feature extraction, feature selection methods, and cross-validation strategies on DB-THY-IR dataset.

Just Accepted

We can observe from Figures 5, 6 and 7 and figures in Appendix A (figures 11, 12 and 13) that among the three re-sampling strategies (5-fold, 10-fold, and LOOCV), LOOCV strategy (green line) gave the best results for all the nine feature extraction methods, individually and on their union. The best performance is achieved using LOOCV as the maximum number of samples for training is available. Henceforth, all the analysis is done on results obtained with LOOCV, shown in Table 6 for DB-DMR-IR, DB-FOOT-IR, and DB-THY-IR datasets, respectively.

Broadly, we can observe the following from Table 6 corresponding to DB-DMR-IR, DB-FOOT-IR and DB-THY-IR datasets, respectively:

- The union of features (UnionFeature_Set) from the existing nine research works (UnionFeature_Set) provided better results in comparison to each of the existing research work individually for all the combinations of 4 FS methods and 4 classifiers across all the datasets. Table 5 states the best-performing models along with their accuracy and F-Measure for all the datasets considered.

Table 5. The classification accuracy(%) and F-measure values (%) obtained from the best performing model using LOOCV strategy on DB-DMR-IR, DB-FOOT-IR, and DB-THY-IR datasets.

DataSets	FE Method	Classification Model	Feature-subset	Accuracy (%)	F-Measure (%)
DB-DMR-IR	UnionFeature_Set	SVM+mRMR	45	94.75	93.80
DB-FOOT-IR	UnionFeature_Set	SVM+mRMR	57	93.14	92.55
DB-THY-IR	UnionFeature_Set	RF+ReF	39	92.06	91.42

- The classification accuracy for all the nine feature extraction methods, individually and in their union, improves with the use of all four FS methods (PC, ChiSq, mRMR, and ReF). In some cases, a reduced feature set gave the same performance as that with all features of a given feature extraction method. Hence, better visualization may be achieved with a reduced subset of features without losing performance.

Sections 5.1 and 5.3 introspect the performance of nine feature extraction methods, considered individually and in combination, respectively. Also, we compare the performance of all the 40 combinations of feature extraction (FE) and FS methods in section 5.2.

5.1 Comparison of Feature Extraction (FE) Methods

In this subsection, we introspect the performance of state-of-the-art feature extraction methods based on the results in Table 6. The following can be observed:

- Without feature selection and among nine individual research works, the feature-set suggested in the work BAR[5] gave the best result for DB-DMR-IR with k-NN and DB-THY-IR with SVM. The features set suggested in the work GUP[25] gave the best performance for DB-FOOT-IR with k-NN and RF. With feature selection also and all the datasets, the feature-set suggested in work BAR[5] gave better performance than the other eight research works.
- Finally, the combination of features (UnionFeature_Set) provides maximum classification accuracy of 94.75% for DB-DMR-IR and 93.14% for DB-FOOT-IR datasets with the combination of mRMR and SVM. The dataset DB-THY-IR gave an accuracy of 92.06% with the combination of RF and ReF. The F-Measure obtained are 93.80%, 92.55%, and 91.42% for DB-DMR-IR, DB-FOOT-IR, and DB-THY-IR datasets, respectively. Similar results on DB-DMR-IR and DB-FOOT-IR are obtained with the combination of ReF and RF.
- Majority of the FE methods (RAS[49], SCH[52], SAT[51], GUP[25] and CRZ[15]) performed better with the combination of [mRMR and SVM] and [Ref and RFC] for DB-DMR-IR dataset. All the feature extraction methods gave the best results with the same combinations ([mRMR and SVM] and [Ref and RFC]) for DB-FOOT-IR and DB-THY-IR datasets.
- GLCM and GLRM-based features suggested in research works ACH[1], CRZ[15], and KAR[30], have not yielded satisfactory results for all the datasets with and Without FS (WFS) methods. Amongst the nine FE methods, the worst performance is shown by the feature-set suggested in KAR[30] for all the datasets, without using any FS method. With feature selection, the average performance of the feature extraction methods KAR[30] and CRZ[15] is minimum for all the datasets.
- Amongst four classifiers, the best performance is achieved with SVM or RF for most of the feature-sets on all the datasets.
- The combinations of [mRMR and SVM] and [Ref and RFC] performed better in the majority of the ten feature-sets for all the datasets.
- DTC did not perform effectively for most of the research works and datasets.

Just Accepted

Table 6. The accuracy (%) and F-Measure values (%) obtained from various combinations of feature extraction methods, feature selection methods, classifiers, and LOOCV strategy on three datasets.

S.No.	Reference	Evaluation of Reference on the Dataset	Feature Selection (FS) method	DB-DMR-IR						DB-FOOT-IR						DB-THY-IR					
				Accuracy				F-Measure			Accuracy			F-Measure			Accuracy			F-Measure		
				SVM	*k*-NN	DTC	RFC	SVM	*k*-NN	DTC	RFC	SVM	*k*-NN	DTC	RFC	SVM	*k*-NN	DTC	RFC	SVM	*k*-NN	DTC	RFC
		DB-DMR-IR-Acc-72.5%,F-	WFS	70	69.6	66.3	69.6	70	67.5	66.23	68.50	66.64	69.66	65	69.37	65.74	70	65	68.50	67.1	65.5	61.6	64.6	66.72	64.07	63.6	62.50
		Meas-70%,	PC	68	90.29	80.86	86.57	85	90.35	80.59	85.04	84.53	86.31	80.73	82.76	83.87	85.71	80	82.31	84.17	84.37	78.5	80.64	83.24	83.44	78.78	80.00
1	RAS[49]	DB-FOOT-IR-Acc-	ChiSq	86.5	87.5	75.71	82.86	85.89	87.97	75.8	81.78	85.05	85.84	74.66	83.52	86.19	84.61	73.71	84.00	79.16	83.17	72.07	85.53	80	83.18	73.1	84.61
		61.5%,F-Meas-60%,	mRMR	**90.5**	89.43	81.67	89.43	**87.5**	90.62	80.27	89.81	87.91	83.28	80.28	87.49	86.8	84.58	79.14	87.12	**87.9**	83.1	82.41	83.4	**87.05**	81.61	82.32	83.54
		DB-THY-IR-Acc-57.5%,F-	ReF	88	90	83.71	90.35	92.07	90	85.51	91.92	87.07	87.93	83.21	**89.67**	86.09	87.43	81.92	**91.54**	85.4	84.2	80.19	**87.9**	84.38	85.97	81.73	**88.49**
		[illegible]	WFS	66.57	68	65.6	70	67.99	66.74	64.18	69.96	67.29	67.65	65.37	70.49	66.14	66.99	64.76	69.01	65	63.3	60.7	70	65	62.87	59.57	68.27
		Meas-65.0%	PC	80	73.71	76.57	86.57	80	72.39	76.5	88.31	76.43	72.64	74.31	84.29	78.1	72.23	72.74	82.32	75.4	71.51	72.57	81.97	73.86	69.88	72.43	83.33
2	SCH[52]	DB-FOOT-IR-Acc-60%,F-	ChiSq	79.5	86.57	79.43	85.14	77.5	88.33	79.16	85.93	78.94	83.04	80.72	81.16	80	82.43	80.14	79.88	78.3	83.67	77.53	81.14	77.08	82.02	76.27	80.00
		Meas-60%,	mRMR	**88.57**	75.14	76.57	87.12	**83.6**	76.62	75.96	87.59	88.38	76.13	77.4	84.96	87.06	74.56	76.51	85.46	85.77	73.84	73.57	82.42	83.85	74.58	73.77	81.32
		DB-THY-IR-Acc-60%,F-	ReF	86.17	75.83	76.29	84.5	84.6	76.3	75	89.69	83.74	75.15	72.82	**89.82**	82.66	73.25	71.52	**89.3**	82.27	74.83	72.29	**86.7**	82.26	74.94	70.3	**87.37**
		[illegible]	WFS	66.57	69.43	66.6	70	65.13	67.66	68.38	69.56	66.04	70.28	64.08	70	64.71	69.65	63.3	72.09	64.77	67.93	61.9	68.6	63.24	66.61	62.87	69.45
		Meas-60%,	PC	78	77.5	76.57	80	77.9	76.24	75	80.00	76.14	74.56	73.99	76.04	76.49	73.53	72.31	77.15	73	75.1	74.37	76.83	74.4	75.31	74.5	74.91
3	SAM[50]	DB-FOOT-IR-Acc-	ChiSq	77.5	78.57	79.43	79.43	77.5	77.93	80	79.05	77.16	77.4	75.27	76.34	75.46	76.28	76.16	75.54	74.9	76.87	76.5	77.23	73.32	78.02	76.5	79.12
		58.75%,F-Meas-55%,	mRMR	80.86	72.29	76.57	79.43	80	70	74.83	78.91	**80.58**	72.02	74.68	78.39	**81.88**	71.61	73.91	79.85	78.86	72.39	72.77	77.73	77.19	70.4	72.51	76.52
		DB-THY-IR-Acc-55%,F-	ReF	79.5	**82.71**	80.14	80	78.18	**82.64**	81.16	80.00	79.28	80.19	76.88	80.19	80.33	81.76	74.94	78.68	77.6	78.61	76.34	**80**	79.28	78.5	76.23	**80**
		[illegible]	WFS	68	70	65.5	70	69.76	70	65.5	69.56	68.11	70	62.17	69.5	69.58	72.01	61.13	69.53	64	68.19	62.9	67.5	62.07	67.4	64.31	66.64
		Meas-76.53%,DB-FOOT-	PC	75	75.14	78.57	81.67	75	76.16	80	83.05	74.74	74.17	78.64	80	73.96	73.06	79.97	80.00	72.24	73.74	76.37	77.07	73.26	72.62	75.57	76.55
4	ACH[1]	DB-FOOT-IR-Acc-	ChiSq	73.71	75.83	72.29	76.57	74.18	73.61	70.65	74.63	74.41	75.32	71.29	73.55	72.63	76.97	70.47	73.11	71.91	73.53	72.59	74.17	70.34	74.47	71.28	74.25
		58.5%,F-Meas-56.6%,	mRMR	78	76.57	79.43	77.5	78	76.69	77.55	77.50	**80.53**	74.62	78.46	80	**81.42**	75.62	77	80.00	**82.3**	72.57	76.43	75	**81.41**	72.19	75.03	75.00
		DB-THY-IR-Acc-56.5%,F-	ReF	80	**84.5**	75	80	80	**85.76**	75	80.00	80.58	80	71.94	80	80	80	71.58	80.00	80.2	77.34	70.14	80	80	77.17	69.94	80.00
		[illegible]	WFS	70	66	65	70	66.6	65	65	70.00	69.34	65.61	65.03	69.84	67.36	64.67	63.48	68.14	66.2	64.1	63.9	70.1	64.42	63.64	61.91	68.65
		Meas-76.53%,DB-FOOT-	PC	80	82.5	76.57	82.29	81.58	82.5	75.99	83.88	77.09	82.5	74.88	79.93	75.5	82.5	75.02	80.05	76.4	77.9	75.57	79.99	75.46	78.64	75.4	81.55
5	SAT[51]	IR-Acc-75%,F-Meas-	ChiSq	80.86	72.29	72.5	75	80.43	72.91	72.5	75.00	77.75	75	73.68	73.67	79.25	75	73.1	72.25	78.36	75	73.8	72.54	76.61	75	73.34	74.37
		73.75%,	mRMR	**88.35**	78	81.67	85	**87.55**	76.59	80.59	85.00	**89.77**	73.34	76.99	85.29	**91.53**	73.79	75.98	86.93	**86.9**	73.3	76.77	83.8	**88.33**	74.46	74.93	83.82
		DB-THY-IR-Acc-74.5%,F-	ReF	86.67	76.67	80.86	87.5	86.72	75	80.99	87.50	84.87	76.51	78.46	85.5	84.14	76.51	76.74	84.10	82.67	74.47	79.36	86.5	81.5	74.34	78.68	85.00
		[illegible]	WFS	62.29	63.33	63.33	65	61.65	62.47	64.54	65.00	60.41	63.24	63.67	64.74	61.63	63.89	61.82	64.61	58.79	62.13	61.43	63.7	57.68	63.47	62.18	63.42
		Meas-60%,DB-FOOT-IR-	PC	70	77	72.5	73.71	70	78.53	72.5	72.14	81.05	76.79	73.09	74.86	80.31	75.76	74.4	73.94	79.66	72.7	72.8	72.71	81.56	71.45	72.49	73.88
6	KAR[30]	Acc-54.5%,F-Meas-52%,	ChiSq	73.55	72.29	72.5	75	71.66	70.65	72.5	75.00	79.97	73.64	74.22	74.87	79.03	72.5	74.12	76.57	76.86	73.99	71.5	73.64	76.83	73.44	71.5	73.06
		DB-THY-IR-Acc-50%,F-	mRMR	80	69.4	70.86	73.71	80	68.63	71.81	72.46	**83.19**	76.02	72.25	76.68	**82.17**	77.56	72.34	75.35	81.6	78.4	73.86	76.61	80.49	77.4	73.53	77.67
		Meas-50%	ReF	72	**83.58**	75.83	80.5	70	**84.05**	75.79	80.60	81.72	81.05	75	81.05	79.84	80	75	82.29	77.86	**82.081**	72.93	79.1	78.7	**82.621**	74.92	78.47
		DB-DMR-IR-Acc-	WFS	68	70.83	71.57	70	65	68.5	66	70.00	69.68	67.55	68.71	70	70.71	67.02	66.94	70.00	68.6	67.93	68.77	68.4	69.26	68.88	70.62	67.13
		70.13%,F-Meas-70%,DB-	PC	75	80	80	80.48	73.54	80	80	82.80	78.83	81.77	76.59	83.58	77.7	81.36	76.02	83.91	77.7	78.7	77.9	**81.14**	78.13	76.83	77.32	**80.69**
7	CRZ[15]	FOOT-IR-Acc-67.6%,F-	ChiSq	72	75.83	77.5	78	70	75.75	77.5	77.19	75.27	74.89	78.16	77.83	76.32	74.28	79.26	77.66	72.24	73.73	75.6	76.4	71.91	73.17	74.76	76.29
		Meas-66.6%,	mRMR	80	75.2	75	80	80	76.98	75	80.00	83.82	75	72.7	76.81	84.95	75	71.21	75.24	80.9	77.8	72.54	78.1	79.67	76.95	72.01	77.56
		DB-THY-IR-Acc-	ReF	70	77.5	76.57	**85.14**	80.51	77.65	76.35	**83.50**	78.28	82.45	75.52	**84.08**	76.79	81.91	75.33	**83.698**	78.4	79.14	72.27	**81.14**	77.33	80.77	70.95	**80.62**
		[illegible]	WFS	68.5	**72.5**	66.5	67	69.34	70	65	67.00	72	71.5	62.81	69.74	70	66.36	61.48	66.03	71.5	65.5	63.7	65.5	70	65.5	64.09	64.16
		Meas-66.66%,DB-FOOT-	PC	86.57	86.57	80	85.14	85.65	85.91	80	85.49	84.94	85	79.2	83.65	83.61	87.87	78.13	84.22	82.9	88.89	78.66	81.97	83.69	87.98	78.22	83.04
8	BAR[5]	IR-Acc-53.5%, F- Meas -	ChiSq	80.86	86.57	76.57	87.43	80.54	84.95	75.43	88.45	83.97	86.18	74.01	83.2	83.75	84.27	73.82	83.80	83.8	85.2	74.01	80.46	84.16	83.4	74.19	78.75
		52%, DB-THY-IR-Acc	mRMR	91.86	89.43	83.71	88	91.71	87.79	83.3	87.78	90	83.09	82.58	87.7	90	82.11	81.62	86.15	**89.1**	81.43	78.67	84.53	**87.2**	81.55	77.65	83.04
		51%,F-Meas-50%	ReF	89.5	88	82.29	**92.28**	89.3	88	81.68	**91.72**	86.5	84.74	80.8	**91.5**	86.14	82.79	81.17	**90.07**	86.09	85.7	81.41	88.45	84.18	87.69	80	86.47
		DB-DMR-IR-Acc-80%,F-	WFS	68.5	71.5	71.5	71.66	68.5	72.06	70	70.00	69.68	73.04	68.57	73.95	68.78	71.17	66.62	75.29	67.3	70.77	70	70.1	67.76	70.42	70	68.52
		Meas-80%,	PC	85.14	88.86	76.57	87.5	84.67	87.29	75.99	89.09	82.5	85.6	75.49	85.28	82.79	83.69	73.62	84.88	80.84	84.16	75.57	86.2	80	82.63	76.47	85.76
9	GUP[25]	DB-FOOT-IR-Acc-75%,F-	ChiSq	86.57	86.57	79.43	85.14	85.51	86	80	85.94	83.29	82.27	77.65	85.7	81.76	82.56	75.74	86.63	81.87	81.67	74.73	83.64	81.55	80	75.13	82.81
		Meas-72.8%,	mRMR	90.71	88	77.6	87.5	92.5	88	77.18	87.50	88.41	81.38	77.36	84.08	86.5	82.66	79.08	83.87	88.51	82.5	74.2	83	87.96	82.5	74.58	83.76
		DB-THY-IR-Acc-	ReF	90	87.5	75.71	**90.86**	90	88	74.98	**90.53**	86.83	84.93	79.15	**90.30**	86.18	85.53	79	**88.835**	87.5	82.9	71.91	**89.56**	87.5	83.72	72.9	**87.87**
		76.75%,F-Meas-75%	WFS	65.14	63.5	62	70	66.06	68	67.59	67.00	55	60	55	66.5	55	55	50	60.00	61.24	63.6	67.33	70.9	59.58	63.39	66.08	72.32
			PC	89.6	93.16	87.5	91.66	92.5	92.65	87.5	92.28	88	90	88.58	90.51	86.18	90.51	86.87	94.67	86.57	91.23	86.1	89.46	84.46	90.24	85.94	90.78
10	UnionFeature_Set		ChiSq	92.17	90.29	87.5	91.5	93.88	91.51	87.5	92.36	89.85	89.48	90	88.48	91.07	89.07	87.54	86.74	90.15	86.5	87.57	86.6	88.28	88.18	86.08	87.70
			mRMR	**94.75**	90	88.5	90	**93.8**	90	85	90.00	**93.14**	92.38	88.68	92.95	**92.55**	90	89.03	90.21	90.07	88.39	86.4	88.8	89.85	86.3	87.55	88.32
			ReF	90	90.29	90	93.67	90.98	91.92	90.85	95.69	88.37	90.23	90	92.5	86.55	88.66	90	91.16	89.5	85.99	85.8	**92.06**	88.84	86.39	84.17	**91.42**

Just Accepted

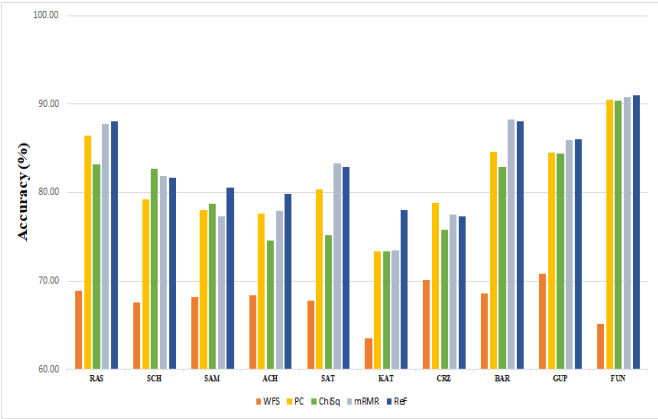


Fig. 8. Comparison of the average classification accuracy values for a given combination of feature extraction and selection method for DB-DMR-IR dataset.

5.2 Comparison of Combinations of FE and FS Methods

In this subsection, we investigate the performance of four FS methods in conjunction with the nine feature extraction methods suggested in the literature individually and their combination (UnionFeature_Set). We will compare the results statistically also.

Figures 8, 14(a) (in Appendix B), and 14(b) (in Appendix B) illustrates the comparison of average classification accuracy (over 4 different classifiers) for a combination of feature extraction and selection method on three datasets, respectively. The following can be observed from Figures 8, 14(a), and 14(b) about the various combinations:

- For each FS method, the average performance improved significantly compared to Without FS, in terms of accuracy and F-Measure, for each feature extraction method.
- The average performance achieved with all the four FS methods on UnionFeature_Set is better than feature selection on all of the individual nine feature-sets suggested in the literature.
- On average, the performance of ReF and mRMR is similar and better than PC and ChiSq for the majority of the feature extraction methods and all the datasets.

We investigated the relative performance of 40 combinations of 10 feature extraction and 4 FS methods for all the datasets separately, with the help of a robust ranking mechanism. This method ranks the combinations based on their net improvement in performance with reference to the worst performance (m_c^{WFS}) obtained among 10 FE methods and without any feature selection and best performance (M_c). Let N and n_c denote the number of combinations (= 40) of feature extraction and selection methods and the number of classifiers (= 4). The average percentage improvement in accuracy, ($g_i, i = \{1, 2, ..., N\}$), of a combination i, is calculated as follows:

$$g_i = \frac{1}{n_c} \sum_{c=1}^{n_c} \frac{a_{ic} - m_c^{WFS}}{M_c - m_c^{WFS}} \times 100 \tag{13}$$

where a_{ic} is the accuracy achieved for the model with combination i and classifier c. The rank of a combination i, r_i , is assigned as $r_a \leq r_b$ if $g_a \geq g_b$.

Table 7 shows net percentage improvement in accuracy for all combinations of feature extraction and selection methods for all the datasets. Figure 9, 15(a) (in Appendix B) and 15(b) (in Appendix B) illustrate the ranking of the combinations obtained by sorting the improvement in accuracy values in descending order for all the datasets, respectively. We observed the following from Table 7 and Figures 9, 15(a), and 15(b):

Just Accepted

Table 7. Performance Improvement (%) in Accuracy for all the combinations of feature extraction and selection methods.

S. No	Combinations of FE and FS Methods	DMR-IR	DB-FOOT-IR	DB-THY-IR
1	UnionFeature_Set + ReF (FUN+ReF)	**93.93**	94.82	93.33
2	UnionFeature_Set + mRMR (FUN+mRMR)	92.81	**99.05**	**93.53**
3	UnionFeature_Set + PC (FUN+PC)	92.05	91.61	93.32
4	UnionFeature_Set + C2Sq (FUN+C2Sq)	91.48	91.64	91.12
5	[5] + mRMR (BAR+mRMR)	84.09	80.82	75.82
6	[49] + ReF (RAS+ReF)	83.65	84.83	79.64
7	[5] + ReF (BAR+ReF)	83.54	81.89	83.10
8	[49] + mRMR (RAS+mRMR)	82.46	78.12	78.79
9	[49] + PC (RAS+PC)	78.04	74.01	70.84
10	[25] + ReF (GUP+ReF)	76.39	80.01	73.96
11	[25] + mRMR (GUP+mRMR)	76.12	71.51	70.77
12	[5] + PC (BAR+PC)	71.81	72.97	75.02
13	[25] + PC (GUP+PC)	71.63	70.62	70.18
14	[25] + C2Sq (GUP+C2Sq)	71.30	70.49	65.82
15	[51] + mRMR (SAT+mRMR)	67.36	67.00	64.67
16	[49] + C2Sq (RAS+C2Sq)	66.72	70.32	64.14
17	[51] + ReF (SAT+ReF)	66.41	67.48	67.09
18	[5] + C2Sq (BAR+C2Sq)	66.35	69.13	66.92
19	[52] + C2Sq (SCH+C2Sq)	65.86	66.40	65.09
20	[52] + mRMR (SCH+mRMR)	62.44	68.25	60.05
21	[52] + ReF (SCH+ReF)	62.12	65.49	60.69
22	[50] + ReF (SAM+ReF)	58.78	60.82	58.08
23	[51] + PC (SAT+PC)	57.79	59.51	55.78
24	[1] + ReF (ACH+ReF)	56.07	57.84	53.29
25	[52] + PC (SCH+PC)	54.16	54.92	48.45
26	[15] + PC (CRZ+PC)	53.33	64.54	60.69
27	[50] + C2Sq (SAM+C2Sq)	52.63	52.71	52.13
28	[30] + ReF (KAR+ReF)	50.32	62.50	57.30
29	[50] + PC (SAM+PC)	50.07	48.67	46.76
30	[1] + mRMR (ACH+mRMR)	49.66	58.31	52.31
31	[1] + PC (ACH+PC)	49.02	54.29	47.06
32	[15] + ReF (CRZ+ReF)	48.39	64.35	56.39
33	[15] + mRMR (CRZ+mRMR)	48.27	53.81	54.80
34	[50] + mRMR (SAM+mRMR)	47.40	52.20	48.41
35	[15] + C2Sq (CRZ+C2Sq)	43.13	52.92	45.75
36	[51] + C2Sq (SAT+C2Sq)	39.90	47.74	46.64
37	[1] + C2Sq (ACH+C2Sq)	38.55	43.99	40.55
38	[30] + PC (KAR+PC)	34.37	51.93	44.92
39	[30] + C2Sq (KAR+C2Sq)	34.27	49.60	43.40
40	[30] + mRMR (KAR+mRMR)	34.23	53.75	55.79

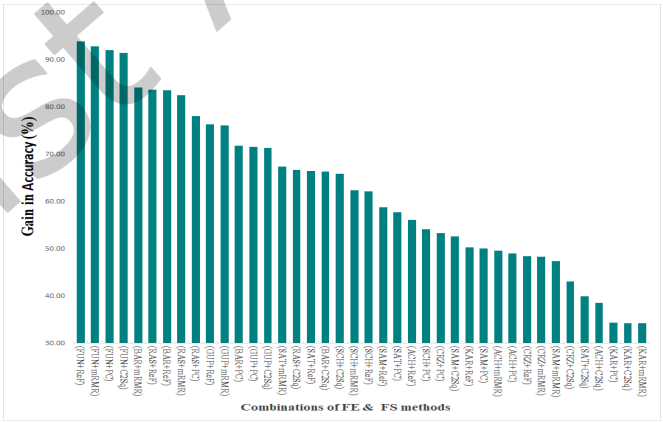


Fig. 9. Ranking of combinations of feature extraction and selection methods based on Performance Improvement (%) for DB-DMR-IR dataset.

Just Accepted

- The improvement in relative accuracy (%) with the combination of UnionFeature_Set and any feature selection is higher than any other combination for all the datasets. The maximum improvement in accuracy is achieved with the ReF and mRMR for all the datasets, respectively.
- The improvement in performance with feature-set used in the research work KAR[30], SAM[50] and ACH[1] is worse for all the datasets.
- ReF and mRMR FS methods achieved higher net profit/improvement in accuracy compared to PC and ChiSq, for most of the research works, in individual and in combination, and all the datasets. ReF and mRMR both take care of the correlation between the features and minimize the redundancy in the feature set.

5.3 Comparison of Features (UnionFeature_Set)

Here, we introspect the performance of UnionFeature_Set, having a total of 467 distinct features, quantitatively and qualitatively. The accuracy and F-Measure values of the experiments with UnionFeature_Set on the proposed framework are stated in the last row of Table 6. Table 8 displays the subset of features for which the best result is obtained for all the datasets, along with their p-values using the T-Test. Tables 9, 10 (in Appendix C), and 11 (in Appendix C) present the color map of correlation matrices of the features with which the best performance is achieved for the three datasets, respectively. The absolute value of the correlation between a feature pair is considered. Following observations are made from Tables 8, 9, 10 (in Appendix C), and 11 (in Appendix C) for the experiments conducted on UnionFeature_Set:

- UnionFeature_Set performed best with a reduced set of 45 and 57 features for DB-DMR-IR, DB-FOOT-IR selected with mRMR and 39 features for DB-THY-IR selected using ReF, wherein the features are from the research works, SCH[52], RAS[49], SAT[51], BAR[5], GUP[25], and ACH[1].
- Majority of the features in feature subset (Table 8) that gave the best results are found to be relevant at the significance level, $\alpha = 0.01$, for all the datasets. Also, most of them (RAS[49], BAR[5], GUP[25] and ACH[1]) are common for the datasets considered.
- Most of the best-performing features (Table 8) are minimally correlated to each other for all the datasets (cell colors are lighter red in shade(orange), Tables 9, 10 and 11).
- Features proposed by BAR[5] are found to be the most prominent features for discriminating against abnormal thermal patterns.
- The thermal patterns for abnormality due to inflammation are best distinguished when the features (Histogram-based features, First Order Statistical features, Texture-based features, and LBP-based features) are computed from the segmented region of interest (BAR[5] and CRZ[15]).

Also, we performed the qualitative analysis of the best-performing models (stated in table 5) using Explainable AI with SHapley Additive exPlanation (SHAP) [38–40]. SHAP explains the prediction given by a classification model using a game theory-based approach. It uses Shapley values to measure the contribution of features in predicting the outcome of the model. We used the Python SHAP library to understand the feature's importance in predicting abnormality in the human body, captured in thermal images. To analyze the importance of the features in UnionFeature_Set we passed all the 467 features with the model (with tuned parameters) to SHAP Explainer to analyze their importance.

Figure 10 displays the absolute summary plot of the models(5) for the three datasets – DB-DMR-IR, DB-FOOT-IR, and DB-THY-IR. The figure also demonstrates the relative importance of features by class. In this figure, the average absolute value of the SHAP values for each variable is taken in order to obtain a bar chart as a function of the contribution of each variable to the prediction of the model. The features are ordered from being most (top) to least (bottom) influential. We plotted the graphs for the five most significant features (out of 467 features) that contribute to the model's prediction. We observed the following from figure 10.

Just Accepted

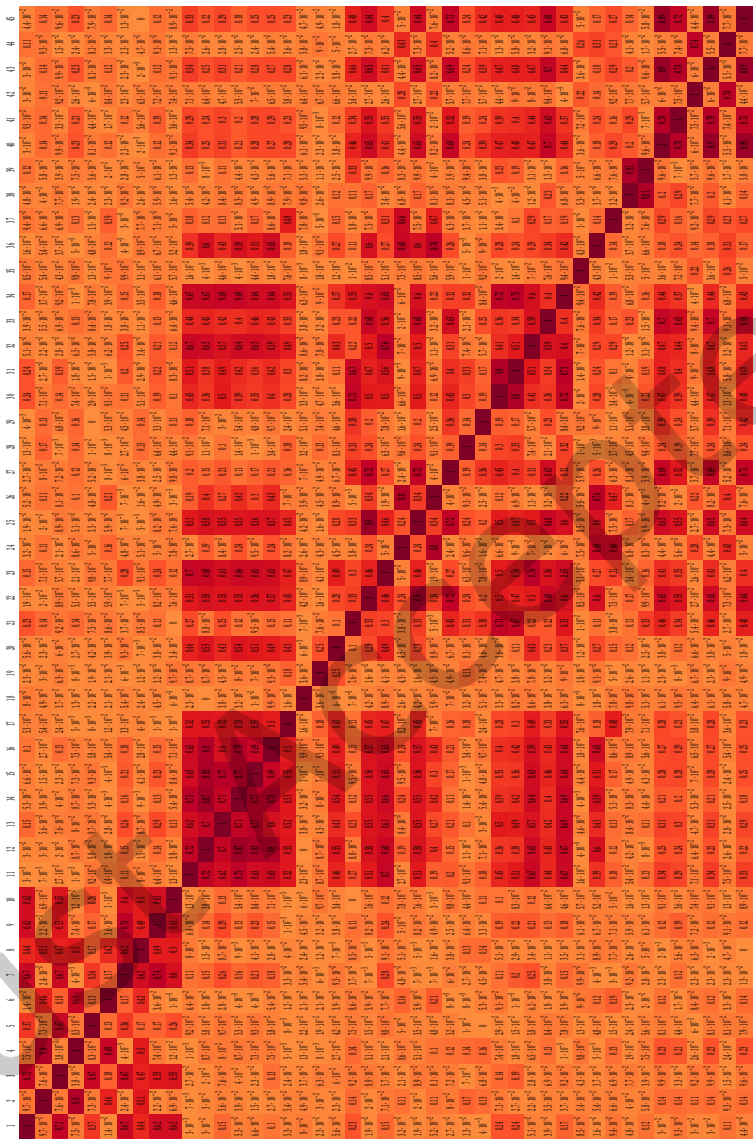
Table 8. The Features and their p-Values with which the best result is obtained for all the datasets.

S. No.	Reference	Feature Name	p-Value for DB-DMR-IR Dataset	p-Value for DB-FOOT-IR Dataset	p-Value for DB-THY-IR Dataset	S. No.	Reference	Feature Name	p-Value for DB-DMR-IR Dataset	p-Value for DB-FOOT-IR Dataset	p-Value for DB-THY-IR Dataset
F1	**BAR[5]**	Texture energy using LS kernel	**9.54E-06**	**0.000736465**	-	F34	**SAT[51]**	Mean (from Local energy matrices, calculated for sub-band 6)	0.036016195	-	-
F2	**BAR[5]**	Contrast(Neighborhood gray-tone difference matrix based)	**3.46088E-05**	**0.000232748**	**2.15761E-05**	F35	**RAS[49]**	GLCM based Energy (from Modified LBP Colour matrix)	0.036059195	0.046864998	0.024161069
F3	**BAR[5]**	Contrast(Gray level difference statistics based)	**5.19114E-05**	**0.000261966**	**3.89146E-05**	F36	**RAS[49]**	GLCM based Energy (from Modified LBP Edge matrix)	0.036136696	0.047042847	**0.000972879**
F4	**BAR[5]**	Coarseness(Neighborhood gray-tone difference matrix based)	**6.09136E-05**	**7.73816E-05**	**0.000301748**	F37	**RAS[49]**	Sum from Modified LBP Colour matrix	0.036277198	0.051967651	-
F5	**BAR[5]**	Busyness (Neighborhood gray-tone difference matrix based)	**6.86098E-05**	**6.47659E-05**	**0.000231966**	F38	**RAS[49]**	Sum from Modified LBP Edge matrix	0.03639456	0.053655003	-
F6	**BAR[5]**	Texture energy using LE kernel	**0.000276149**	**0.00526576**	**0.008670024**	F39	**SCH[52]**	Contrast (Gray Level Cross co-occurrence matrix from Laplace filter applied images)	0.036920254	0.030455145	0.031042951
F7	**BAR[5]**	Eccentricity	**0.000321131**	**0.000414763**	-	F40	**SCH[52]**	Contrast (Gray Level Cross co-occurrence matrix)	0.038373094	-	-
F8	**BAR[5]**	Regularity/roughness(Statistical feature matrix based)	**0.000352973**	-	-	F41	**SAT[51]**	Variance (from Local energy matrices, calculated for sub-band 12)	0.056953143	**0.001364869**	0.075566216
F9	**BAR[5]**	Euler number	**0.000471395**	**0.001101846**	-	F42	**SCH[52]**	Symmetry (Gray Level Cross co-occurrence matrix from Laplace filter applied images)	0.066294345	0.118365239	-
F10	**BAR[5]**	Convexity	**0.000502856**	**0.000353501**	**1.9523E-05**	F43	**SAT[51]**	Entropy (from Local energy matrices, calculated for sub-band 7)	0.069901974	0.122022261	-
F11	**BAR[5]**	Texture energy using ES kernel	**0.001455341**	**0.001712522**	**0.001762114**	F44	**SCH[52]**	Difference Histogram,	0.083531184	0.09977041	0.090448805
F12	**ACH[1]**	Moment 3	**0.001557873**	**0.003381957**	**0.006502389**	F45	**SCH[52]**	Difference of positive and negative (from Normalised Histograms)	0.0852041	0.107557706	**0.000670753**
F13	**CRZ[15]**	Number of pixels (segmented using Differential Evolution)	**0.001719172**	**0.000356737**	**0.008154521**	F46	**BAR[5]**	Angular sum from Fourier power spectrum	-	**4.29E-06**	**0.000114174**
F14	**KAR[30]**	Kurtosis(from Curvelet transformed)	**0.001733792**	**0.000955772**	**0.000222863**	F47	**BAR[5]**	Coarseness(Statistical feature matrix based)	-	**2.80144E-05**	**0.000349501**
F15	**ACH[1]**	Moment 4	**0.002101889**	**0.004432993**	**0.001562873**	F48	**BAR[5]**	Complexity(Neighborhood gray-tone difference matrix based)	-	**8.25699E-05**	**0.000438763**
F16	**ACH[1]**	Angular Second Moment	**0.003475957**	**0.000263501**	-	F49	**BAR[5]**	Energy (Gray level difference statistics based)	-	**0.000957879**	-
F17	**GUP[25]**	Correlation from GLCM Sub-Matrix	**0.003475957**	**0.005187852**	-	F50	**BAR[5]**	Entropy (Gray level difference statistics based)	-	**0.000993772**	-
F18	**GUP[25]**	Difference variance from GLCM Sub-Matrix	**0.004481993**	**0.008749024**	**0.002048889**	F51	**BAR[5]**	Homogeneity(Gray level difference statistics based)	-	**0.001426869**	-
F19	**ACH[1]**	Entropy	**0.006473389**	**0.004462993**	-	F52	**BAR[5]**	Mean (Gray level difference statistics based)	-	**0.001433341**	-
F20	**BAR[5]**	Texture energy using LL kernel	**0.006486389**	**0.000616753**	**0.00902678**	F53	**BAR[5]**	Periodicity(Statistical feature matrix based)	-	**0.000838435**	**0.000529959**
F21	**GUP[25]**	Energy from GLCM Sub-Matrix	0.014647123	**0.00899278**	-	F54	**BAR[5]**	Radial sum from Fourier power spectrum	-	**0.000934663**	-
F22	**GUP[25]**	Homogeneity from GLCM Sub-Matrix	0.015107258	-	-	F55	**BAR[5]**	Solidity	-	**0.001122531**	-
F23	**GUP[25]**	Information measure of correlation from GLCM Sub-Matrix	0.022027663	**0.001099846**	0.015121258	F56	**BAR[5]**	Texture energy using EE kernel	-	**0.00121398**	**0.005243872**
F24	**GUP[25]**	Otsu Threshold	0.023440631	0.029550677	0.022001663	F57	**BAR[5]**	Texture energy using SS kernel	-	**0.00058523**	-
F25	**GUP[25]**	Sum entropy from GLCM Sub-Matrix	0.024172069	0.030451145	0.023452631	F58	**BAR[5]**	Texture strength(Neighborhood gray-tone difference matrix based)	-	**0.001730656**	-
F26	**GUP[25]**	Sum variance from GLCM Sub-Matrix	0.027386128	0.031046951	0.024097069	F59	**ACH[1]**	Long Runs Emphasis	-	-	**0.00522876**
F27	**RAS[49]**	Difference between Cooccurrence levels based energy (from Modified LBP Colour matrix)	0.027408579	0.032323166	0.027390128	F60	**SCH[52]**	Distance of difference maximum from the centre(from Fourier Spectrum of left and right breast),	-	0.109284161	**0.00167705**
F28	**SAT[51]**	Skewness (from Local energy matrices, calculated for sub band 12)	0.027454128	-	-	F61	**SCH[52]**	Energy (Gray Level Cross co-occurrence matrix from Laplace filter applied images)	-	0.114465656	**0.001557873**
F29	**RAS[49]**	Difference between Cooccurrence levels based energy (from Modified LBP Edge matrix)	0.030428145	0.0352833	**5.89856E-05**	F62	**SCH[52]**	Mutual Information,	-	0.117180514	**0.008740024**
F30	**RAS[49]**	Difference of GLCM Matrixes	0.030987951	0.036041195	0.021953663	F63	**SAT[51]**	Kurtosis (from Local energy matrices, calculated for sub band 20)	-	-	**0.000233966**
F31	**SAT[51]**	Kurtosis (from Local energy matrices, calculated for sub band 12)	0.032310166	0.131642063	0.014568123	F64	**SAT[51]**	Skewness (from Local energy matrices, calculated for sub band 2)	-	-	**0.000250748**
F32	**RAS[49]**	Euclidean distance of histograms (from Modified LBP Colour matrix)	0.032340166	0.040762057	0.027479579	F65	**SAT[51]**	Variance (from Local energy matrices, calculated for sub band 9)	-	0.023478631	-
F33	**RAS[49]**	Euclidean distance of histograms (from Modified LBP Edge matrix)	0.0353093	0.044455997	-						

Just Accepted

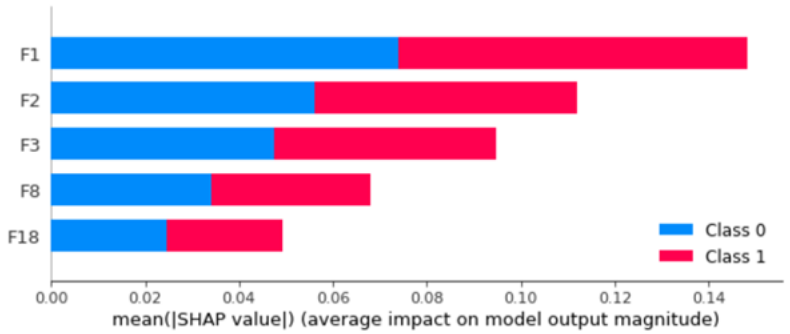
- F1 - Texture energy using LS kernel BAR[5](table 8) is the most influential feature for predicting the abnormality in DB-DMR-IR dataset. Features (table 8) F3 – Contrast (Gray level difference statistics based) BAR[5] and F29 - Difference between Co-occurrence levels based energy (from Modified LBP Edge matrix) RAS[49] are the most influential for DB-FOOT-IR and DB-THY-IR datasets respectively.
- Features (table 8) F3 - Contrast (Gray level difference statistics based) (BAR[5] and F18 - Difference variance from GLCM Sub-Matrix GUP[25] are significant features in all the models for predicting abnormality in the human body.
- Texture and Energy - based Features extracted from the ROI are observed to be more influential compared to the ones extracted from the entire thermal image for all the datasets.

Table 9. The colour map of the correlation matrix for the features with which the best result is obtained for DB-DMR-IR dataset.

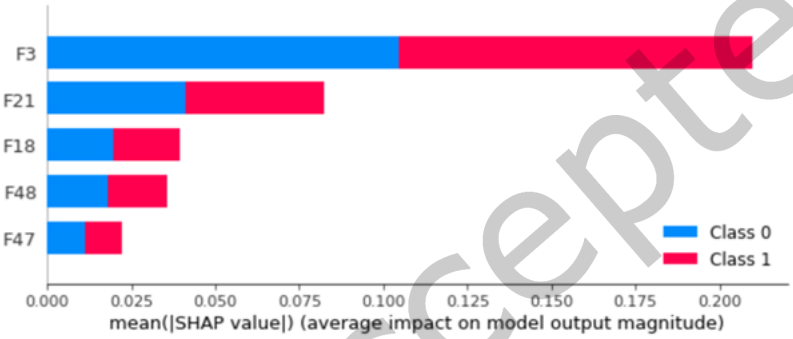


Just Accepted

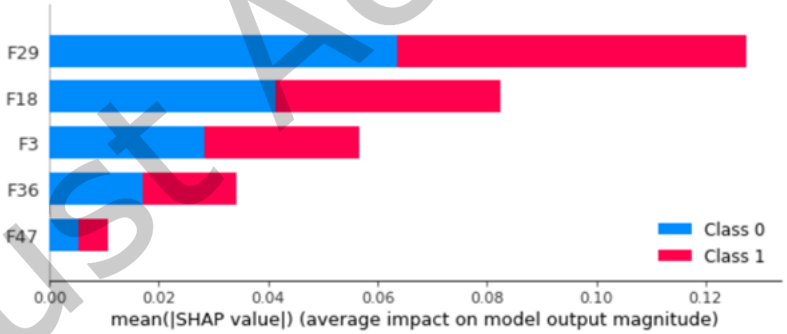
Thus, we obtain a feature subset from UnionFeature_Set (Table 8) that is most relevant and not redundant for diagnosing inflammation using Thermography. Also, the respective feature-sets effectively distinguish normal and abnormal thermal patterns for the datasets.



(a) DB-DMR-IR



(b) DB-FOOT-IR



(c) DB-THY-IR

Fig. 10. The absolute summary plot of the best performing models using UnionFeature_Set for the three datasets.

Just Accepted

6 CONCLUSION AND FUTURE WORK

Thermography-based technology for medical screening for various diseases, aided with machine learning, has been studied widely on private datasets. However, state-of-the-art methods are not compared using a common dataset, feature selection methods, classifiers, or a sampling strategy. Hence, it is not suitable to comment on the

approaches proposed in the literature. Motivated by this, we systematically compared nine feature extraction approaches on a platform using publicly available Datasets, Feature Selection mechanisms, Classifiers, and a 2-level sampling strategy. Also, we evaluated the union of feature sets proposed in the nine research works on the same platform.

Experimental results demonstrate that LOOCV gave the best results for all the ML pipelines considered. The classification accuracy of each feature extraction improved with the inclusion of FS method. Also, the combination of Feature-Sets (UnionFeature_Set) and any of the four FS methods gave a better performance in terms of average accuracy and net improvement in accuracy than individual feature extraction methods. The classification model with UnionFeature_Set gave the best performance for all the datasets. This is shown using several statistical tools. Finally, we obtained a feature-subset of 45, 57, and 39 features (from UnionFeature_Set) filtered by mRMR feature selection method for DB-DMR-IR and DB-FOOT-IR, and ReF FS method for DB-THY-IR, respectively. The feature-subset obtained for all the datasets are relevant, non-redundant, and distinguishes normal and abnormal thermal patterns with an accuracy of 94.75% on the DB-DMR-IR, 93.14% on DB-FOOT-IR, and 92.06% on DB-THY-IR dataset. In the future, we plan to develop a CAD-based system/model that segments the abnormal ROI and extracts robust features to screen abnormality/inflammation from thermal images of the human body. Also, we plan to develop a dataset of a considerable number of samples so that the Deep Learning model can be applied.

ACKNOWLEDGMENTS

We sincerely thank Dr. Satyavir Yadav (Assistant Professor, Department of Cardiology, All India Institute of Medical Sciences, Delhi, India) for reviewing the medical aspect of the paper and providing his valuable feedback for the same. Also, we thank Dr. Rishu Sangal, Radiologist, B. L. Kapoor - Max Super Speciality Hospital, India, for reviewing the manuscript and providing the feedback.

REFERENCES

[1] U Rajendra Acharya, Eddie Yin-Kwee Ng, Jen-Hong Tan, and S Vinitha Sree. 2012. Thermography based breast cancer detection using texture features and support vector machine. *Journal of medical systems* 36, 3 (2012), 1503–1510.

[2] Muhammad Adam, Eddie YK Ng, Shu Lih Oh, Marabelle L Heng, Yuki Hagiwara, Jen Hong Tan, Jasper WK Tong, and U Rajendra Acharya. 2018. Automated detection of diabetic foot with and without neuropathy using double density-dual tree-complex wavelet transform on foot thermograms. *Infrared Physics & Technology* 92 (2018), 270–279.

[3] MA Aweda, AO Adeyomoye, and GA Abe. 2012. Thermographic analysis of thyroid diseases. *Adv. Appl. Sci. Res* 3, 4 (2012), 2027–2032.

[4] S. Bagavathiappan, J. Philip, T. Jayakumar, B. Raj, P.N.S.R. Rao, M. Varalakshmi, and V. V. Mohan. 2010. Correlation between plantar foot temperature and diabetic neuropathy: a case study by using an infrared thermal imaging technique. *J. Diabetes Sci. Technol.* (2010), 1386–1392.

[5] Shawli Bardhan and Mrinal Kanti Bhowmik. 2019. 2-Stage classification of knee joint thermograms for rheumatoid arthritis prediction in subclinical inflammation. *Australasian physical & engineering sciences in medicine* 42, 1 (2019), 259–277.

[6] Shawli Bardhan, Mrinal Kanti Bhowmik, Satyabrata Nath, and Debotosh Bhattacharjee. 2015. A review on inflammatory pain detection in human body through infrared image analysis. In *2015 International Symposium on Advanced Computing and Communication (ISACC)*. 251–257. https://doi.org/10.1109/ISACC.2015.7377350

[7] OH Beahrs. 1979. Report to the working group to review the National Cancer Institute-American Cancer Society breast cancer demonstration projects. *J Natl Caner Inst* 62 (1979), 639–709.

[8] Christopher M. Bishop. 2006. *Pattern Recognition and Machine Learning*. Springer. /bib/bishop/Bishop2006/Pattern-Recognition-and-Machine-Learning-Christophe-M-Bishop.pdf,/bib/bishop/Bishop2006/978-0-387-31073-2_sm.pdf,https://www.microsoft.com/en-us/research/people/cmbishop/#!prml-book

[9] Tiago B. Borchartt, Aura Conci, Rita C.F. Lima, Roger Resmini, and Angel Sanchez. 2013. Breast thermography from an image processing viewpoint: A survey. *Signal Processing* 93, 10 (2013), 2785 – 2803. https://doi.org/10.1016/j.sigpro.2012.08.012 Signal and Image Processing Techniques for Detection of Breast Diseases.

[10] Leo Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone. 1984. *Classification and Regression Trees*. Wadsworth Publishing Company, Belmont, California, U.S.A.

Just Accepted

[11] Busra Can, Ozgur Kara, Muhammet Cemal Kizilarslanoglu, Gunes Arik, Gozde Sengul Aycicek, Fatih Sumer, Ramazan Civelek, Canan Demirtas, and Zekeriya Ulger. 2017. Serum markers of inflammation and oxidative stress in sarcopenia. *Aging clinical and experimental research* 29, 4 (2017), 745–752.

[12] Martha Rebeca Canales-Fiscal, Rocío Ortiz López, Regina Barzilay, Víctor Treviño, Servando Cardona-Huerta, Luis Javier Ramírez-Treviño, Adam Yala, and José Tamez-Peña. 2021. COVID-19 classification using thermal images: thermal images capability for identifying COVID-19 using traditional machine learning classifiers. In *Proceedings of the 12th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics*. 1–5.

[13] C. Cortes and V. Vapnik. 1995. Support-vector networks. *Machine Learning* 20 (1995), 273–297.

[14] T. Cover and P. Hart. 1967. Nearest neighbor pattern classification. *IEEE Transactions on Information Theory* 13, 1 (1967), 21–27. https://doi.org/10.1109/TIT.1967.1053964

[15] Israel Cruz-Vega, Hayde Peregrina-Barreto, Jose de Jesus Rangel-Magdaleno, and Juan Manuel Ramirez-Cortes. 2019. A comparison of intelligent classifiers of thermal patterns in diabetic foot. In *2019 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*. IEEE, 1–6.

[16] Nicholas A Diakides and Joseph D Bronzino. 2007. *Medical infrared imaging*. CRC press.

[17] Mahnaz Etehadtavakol, Zahra Emrani, and Eddie Yin Kwee Ng. 2019. Rapid extraction of the hottest or coldest regions of medical thermographic images. *Medical & biological engineering & computing* 57, 2 (2019), 379–388.

[18] Mahnaz Etehadtavakol and Eddie YK Ng. 2017. Potential of thermography in pain diagnosing and treatment monitoring. In *Application of infrared to biomedical sciences*. Springer, 19–32.

[19] Mahnaz Etehadtavakol, Eddie YK Ng, and Mohammad Hassan Emami. 2017. Potential of infrared imaging in assessing digestive disorders. In *Application of infrared to biomedical sciences*. Springer, 1–18.

[20] Oliver Faust, U Rajendra Acharya, EYK Ng, Tan Jen Hong, and Wenwei Yu. 2014. Application of infrared thermography in computer aided diagnosis. *Infrared Physics & Technology* 66 (2014), 160–175.

[21] Sheeja V Francis, M Sasikala, and S Saranya. 2014. Detection of breast abnormality from thermograms using curvelet transform based feature extraction. *Journal of medical systems* 38, 4 (2014), 23.

[22] Michel Gautherie. 1983. Thermobiological assessment of benign and malignant breast diseases. *American journal of obstetrics and gynecology* 147, 8 (1983), 861–869.

[23] Usha Rani Gogoi, Mrinal Kanti Bhowmik, Anjan Kumar Ghosh, Debotosh Bhattacharjee, and Gautam Majumdar. 2017. Discriminative feature selection for breast abnormality detection and accurate classification of thermograms. In *2017 International Conference on Innovations in Electronics, Signal Processing and Communication (IESC)*. IEEE, 39–44.

[24] José R González, Charbel Damião, and Aura Conci. 2017. An Infrared Thermal Images Database and a New Technique for Thyroid Nodules Analysis. *Studies in health technology and informatics* 245 (2017), 384—387. http://europepmc.org/abstract/MED/29295121

[25] Trasha Gupta, Rajni Jindal, and S Indu. 2020. Empirical Analysis of Thermography Effectiveness for Health Diagnosis. In *2020 International Conference on Computational Science and Computational Intelligence (CSCI)*. IEEE, 830–835.

[26] Peter E Hart, David G Stork, and Richard O Duda. 2000. *Pattern classification*. Wiley Hoboken.

[27] Daniel Hernández-Contreras, Hayde Peregrina-Barreto, Jose Rangel-Magdaleno, and Francisco Renero-Carrillo. 2019. Plantar Thermogram Database for the Study of Diabetic Foot Complications. (2019). https://doi.org/10.21227/tm4t-9n15

[28] H.Usuki, T.Ikeda, Y.Igarashi, I.Takahashi, A. Fukami, T.Yokoe, H. Sonoo, and K.Asaishi. 1998. What kinds of non-palpable breast cancer can be detected by thermography? *BiomedicalThermology* 4, 18 (1998), 8–12.

[29] Erik Ingelsson, Johan Ärnlöv, Johan Sundström, and Lars Lind. 2005. Inflammation, as measured by the erythrocyte sedimentation rate, is an independent predictor for the development of heart failure. *Journal of the American College of Cardiology* 45, 11 (2005), 1802–1806.

[30] R Karthiga and K Narasimhan. 2021. Medical imaging technique using curvelet transform and machine learning for the automated diagnosis of breast cancer from thermal image. *Pattern Analysis and Applications* (2021), 1–11.

[31] Kosar Khaksari, Thien Nguyen, Brian Y Hill, Timothy Quang, John Perrault, Viswanath Gorti, Ravi Malpani, Emily Blick, Tomas Gonzalez Cano, Babak Shadgan, et al. 2021. Review of the efficacy of infrared thermography for screening infectious diseases with applications to COVID-19. *Journal of Medical Imaging* 8, S1 (2021), 010901.

[32] Kenji Kira, Larry A Rendell, et al. 1992. The feature selection problem: Traditional methods and a new algorithm. In *Aaai*, Vol. 2. 129–134.

[33] Lauren N Ko, Adam B Raff, Anna C Garza-Mayers, Allison S Dobry, Antonio Ortega-Martinez, R Rox Anderson, and Daniela Kroshinsky. 2018. Skin surface temperatures measured by thermal imaging aid in the diagnosis of cellulitis. *Journal of Investigative Dermatology* 138, 3 (2018), 520–526.

[34] Lauren N Ko, Adam B Raff, Anna C Garza-Mayers, Allison S Dobry, Antonio Ortega-Martinez, R Rox Anderson, and Daniela Kroshinsky. 2018. Skin surface temperatures measured by thermal imaging aid in the diagnosis of cellulitis. *Journal of Investigative Dermatology* 138, 3 (2018), 520–526.

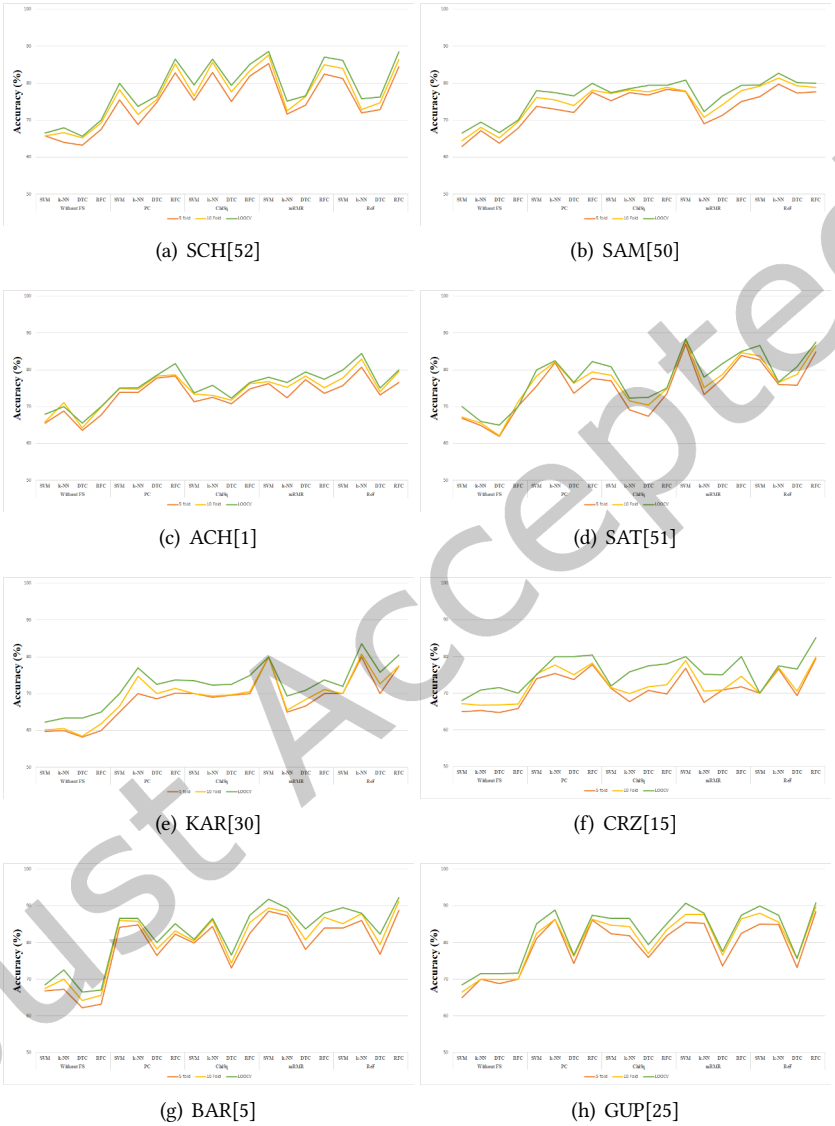
[35] Igor Kononenko. 1994. Estimating attributes: Analysis and extensions of RELIEF. In *European conference on machine learning*. Springer, 171–182.

Just Accepted

[36] R. Lawson. 1956. Implications of surface temperatures in the diagnosis of breast cancer. (1956), 309–310.
[37] Huan Liu and Rudy Setiono. 1995. Chi2: Feature selection and discretization of numeric attributes. In *Proceedings of 7th IEEE International Conference on Tools with Artificial Intelligence*. IEEE, 388–391.
[38] Haochen Liu, Yiqi Wang, Wenqi Fan, Xiaorui Liu, Yaxin Li, Shaili Jain, Yunhao Liu, Anil Jain, and Jiliang Tang. 2022. Trustworthy AI: A Computational Perspective. *ACM Trans. Intell. Syst. Technol.* 14, 1, Article 4 (nov 2022), 59 pages. https://doi.org/10.1145/3546872
[39] Scott M. Lundberg, Gabriel Erion, Hugh Chen, Alex DeGrave, Jordan M. Prutkin, Bala Nair, Ronit Katz, Jonathan Himmelfarb, Nisha Bansal, and Su-In Lee. 2020. From local explanations to global understanding with explainable AI for trees. *Nature Machine Intelligence* 2, 1 (2020), 2522–5839.
[40] Scott M Lundberg and Su-In Lee. 2017. A Unified Approach to Interpreting Model Predictions. In *Advances in Neural Information Processing Systems 30*, I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett (Eds.). Curran Associates, Inc., 4765–4774. http://papers.nips.cc/paper/7062-a-unified-approach-to-interpreting-model-predictions.pdf
[41] G Machin, A Whittam, S Ainarkar, J Allen, J Bevans, M Edmonds, B Kluwe, A Macdonald, N Petrova, P Plassmann, et al. 2017. A medical thermal imaging device for the prevention of diabetic foot ulceration. *Physiological measurement* 38, 3 (2017), 420.
[42] Eddie YK Ng and Mahnaz Etehadtavakol. 2017. *Application of infrared to biomedical sciences*. Springer.
[43] Kalliopi Pafili and Nikolaos Papanas. 2015. Thermography in the follow up of the diabetic foot: best to weigh the enemy more mighty than he seems.
[44] KARL PEARSON. 1920. NOTES ON THE HISTORY OF CORRELATION. *Biometrika* 13, 1 (10 1920), 25–45. https://doi.org/10.1093/biomet/13.1.25
[45] Hanchuan Peng, Fuhui Long, and Chris Ding. 2005. Feature Selection Based on Mutual Information: Criteria of Max-Dependency, Max-Relevance, and Min-Redundancy. *IEEE Trans. Pattern Anal. Mach. Intell.* 27, 8 (Aug. 2005), 1226–1238. https://doi.org/10.1109/TPAMI.2005.159
[46] Stephen P Power, Fiachra Moloney, Maria Twomey, Karl James, Owen J O'Connor, and Michael M Maher. 2016. Computed tomography and patient risk: facts, perceptions and uncertainties. *World journal of radiology* 8, 12 (2016), 902.
[47] Jose Ignacio Priego Quesada, Marcos Roberto Kunzler, and Felipe P Carpes. 2017. Methodological aspects of infrared thermography in human assessment. In *Application of infrared thermography in sports science*. Springer, 49–79.
[48] Hairong Qi, Phani Teja Kuruganti, and Wesley E Snyder. 2012. Detecting breast cancer from thermal infrared images by asymmetry analysis. *Medicine and Medical Research* 38 (2012).
[49] Rozita Rastghalam and Hossein Pourghassem. 2016. Breast cancer detection using MRF-based probable texture feature and decision-level fusion-based classification using HMM on thermography images. *Pattern Recognition* 51 (2016), 176–186.
[50] J Saminathan, M Sasikala, VB Narayanamurthy, K Rajesh, and R Arvind. 2020. Computer aided detection of diabetic foot ulcer using asymmetry analysis of texture and temperature features. *Infrared Physics & Technology* 105 (2020), 103219.
[51] Dayakshini Sathish, Surekha Kamath, Keerthana Prasad, and Rajagopal Kadavigere. 2019. Role of normalization of breast thermogram images and automatic classification of breast cancer. *The Visual Computer* 35, 1 (2019), 57–70.
[52] Gerald Schaefer, Michal Závišek, and Tomoharu Nakashima. 2009. Thermography based breast cancer analysis using statistical features and fuzzy classification. *Pattern Recognition* 42, 6 (2009), 1133–1137.
[53] G. Serbu. 2009. Infrared imaging of the diabetic foot. In *In: Proceedings on InfraMation, Romanian Society of Thermography*. 5–20.
[54] LF Silva, DCM Saade, GO Sequeiros, AC Silva, AC Paiva, RS Bravo, and Aura Conci. 2014. A new database for breast research with infrared image. *Journal of Medical Imaging and Health Informatics* 4, 1 (2014), 92–100.
[55] Valmir Oliveira Silvino, Regis Bernardo Brandim Gomes, Sérgio Luiz Galan Ribeiro, Davyson de Lima Moreira, and Marcos Antonio Pereira dos Santos. 2020. Identifying febrile humans using infrared thermography screening: Possible applications during COVID-19 outbreak. *Revista Contexto & Saúde* 20, 38 (2020), 5–9.
[56] Vera A van Atteveld, Jeanine M Van Ancum, Esmee M Reijnierse, Marijke C Trappenburg, Carel GM Meskers, and Andrea B Maier. 2019. Erythrocyte sedimentation rate and albumin as markers of inflammation are associated with measures of sarcopenia: a cross-sectional study. *BMC geriatrics* 19, 1 (2019), 1–8.
[57] Ricardo Vardasca, Carolina Magalhaes, and Joaquim Mendes. 2019. Biomedical applications of infrared thermal imaging: current state of machine learning classification. In *Multidisciplinary Digital Publishing Institute Proceedings*, Vol. 27. 46.
[58] Valanarasi Antony Santiagu Vaz. 2014. Diagnosis of hypo and hyperthyroid using MLPN network. *International Journal of Innovative Research in Science, Engineering and Technology* 3, 7 (2014), 14314–14323.

Just Accepted

A VISUAL COMPARISON OF ACCURACY FOR A COMBINATION OF VARIOUS FEATURE EXTRACTION METHODS, FS METHODS, CLASSIFIERS AND CROSS-VALIDATION APPROACHES.



(a) SCH[52] (b) SAM[50]

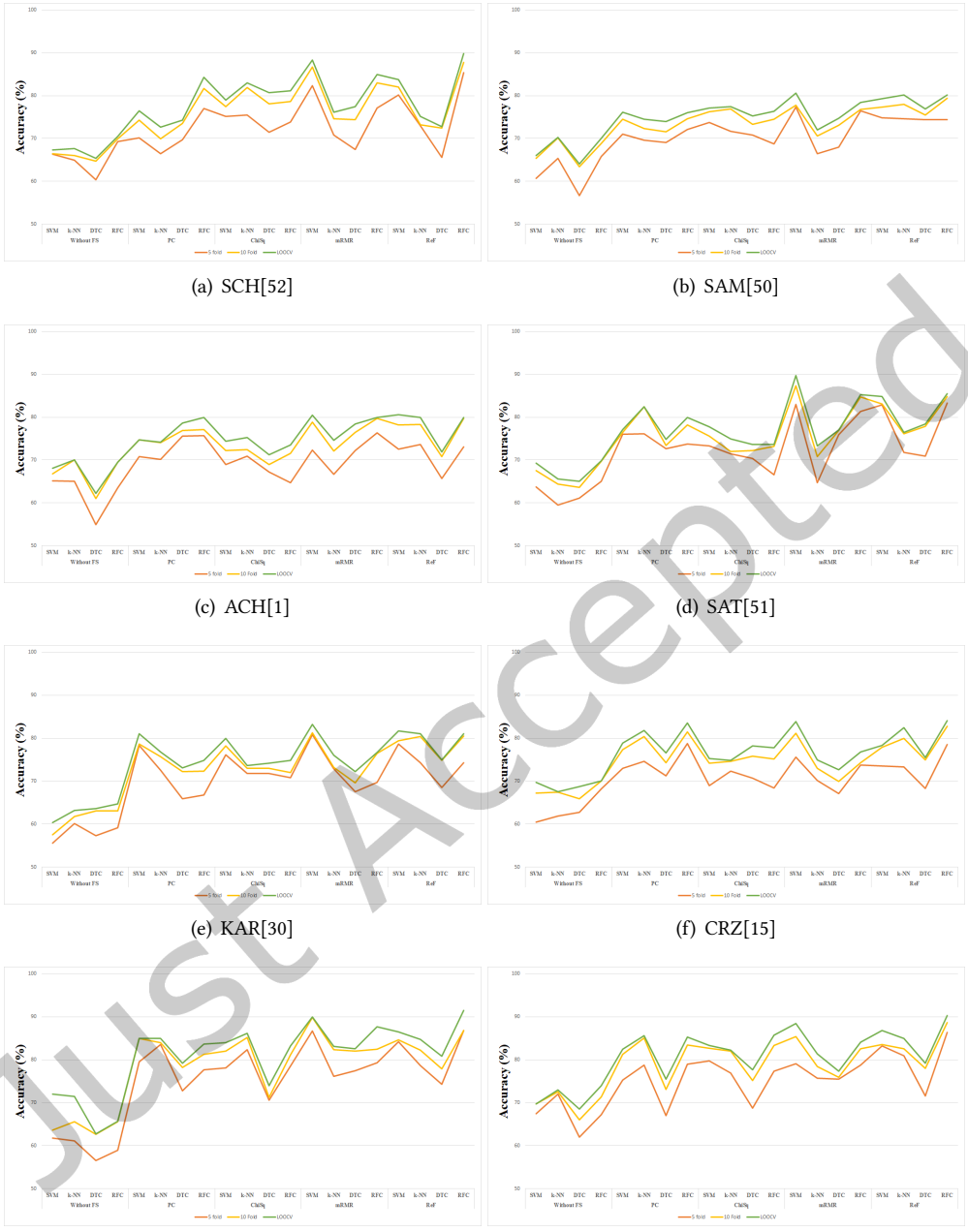
(c) ACH[1] (d) SAT[51]

(e) KAR[30] (f) CRZ[15]

(g) BAR[5] (h) GUP[25]

Fig. 11. Comparison of classification accuracy values for a given combination of feature extraction, feature selection methods and cross-validation strategies on DB-DMR-IR dataset.

Just Accepted



(a) SCH[52] (b) SAM[50]

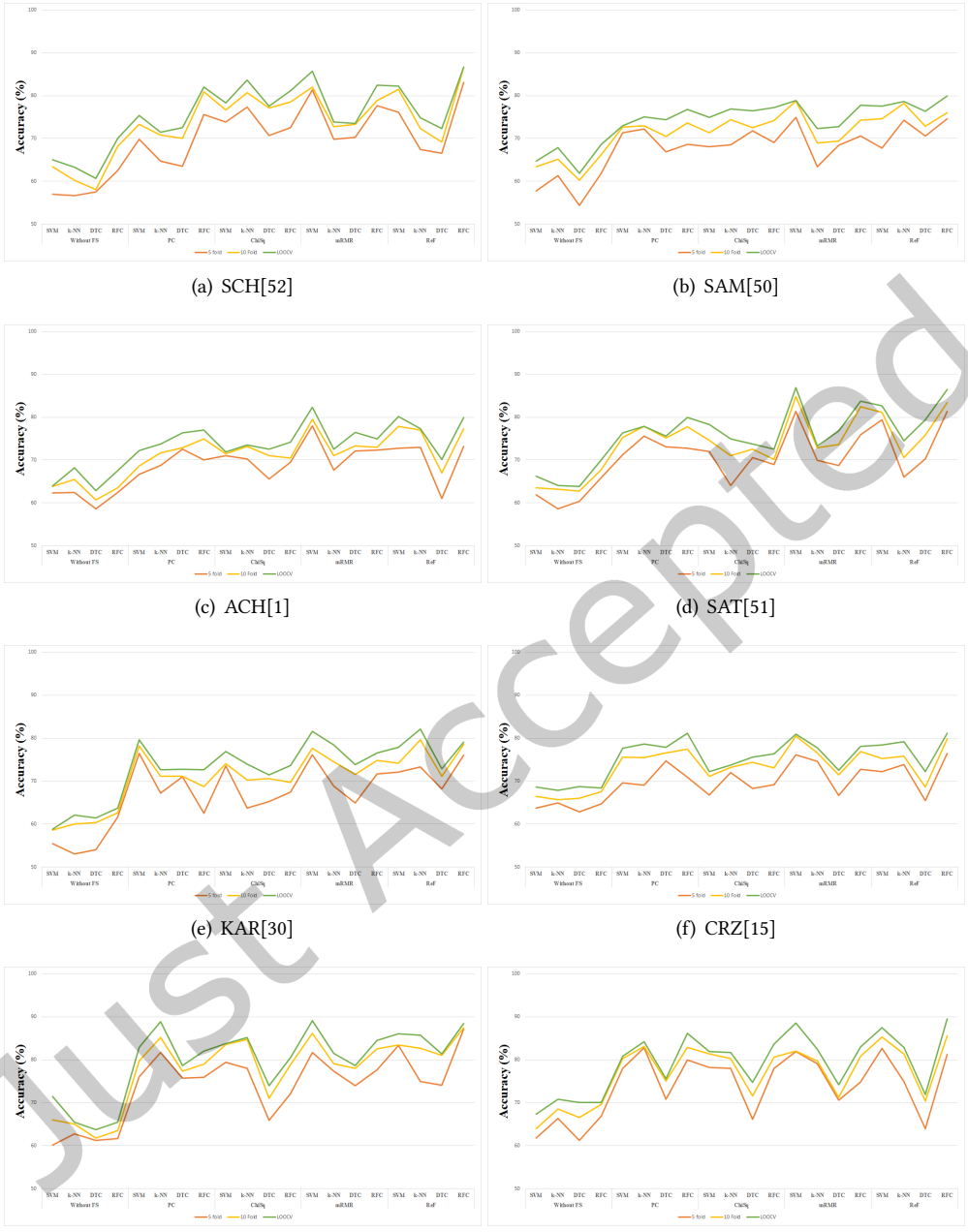
(c) ACH[1] (d) SAT[51]

(e) KAR[30] (f) CRZ[15]

(g) BAR[5] (h) GUP[25]

Fig. 12. Comparison of classification accuracy values for a given combination of feature extraction, feature selection methods and cross-validation strategies on DB-FOOT-IR dataset.

Just Accepted



(a) SCH[52] (b) SAM[50]

(c) ACH[1] (d) SAT[51]

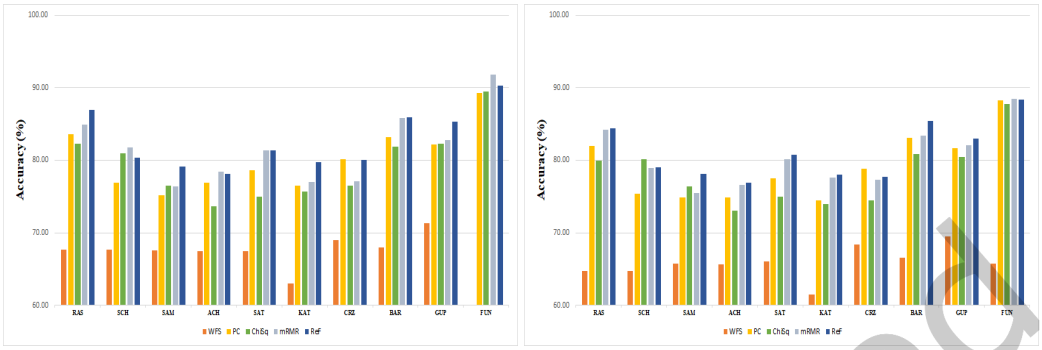
(e) KAR[30] (f) CRZ[15]

(g) BAR[5] (h) GUP[25]

Fig. 13. Comparison of classification accuracy values for a given combination of feature extraction, feature selection methods and cross-validation strategies on DB-THY-IR dataset.

Just Accepted

B COMPARISON OF COMBINATION OF FEATURE EXTRACTION AND SELECTION METHODS



(a) (b)

Fig. 14. Comparison of the average classification accuracy values for a given combination of feature extraction and selection method for DB-FOOT-IR (a) and DB-THY-IR (b) datasets.

(a) (b)

Fig. 15. Ranking of combinations of feature extraction and selection methods on the basis of Performance Improvement (%) for DB-FOOT-IR (a) and DB-THY-IR (b) datasets.

Just Accepted

C THE COLOUR MAP OF CORRELATION MATRIX

Just Accepted

Table 10. The colour map of correlation matrix for the features with which the best result is obtained for DB-FOOT-IR dataset.

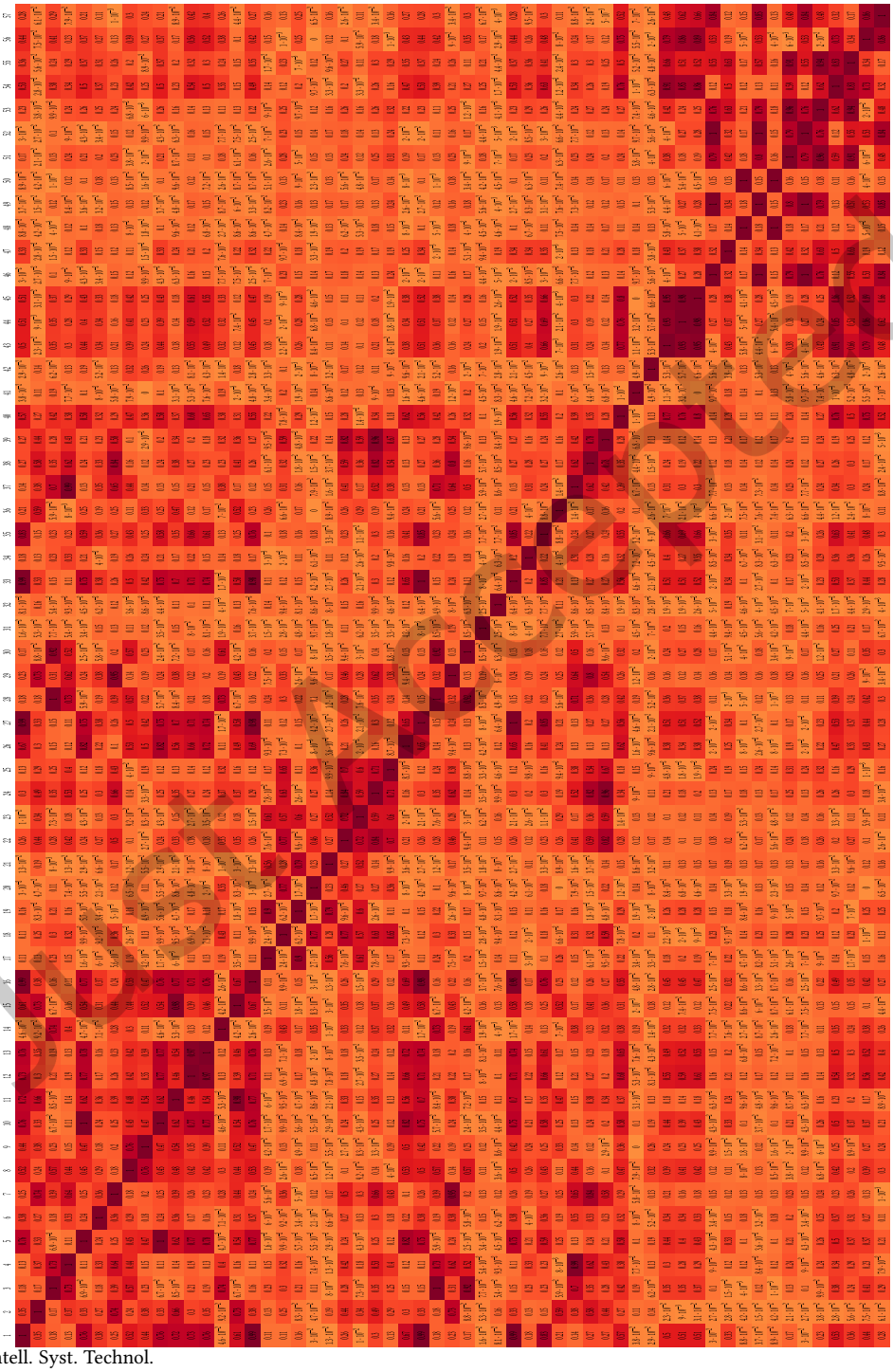
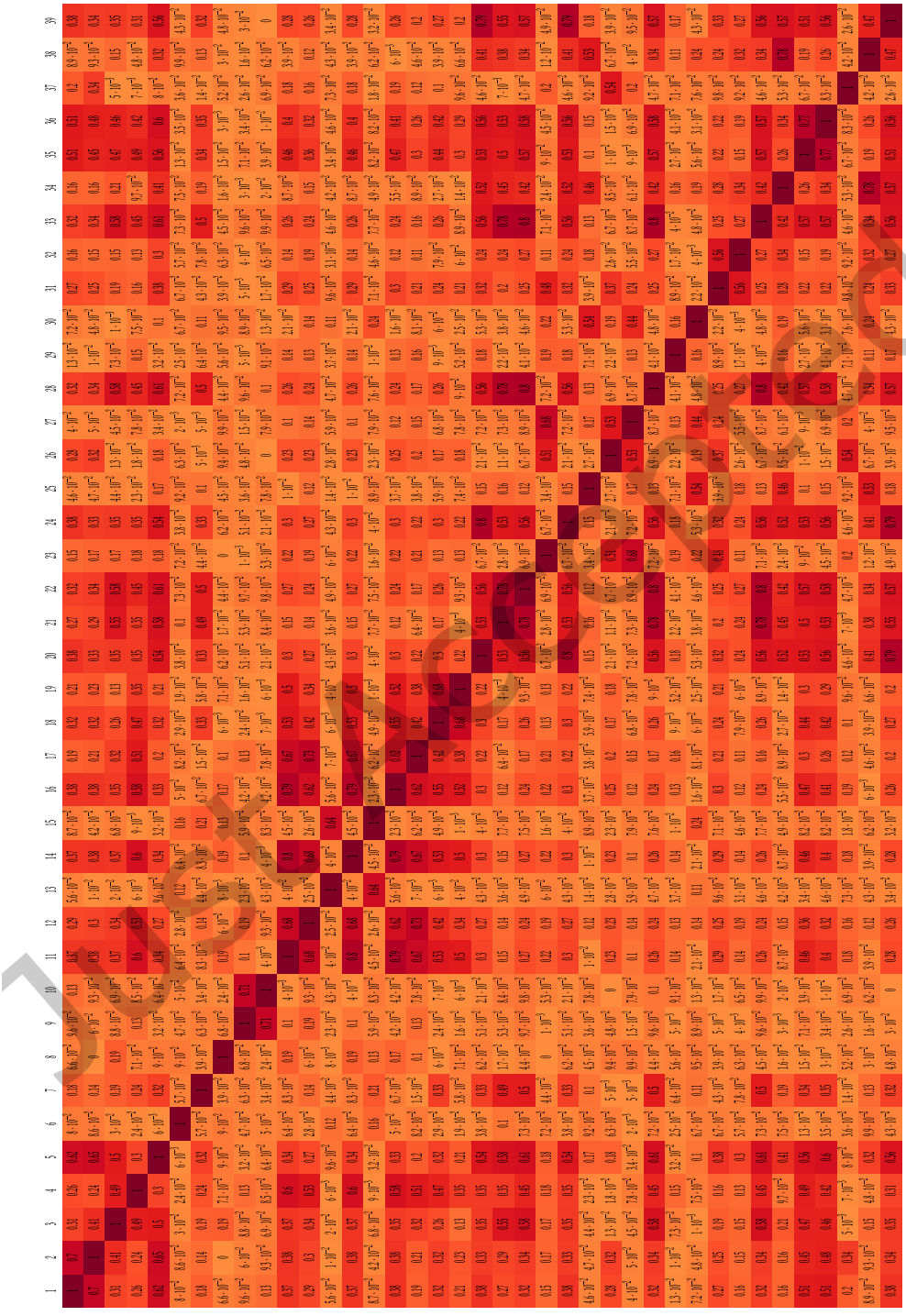


Table 11. The colour map of correlation matrix for the features with which the best result is obtained for DB-THY-IR dataset.



Unpublished

2023 15th International Conference on COMmunication Systems & NETworkS (COMSNETS) | 978-1-6654-7706-2/23/$31.00 ©2023 IEEE | DOI: 10.1109/COMSNETS56262.2023.10041340

Energy Efficient Multiple Visible Light Communication Access Points (VAP)

Vishal Kumar
Delhi Technological University, New Delhi
vishalkumar_2k21phdec07@dtu.ac.in

Rohit Kumar
Delhi Technological University, New Delhi
rohit.kumar@dtu.ac.in

I. PROBLEM STATEMENT

The rise in various activities such as the internet of things, gaming platforms, and online teaching platforms have led to expanding demands for boundless data in wireless communication. This in turn stipulates a broad range of spectrum and high data rates. Being a traditional way of communication, the radio frequency (RF) spectrum is overburdened day by day. Amidst the innumerable results, visible light communication (VLC) has been considered as a reassuring wireless technique to acknowledge the insufficiency of RF spectra, due to its superiority in transmission capacity, channel capacity, and link safety. It is one of the best solutions for indoor communication in the existing literature. In spite of its several advantages, VLC suffers from a major challenge of the energy efficiency of VAP. In VLC technology, the main aim is to employ visible light for communication as well as luminosity. Thus, there must be proper trade-off between these two. In this article, we propose an energy-efficient VLC system based on the strategic sleeping of VAP. Herein, multiple access points (APs) are considered, and based on the proposed methodology, theoretical and numerical analysis is done for energy consumption. Further, the ideal probability of the VAP is derived and analyzed from which it is inferred that it is needless for all the VAPs to be active all the time for communication with the User equipment (UE) at a specific position. Additionally, a comparison between Multiple VAP and Single VAP is carried out.

II. STATE OF THE ARTS

Approximately 66 % of the total power is apprehended to be consumed by the nodes at an idle state in wireless communication [1]. That shows a high amount of energy is underutilized or wasted. One of the recent technology to reduce this power consumption substantially is by utilizing a Wake-up system [2]. In [3], an algorithm is proposed in order to reduce the overall power expenditure by putting the small base stations (BS) in sleeping or energy-saving mode depending upon the demand of traffic for static traffic and dynamic traffic models. In [4], techniques for opportunistic sleep mode have been proposed. The suggested proposal expedites the minuscule cell BSs in RF communication to escalate their downlink capability by an appropriate trade-off between quality of service (QOS) exigency of users and well-organized stability of load amid various base stations. In these two works, sleeping techniques are used for RF technology. In [5], strategic sleeping for VLC access points is proposed for single VAP and single UE. In a single VAP, as the UE goes away from the VAP inside the room, the signal suffers from fading. To reduce the effect of fading in this article we have proposed multiple VAPs. In areas such as hospitals, RF is prohibited as it can cause a serious threat to patients, this proposed energy-efficient VLC system can be utilized. Other major applications lie in mines or airplanes where an energy-efficient system is of utmost importance.

III. PROPOSED SOLUTION APPROACH

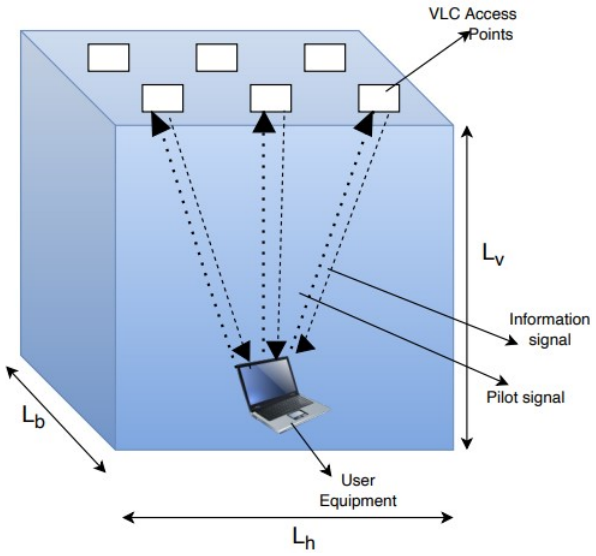


Fig. 1. System Model

In Fig.(1), we demonstrate the overall proposed system model. Here, 'k' number of VAPs are embedded in the ceiling of the room. The UE can be anywhere in the room. When UE wants to access a VAP then it broadcast a pilot signal (PS) to all the VAPs. The UE has the capability to track the signal-to-noise ratios (SNRs) of all the signals received from all the VAPs and hold them for a few seconds and choose the VAP whose signal wavelength is associated with the highest SNR. After receiving the PS, the VAP comes out of the idle state and starts transmitting the data. Once the connection is established rest all VAPs should be put in sleeping mode. One of the

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 02,2023 at 06:45:49 UTC from IEEE Xplore. Restrictions apply.

challenges in Multiple access points is signal interference due to the 'k' number of VAPs. To mitigate the effect of signal interference, the signal modulated at each VAP is dedicated to a different wavelength of visible light.

TABLE I: Notations and Descriptions

Notations	Descriptions
l_v	Vertical distance between plane of VAP and UE
v_k	Horizontal distance between k^{th} VAP and the origin
μ_v	Node density of VAP
μ_u	Node density of UE
$U(\beta)$	Optical filter gain
$G(\beta)$	Optical concentrator gain
A	Physical area of the PD
R_p	Responsivity of the PD
β	Angle of irradiance
N	Internal refractive index
ψ_c	FOV at semi-angle of the concentrator
$\phi_{1/2}$	LED's semi-angle at half power
ϑ_k	Indicator of the transmission condition of VAP
P_a	Power consumed by that VAP which has the highest SNR
PL	Power consumed to send the pilot signals
$\hbar$	Coverage area of VAP
$\mu(\hbar)$	Standard Lebesgue measure
$1_{\hbar}(v_k)$	Metric of random counting
δ	Shape parameter

Multiple VAP's horizontal placements are modeled using a 2D homogeneous Poisson point process (PPP). Here, We assume that VAP is at a vertical distance of 'l_v' from the plane of UE. This PPP model is expressed as

$$\Psi_v = \{v_k, k \in \mathbb{N}\} \subset \mathbb{R}^2 \tag{1}$$

Here, 'v_k' is the horizontal separation between k^{th} VAP and the origin. Node density of VAP and UE is denoted as 'μ_v' and 'μ_u' respectively. DC channel gain of VLC system is given by [6] as

$$h_{VLC} = \frac{A(m+1)R_p}{2\pi d^2}\cos^m(\alpha)U(\beta)G(\beta)\cos(\beta) \tag{2}$$

where, '$U(\beta)$' is optical filter gain, '$G(\beta)$' is optical concentrator gain, 'A' is the physical area of the photodetector (PD), 'R_p' is responsivity of the PD, 'd' is the Euclidean distance between VAP and UE, 'β' is the angle of irradiance for VLC system.

$$m = \frac{-\ln 2}{\ln\left(\cos\left(\phi_{1/2}\right)\right)} \tag{3}$$

$$G(\beta) = \begin{cases} \frac{N^2}{\sin^2(\psi_c)} & ;0 \le \beta \le \psi_c \\ 0 & ;\beta > \psi_c \end{cases} \tag{4}$$

where, internal refractive index is denoted by 'N', 'ψ_c' is field of view (FOV) at semi-angle of the concentrator, '$\phi_{1/2}$' is the LED's semi-angle at half power. By utilizing the geometrical feature of the VLC link [7],

$$\cos(\alpha) = \frac{l_v}{d} = \cos(\beta) \tag{5}$$

$$d = \sqrt{l_v^2 + v_k^2} \tag{6}$$

'ϑ_k' is the indicator of the transmission condition of VAP such that $\vartheta_k = 1$ indicates that VAP is active and transmission took place, $\vartheta_k = 0$ indicates VAP's sleep or idle state. Power consumed by VAP is denoted as P_{VAP} and expressed as

$$P_{VAP} = \begin{cases} P_a + PL & ;\vartheta_k = 1 \\ PL & ;\vartheta_k = 0 \end{cases} \tag{7}$$

where, P_a is the power consumed by that VAP which has the highest SNR i.e. γ_{max} and connected for the communication, PL is the power consumed to send the pilot signals. It is determined by idle probability that the VAP will be in an idle state when no UEs are making data traffic requests so that VLC AP can be decisively shut off to save energy. Let's say $\hbar \subset \mathbb{R}^2$ is the coverage area of k^{th} VAP. Probability of n UE inside $\hbar$ is

$$P\left[\sum_{v_k} 1_{\hbar}(v_k) = n\right] = \mathbb{E}\left[\frac{(\mu_u\mu(\hbar))^n}{n!}\exp\{-\mu_u\mu(\hbar)\}\right] \tag{8}$$

where, $\mu(\hbar)$ is the standard Lebesgue measure of $\hbar$, $1_{\hbar}(v_k)$ is metric of random counting of $\hbar$ and is expressed

$$1_{\hbar}(v_k) = \begin{cases} 1 & ;v_k \in \hbar \\ 0 & ;\text{otherwise} \end{cases} \tag{9}$$

The probability density function (PDF) of '$\mu(\hbar)$' with 'δ' as shape parameter is given as [7]

$$f_{\mu(\hbar)}(t) = \frac{(\delta\mu_v)^{\delta}}{\Gamma(\delta)}t^{\delta-1}\exp\{-\delta\mu_v t\} \tag{10}$$

The final expression for the probability of 'n' UEs present in the coverage area of VAP is derived

$$P\left[\sum_{v_k} 1_{\hbar}(v_k) = n\right] = \frac{\Gamma(n+\delta)}{n!\Gamma(\delta)}\left(\frac{\rho}{\delta+\rho}\right)^n\left(\frac{\delta}{\delta+\rho}\right)^{\delta} \tag{11}$$

Here, ρ is density ratio and given as $\rho = \frac{\mu_u}{\mu_v}$. Idle probability can be obtained by putting $n = 0$ and hence expressed as

$$P_{idle} = \left(\frac{\delta}{\delta+\rho}\right)^{\delta} \tag{12}$$

IV. PRELIMINARY RESULTS

Discussions of numerical and simulation results are presented in this section. Dimension of the room is considered to be of $l_v \times l_h \times l_b$ where specifically we have taken dimension as $(4\times10\times10)\ m^3$. δ is taken as 3.5 in this work. The transmitted power of one LED is $20mW$ and 40×40 LEDs array is used. Semi-angle at half power and FOV of the PD of our VLC

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 02,2023 at 06:45:49 UTC from IEEE Xplore. Restrictions apply.

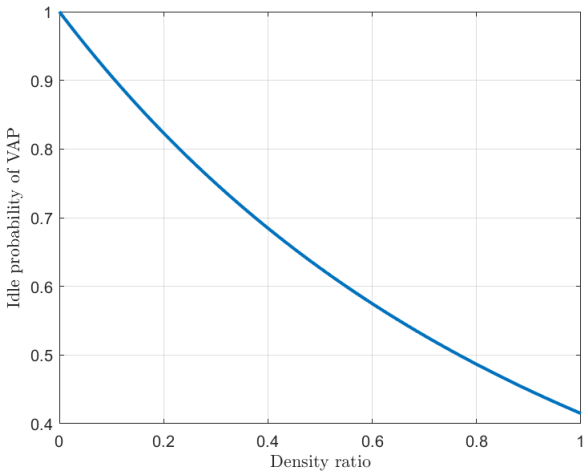


Fig. 2. Characteristic of Idle probability

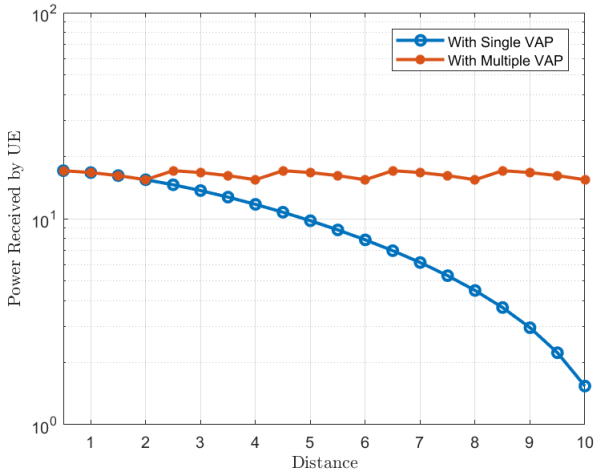


Fig. 4. Comparison between Multiple VAP and Single VAP

system is 70^o. From Fig. (2), it can be clearly depicted that with increment in the number of UEs in a coverage region of a specific VAP, that means the density ratio increases, and hence Probability of a VAP being idle reduces. The simulation results for energy consumption for the proposed Multiple VAPs with and without Strategic sleeping are shown in Fig.(3). From this graph, it is clearly evident that we can save a lot of energy by utilizing the proposed approach. Further, we try

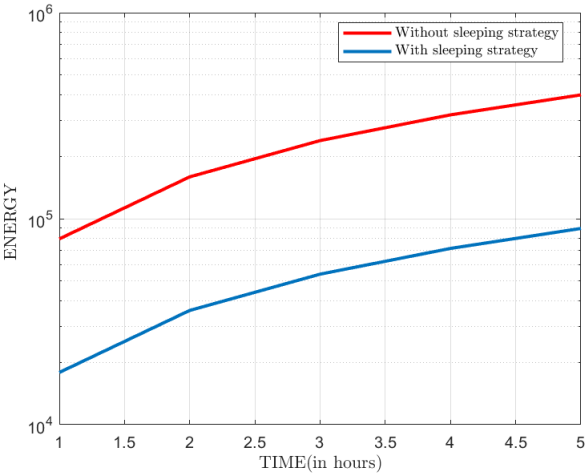


Fig. 3. Amount of Energy saving as a function of time

to analyze the impact of power received by UE at different positions for single VAP and multiple VAPs. From Fig. (4), we can observe the power received at different positions by UE for single VAP reduces while for multiple VAPs it remains uniform throughout the dimension of our room.

V. FUTURE WORK

This work can be extended by analyzing the different parameters like throughput and latency. As, we need to take care that while changing the state of a VAP, there should not be a loss of data. Re-configurable Intelligent Surface (RIS) can be incorporated to reduce the fading effect of signal when UE is mobile. While analyzing the energy efficiency, we can analyze the security performance of the system. Machine Learning can also be utilized to make the proposed system more reliable.

VI. CONCLUSION

In this paper, an energy-efficient Multiple VAP system is proposed. Through the analysis of idle probability, it is shown that there is a probability of multiple VAPs being in an idle state. So, we can switch it OFF or put it in an energy-saving mode to save a lot of energy. One advantage of the proposed work lies in the application of multiple VAPs to reduce signal fading.

REFERENCES

[1] Key World Energy Statistics 2014, International Energy Agency, 2014.
[2] Oller, J., Demirkol, I., Casademont, J., Paradells, J., Gamm, G.U. and Reindl, L., 2013, November. Wake-up radio as an energy-efficient alternative to conventional wireless sensor networks MAC protocols. In Proceedings of the 16th ACM international conference on Modeling, analysis simulation of wireless and mobile systems (pp. 173-180).
[3] C. Liu, B. Natarajan and H. Xia, "Small Cell Base Station Sleep Strategies for Energy Efficiency," in IEEE Transactions on Vehicular Technology, vol. 65, no. 3, pp. 1652-1661, March 2016, doi: 10.1109/TVT.2015.2413382.
[4] S. Samarakoon, M. Bennis, W. Saad and M. Latva-aho, "Opportunistic sleep mode strategies in wireless small cell networks," 2014 IEEE International Conference on Communications (ICC), 2014, pp. 2707-2712, doi: 10.1109/ICC.2014.6883733. S. Samarakoon, M. Bennis, W. Saad and M. Latva-aho, "Opportunistic sleep mode strategies in wireless small cell networks," , 2014, pp. 2707-2712, doi: 10.1109/ICC.2014.6883733..
[5] V. Kumar, A. Gupta and P. Garg, "Strategic Sleeping of Visible Light Communication Access Point," 2021 6th International Conference on Signal Processing, Computing and Control (ISPCC), 2021, pp. 387-392, doi: 10.1109/ISPCC53510.2021.9609445.
[6] L. Yin, W. O. Popoola, X. Wu and H. Haas, "Performance Evaluation of Non-Orthogonal Multiple Access in Visible Light Communication," in IEEE Transactions on Communications, vol. 64, no. 12, pp. 5162-5175, Dec. 2016, doi: 10.1109/TCOMM.2016.2612195.
[7] Ferenc, Járai-Szabó, and Zoltán Néda. "On the size distribution of Poisson Voronoi cells." Physica A: Statistical Mechanics and its Applications 385, no. 2 (2007): 518-526.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 02,2023 at 06:45:49 UTC from IEEE Xplore. Restrictions apply.

2023 13th International Conference on Cloud Computing, Data Science & Engineering (Confluence) | 978-1-6654-6263-1/23/$31.00 ©2023 IEEE | DOI: 10.1109/Confluence56041.2023.10048836

Estimating Medical Insurance Cost using Linear Regression with HyperParameterization, Decision Tree and Random Forest Models

Sanjay Patidar
Department of Software Engineering
Delhi Technological University
New Delhi, India
sanjaypatidar@dtu.ac.in

Saksham Dudi
Department of Software Engineering
Delhi Technological University
New Delhi, India
sakshamdudi@gmail.com

Rohit
Department of Software Engineering
Delhi Technological University
New Delhi, India
rohityadavry885@gmail.com

***Abstract*—The total health expenditure refers to the total public and private funds spent on health services and amenities, medical and surgical bills and all other healthcare facilities provided. The financing for health is of great significance and plays a crucial role in health systems. To enhance the productivity of human capital, the efficiency and delivery of healthcare services must be uplifted. Reports have shown that from year 2000 to 2018 there had been a gradual escalation of global health expenses and is standing on around 10% of the total GDP of the world. Out of pocket expenses are also high in least developed nations that have lower per capita income. Even though the World Health Organization (WHO) sanctions loans to these countries, these nations are bound to use the money on industrialization only and not their healthcare, education and public welfare sectors. With inflation, the expenses of first-rate healthcare are also rising which makes it fundamental to have health and life insurance plans. Health insurance schemes insured around 514 million people in India in the year 2021, most of which were covered under government schemes only. Since the advent of COVID-19 people have realized the need for having a insurance plan. Most of the companies that are based on the health insurance sector use predictive modelling to improve their services and business process. Machine Learning (ML) algorithms are used to train a model and provide insurance costs estimations. Past data is searched for any pattern or trend in the behaviour history of consumers and then future estimations are evaluated. The proposed project is comprised of different regression models like Linear regression with hyperparameterization , regressors like Decision Tree and Random forest to estimate the approximate insurance expenditure.**

***Index Terms*—Machine Learning, Regression Models, Linear Regression, HyperParameterization, Decision Tree, Random Forest, Insurance**

I. Introduction

Human health plays a vital role and is of significant importance in our society, our country and the world. It has always been put at risk when inevitable incidents such as global pandemics or epidemics occur, which can do unimaginable and unforeseen damage to individuals and their families. Recently, COVID-19 pandemic struck people in the world and caused a scarring destruction which no one was ready for. The high medical expenses also took toll on individuals and their families let alone the disease itself. This showed people that no one can predict medical exigencies and the huge financial upheaval can be caused which will be difficult to deal with. A health insurance plan would prove beneficial in times like these. Such an insurance entitles the policyholder to reimbursement for their surgical and medical charges.

World Health Organization (WHO) at [9] recommends that a country should spend at least 5% of their GDP on their health infrastructure. But, in contrast to other countries, out of the total annual GDP, Indian government at [8] only spent roughly around 3% on public healthcare in the financial year 2021-22. Providing good medical facilities with the medical expenses going high can be a thing to worry about for most of the people. Besides that, people themselves should have the idea of the approximate cost for their health insurance beforehand as discussed in [10]. There is a dire need for people to have a health insurance cover for themselves and their families which will provide a financial backup whenever needed. This paper provides a good estimation of the amount for a single person.

Nowadays, the system of health insurance uses Machine Learning (ML) models for estimation of the cost and business up-swinging. Models are trained using the insurance data from the past using different datasets. Different algorithms that are used in those ML models study the historical data and provide the results following the trend shown in the data. We have taken different parameters into consideration like age, BMI, smoker, children, charges, gender and region. Health insurance cost estimation system undergo through 4 different processes which are collection of dataset, pre-processing of data, importance of features, training and testing models. After having successfully trained the regression models, the test data is used for the final estimation using these models. Finally, we compared the models and estimated the cost that is best analyzed.In this work we have divided our work in 5 different sections literature review, methodology used, experiment results and conclusion.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 01,2023 at 07:24:20 UTC from IEEE Xplore. Restrictions apply.

II. Literature Review

In the field of medical insurance many projects for cost estimation are being worked on, which makes it a development prone area for both the medical and the insurance fields together we referred some of the existing work for our research :

Bertsimas et al. [1] estimated the cost of insurance of a year of nearly about five hundred thousand users using the cost of insurance of previous 2 years by using algorithms based on clustering and data mining techniques and on trees like classification trees. S. Sushmita et al. [2] "used characteristics like past history of the person and any previous expenditure record of a person to estimate his medical cost in future.For this they used period of three, six ,nine and twelve months to estimate the cost and implemented it using random forest and linear regression algorithm to estimate the cost."

J.L. Moran et al. [3] "for estimation of cost of health insurance user made a comprehensive model using linear regression method and estimated the cost of ICU using the characteristics data , diagnose reports , treatment time in hospital" Ramya D. et al. [4] discussed about multinomial logistic regression and random forest which attains best accuracy among all works present related to this estimation.Also section 3 of this work discusses the influence of machine learning inf the field of insurance and at section 4 establishes a comparison between above 2 algorithms on different aspects like F1 score , precision, recall , accuracy.

S. Panda et al. [6] uses regression models such as linear,multiple and polynomial and achieved accuracy of 75% in the model which is based on multiple linear regression and of 80.97% in polynomial regression also establish relation between the degree of polynomial regression and accuracy which achieves maxima when degree is 3 and beyond this accuracy decreases to 68 percent

Gregori D. et al. [7] have worked on the the linear regression model and the ordinary least squares method to study multiple variable modeling of healthcare. K. Bhatia et al. [11] focuses on estimating the health insurance medical cost using the linear regression and also established relationship between different the cost and the features present in the data set which is the USA medical cost personal dataset.Using linear regression they achieved a accuracy of 81.3 and section 5 of this work also provides for a website based interface for a user. N. V. Sailaja et al. [12] discussed various regression algorithms and finally implement an stacking regressor hybrid model which achieves an accuracy of 85 percent which was more then any on the regression techniques used singly

Angela D Kafuria et al. [13] discussed the estimation of cost using different parameters in the dataset.The study used models like multiple linear regressor, k-nearest Neighbours, and random forest regression which achieves a R2 score of 85.3% and MAE of 2726.4 and RMSE of 4783.8. Mladenovic et al. [14] studied the dependence of cost on various parameters present in the dataset individually and collectively and find out that smoke plays the most significant part in estimating the insurance cost.

III. Methodology used

Machine learning models are majorly used to predict information without having any prior knowledge. It helps developers to develop models which can work for variety of data the basic path followed in machine learning projects is :

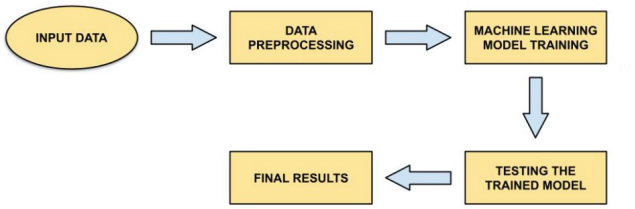


Fig. 1. Machine learning Model flow

A. Dataset

USA's medical cost personal data set in the primary source data set which is taken from [5]. It contains characteristics like 'age', 'gender' , 'charger' which are further mentioned in figure [2]. Originally the data was stored in the csv format which contains 1338 records with 6 characteristics. The data set was not in a standard format and hence we can not use it directly for training our machine learning models and hence we need to do data cleaning and pre-processing first and then we train our machine learning models In this we also found out that all the characteristics which are present in the taken data set don't influence the cost estimation that much as discussed in next subsection.

#	Column	Non-Null Count	Dtype
0	age	1338 non-null	int64
1	sex	1338 non-null	object
2	bmi	1338 non-null	float64
3	children	1338 non-null	int64
4	smoker	1338 non-null	object
5	region	1338 non-null	object
6	charges	1338 non-null	float64

Fig. 2. Initial Data set

B. Data Pre-processing

We performed some operations given below and the final transformed data set looks like in Fig [3]

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 01,2023 at 07:24:20 UTC from IEEE Xplore. Restrictions apply.

- Checking missing value : we did not find any missing value in the data set
- Checking duplicate values : there was 1 duplicate in the data set and that will be removed while processing the data set
- Feature Engineering :
  i) we create a 4 new columns in the data set which is named as 'statusnormal','statusobse', 'statusoverweight','statusunderwight' which are set 1 if true and 0 if false based on the 'bmi' score of the person
  ii) we create 4 other coloumns based on region of the person 'regnortheast', 'regnorthwest','regsoutheast', 'regsouthwest' and are assigned values 1 if belong to that region and 0 if do not and the column of region is erased
  iii) 2 columns are created 'sexmale' , 'sexfemale' and are set true false accordingly
- Transforming the characteristics: here we encode the values of columns 'sex' , 'region' , 'weightstatus' , and 'smoker
- Separating the Categorical and numerical values
- Splitting of data sets into 2 different parts, one for training and one for testing the model

```
Data columns (total 15 columns):
 #   Column              Non-Null Count  Dtype
---  ------              --------------  -----
 0   age                 1337 non-null   int64
 1   bmi                 1337 non-null   float64
 2   children            1337 non-null   int64
 3   smoker              1337 non-null   int64
 4   charges             1337 non-null   float64
 5   sex_female          1337 non-null   uint8
 6   sex_male            1337 non-null   uint8
 7   reg_northeast       1337 non-null   uint8
 8   reg_northwest       1337 non-null   uint8
 9   reg_southeast       1337 non-null   uint8
 10  reg_southwest       1337 non-null   uint8
 11  status_normal       1337 non-null   uint8
 12  status_obese        1337 non-null   uint8
 13  status_overweight   1337 non-null   uint8
 14  status_underweigth  1337 non-null   uint8
```

Fig. 3. Tranformed Data set with addtion of new parametres

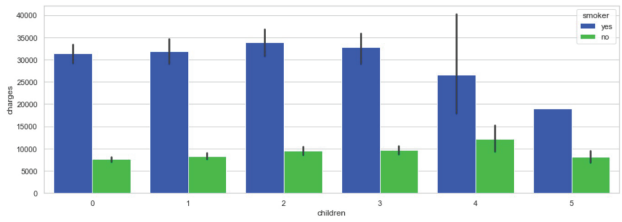


Fig. 4. 'smoke' parameter's importance in determining the cost

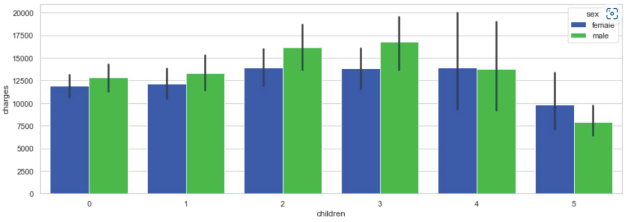


Fig. 5. shows parameter 'sex' is less important in estimating the cost

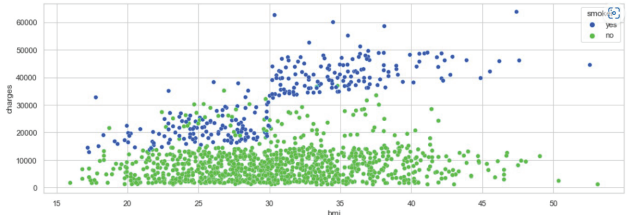


Fig. 6. shows relationship between bmi and charges

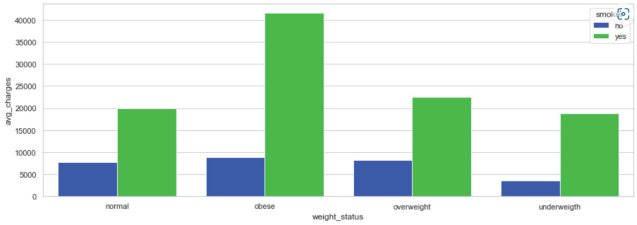


Fig. 7. shows higher 'bmi' and 'smoke' true then cost is high

C. Feature Weightage

analysing the data set with its characteristics by plotting them using python modules we get to know that some of the characteristics like 'smoke' plays important role in estimating the cost as in figure [4] while other characteristics like 'region' do not influence the final estimation cost that much as in figure [5] and figure [6] represents the relation between 'bmi' and 'cost' and figure [7] represent the effect of 'bmi' and 'smoke' parameter on cost

D. Models And algorithms Used

1) Linear Regression: In this type of regression model there is a linear dependence of a variable on another variable. Suppose there are two variables, X and Y namely, X being an independent variable and Y being a dependent variable, which means Y's value depends on X's value linearly in the form of the equation in the where a and b are the coefficients of regression

$$Y = a + bX \tag{1}$$

and a is the value of y when the curve cuts the y axis at that position the value of x is zero as shown in the equation [1] There might be some error in the value of y so we add some

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 01,2023 at 07:24:20 UTC from IEEE Xplore. Restrictions apply.

term so that we can get the correct value as mentioned in the equation [2]

$$Y = a + bX + \varepsilon \quad (2)$$

2) Decision Tree Regressor: It is generally used for supervised learning. In this algorithm we create a tree which consist of nodes of 3 types as structure represented in Figure [8]:
i) Root node : staring node which contains the entire data set and it is split further in different interior and leaf nodes
ii) Interior node: These nodes reflect the characteristic features of the given data set and the branches reflect the decisions rules under which the complete decision tree is formed.
iii) Leaf nodes: These nodes are at the bottom of the tree and don't have any child and represent the final outcome of decisions which were followed in the branches of the tree.

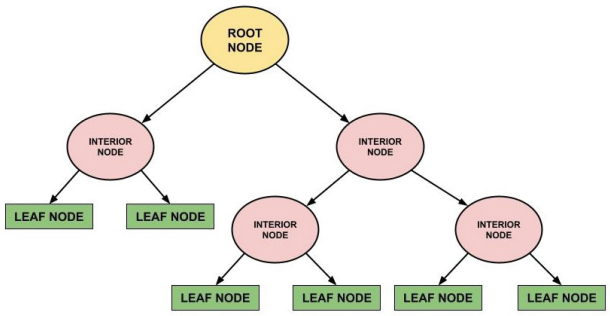


Fig. 8. Decision Tree Structure

3) Random Forest Classifier: It works on the same technique as in the decision trees but in this we have large no of decision tress on various data sets and it uses average of the result of the decision trees to improve the accuracy of the final output. Mostly it gives better results over a single decision tree, but it might also depend on the data set which is used. The flow of model is mentioned in Figure [9]

4) : Tuning HyperParameter: In machine learning it is a characteristics whose value is control or used to control the learning process of the machine learning model.The combination of Hyperparameter with the models gives the best accuracy , minimize the errors in most of the cases.
After using it in Linear regression we get achieved MAE: 4305.31 , RMSE : 6221.51 , R2 Score : 0.77 and Test accuracy of 77.37 percent

IV. Experiment Results

This section represents the flow of the data in Fig [10] and the results of the work on the basis of parameters R2 Score, MAE, RMSE and Test Accuracy are mentioned in Table 1:

Data Analysis results:

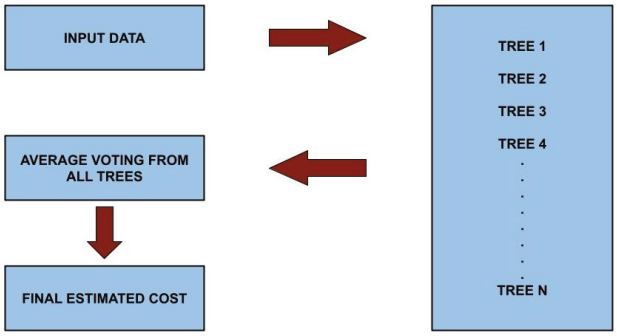


Fig. 9. Random Forest Classifier

TABLE I
Experimental Results

Regressor	R2 Score	MAE	RMSE	Accuracy
Linear Regressor	0.77	4305.20	6209.88	74.26
Decision Tree	0.78	2798.83	6067.50	78.48
Random forest	0.86	2608.55	4841.88	86.29

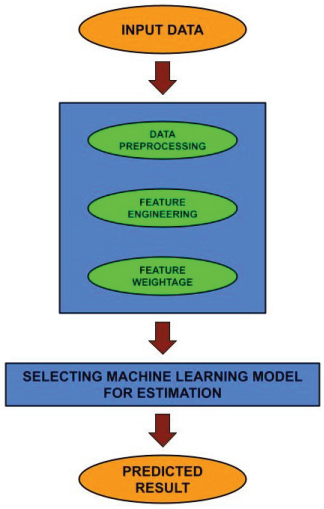


Fig. 10. Data Flow in Suggested Model

- Most of the people in the data set are non smokers and obese
- characteristics like sex, region has an almost balanced charges
- person in the data set who both smoke and have higher value of BMI have a high cost
- person in the data set with high age and smoke characteristics as true then the cost of the insurance is more
- person in the data set with smokes parameter as true and is obese then the cost is higher

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 01,2023 at 07:24:20 UTC from IEEE Xplore. Restrictions apply.

V. Conclusion

This work represented comparative analysis of different models with improved accuracy on USA medical cost dataset.We established the relation of dependence of cost on various parameters in the dataset like 'age','sex','BMI','smoking habits' and got the result that 'smoke' parameter is the most important in determining the cost and then we used different regression models to predict the cost.Linear Regression achieves 74.26% accuracy and decision tree achieves accuracy of 78.48% and Random forest achieves best accuracy of 86.29% and with RMSE 4841.88 and MAE 2608.55, Decision tree with MAE of 2798.83 ,RMSE of 6067.50 and accuracy of 78.48% and on using hyperparamaterization on linear regression the MAE increases to 4305.31 and the RMSE to 6221.51.

References

[1] Bertsimas, M. V. Bjarnad´ottir, M. A. Kane, J. C. Kryder, R. Pandey, S. Vempala, and G. Wang, "Algorithmic prediction of health-care costs," Operations Research, vol. 56, no. 6, pp. 1382–1392, 2008.

[2] S. Sushmita, S. Newman, J. Marquardt, P. Ram, V. Prasad, M. D. Cock, A. Teredesai et al., "Population cost prediction on public healthcare datasets," in Proceedings of the 5th International Conference on Digital Health 2015. ACM, 2015, pp. 87–94. Association for Computing Machinery, New York, NY, USA, 87–94.

[3] J. L. Moran, P. J. Solomon, A. R. Peisach, and J. Martin, "New models for old questions: generalized linear models for cost prediction," Journal of evaluation in clinical practice, vol. 13, no. 3, pp. 381–389, 2007.

[4] R. D, M. S. K and D. J, "Health Insurance Cost Prediction using Machine Learning Algorithms," 2022 International Conference on Edge Computing and Applications (ICECAA), 2022, pp. 1381-1384, doi: 10.1109/ICECAA55415.2022.9936153.

[5] https://www.kaggle.com/code/miriambrandt/health-insurance-medical-cost-forecast/data

[6] S. Panda, B. Purkayastha, D. Das, M. Chakraborty and S. K. Biswas, "Health Insurance Cost Prediction Using Regression Models," 2022 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COM-IT-CON), 2022, pp. 168-173, doi: 10.1109/COM-IT-CON54601.2022.9850653.

[7] Gregori, D., Petrinco, M., Bo, S., Desideri, A., Merletti, F. and Pagano, E., 2011. Regression models for analyzing costs and their determinants in health care: an introductory review. International Journal for Quality in Health Care, 23(3), pp.331-341.

[8] Health Insurance of India's missing middle, Niti Ayog India, Oct 2021, [Online]. Available: https://www.niti.gov.in

[9] Global Expenditure on Health, WHO annual report 2021, [Online]. Available:https://www.who.int/newsroom/events/detail/2021/12/15/default-calendar/global-spending-on-health-2021

[10] Health Insurance Plan Choice. MDM Policy & Practice. 2018; 3(1):238146831878109. doi:10.1177/2381468318781093.

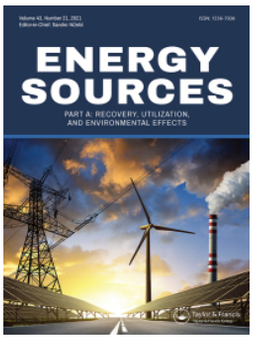
[11] K. Bhatia, S. S. Gill, N. Kamboj, M. Kumar and R. K. Bhatia, "Health Insurance Cost Prediction using Machine Learning," 2022 3rd International Conference for Emerging Technology (INCET), 2022, pp. 1-5, doi: 10.1109/INCET54531.2022.9824201.

[12] N. V. Sailaja, M. Karakavalasa, M. Katkam, D. M, S. M and D. N. Vasundhara, "Hybrid Regression Model for Medical Insurance Cost Prediction and Recommendation," 2021 IEEE International Conference on Intelligent Systems, Smart and Green Technologies (ICISSGT), 2021, pp. 93-98, doi: 10.1109/ICISSGT52025.2021.00029.

[13] Angela D. Kafuria, "Predictive Model for Computing Health Insurance Premium Rates Using Machine Learning Algorithms", IJC, vol. 44, no. 1, pp. 21–38, Aug. 2022.

[14] Mladenovic, S.S., Milovancevic, M., Mladenovic, I., Petrovic, J., Milovanovic, D., Petković, B., Resic, S. and Barjaktarović, M., 2020. Identification of the important variables for prediction of individual medical costs billed by health insurance. Technology in Society, 62, p.101307.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 01,2023 at 07:24:20 UTC from IEEE Xplore. Restrictions apply.



Energy Sources, Part A: Recovery, Utilization, and Environmental Effects

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/ueso20

Experimental approach and optimization of hot air drying tomato waste using response surface methodology

Djamel Mennouche, Tarik Hadibi, Abdelghani Boubekri, Amina Boudouaya, Abdellouahab Benseddik, Müslüm Arıcı, Anil Kumar & Wang Yunfeng

To cite this article: Djamel Mennouche, Tarik Hadibi, Abdelghani Boubekri, Amina Boudouaya, Abdellouahab Benseddik, Müslüm Arıcı, Anil Kumar & Wang Yunfeng (2023) Experimental approach and optimization of hot air drying tomato waste using response surface methodology, Energy Sources, Part A: Recovery, Utilization, and Environmental Effects, 45:1, 1048-1065, DOI: 10.1080/15567036.2023.2175937

To link to this article: https://doi.org/10.1080/15567036.2023.2175937

Published online: 09 Feb 2023.

Submit your article to this journal

Article views: 24

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
2023, VOL. 45, NO. 1, 1048–1065
https://doi.org/10.1080/15567036.2023.2175937

Taylor & Francis
Taylor & Francis Group

Check for updates

Experimental approach and optimization of hot air drying tomato waste using response surface methodology

Djamel Mennouche[a], Tarik Hadibi[b], Abdelghani Boubekri[a], Amina Boudouaya[c], Abdellouahab Benseddik[d], Müslüm Arıcı[e], Anil Kumar[f], and Wang Yunfeng[b]

[a]Laboratory of New and Renewable Energies Development in Arid and Saharan Zones, Faculty of Material Sciences, Kasdi Merbah University, Ouargla, Algeria; [b]Solar Energy Research Institute, Yunnan Normal University, Kunming, China; [c]Laboratory of Chemistry and Environmental Science (LCSE), Tahri Mohammad Bechar University, Bechar, Algeria; [d]Applied Research Unit in Renewable Energies, URAER, Renewable Energy Development Center, CDER, Ghardaïa, Algeria; [e]Mechanical Engineering Department, Engineering Faculty, Kocaeli University, Kocaeli, Turkey; [f]Mechanical Engineering Department, Delhi Technological University, Delhi, India

ABSTRACT

In the present work, an experimental investigation and mathematical modeling of tomato waste hot air drying (HAD) were conducted for drying temperatures of 50 and 60°C, and air velocities of 1 and 2 m s^{-1}. Kinetics, mathematical modeling, and optimization of biological properties using response surface methodology (RSM) were investigated. Increasing the drying temperature from 50 to 60°C shortened the drying time by 5 and 20 min, respectively, for 1 and 2 m s^{-1} air velocities. The proposed mathematical model with the highest correlation coefficient of 0.9997 well described the moisture variation of tomato waste. Significantly high diffusivity was observed with values of 1.24×10^{-6} and 1.67×10^{-6} m^2 s^{-1}. Obtained analysis showed an increase in polyphenol (PPT) and flavonoids (TFC) and a decrease in lycopene content (LYC) and ascorbic acid (AA). Phenomenological modeling and RSM methodology indicated that both drying temperature and air velocity affected the content of TFC, LYC, and AA, whereas, for PPT, the temperature has a positive effect in contrast with negative of air velocity. Optimum drying time, PPT, TFC, LYC, and AA are 44.6 min, 227 mg GAE/g, 32.2 mg GAE/g, 517.92 mg/100 g, and 1.91 mg AAE/100 g, respectively. Optimal drying temperature and velocity were, respectively, 56.7°C and 1 m s^{-1}.

ARTICLE HISTORY
Received 12 September 2022
Revised 26 October 2022
Accepted 31 October 2022

KEYWORDS
Hot air drying; tomato waste; RSM; modeling; antioxidants

Introduction

In industrial food processing, various techniques are used for food preservation, including boiling, heating, cooling, drying, etc. (Singh, Patel, and Singh 2021). Otherwise, processing industries of fruits and vegetables face the problem of waste, which has a high nutritional value as the main product. Usually, this waste is thrown away or used as animal feed. Murugavelh et al. (Murugavelh et al. 2019). reported 20%–30% of organic waste in the food industries' total waste.

Due to their year-round availability and cultivation, tomato wastes, such as peel and seeds, is a common waste product from the fruit industry. On the one hand, the industrial processing of tomato waste permits many carotenoids and other antioxidants, i.e., tomato peel includes 80%–90% of lycopene content (Riggi, Patané, and Ruberto 2008) or is a good source of polyphenolic antioxidants (Savatović et al. 2010). On the other hand, several industrial and semi-industrial processes were used to valorize tomato waste as peel and seed and conserve the main nutritional values and biologically

CONTACT Tarik Hadibi tarikhadibi@gmail.com Solar Energy Research Institute, Yunnan Normal University, Kunming 650500, China

© 2023 Taylor & Francis Group, LLC

active antioxidants. Drying is considered an effective industrial processing method to reach the abovementioned purposes.

Hence, various studies on drying tomato waste using several techniques have been done. Badaoui et al (Badaoui et al. 2019). investigated modeling the greenhouse drying of tomato pomace waste. The solar dryer was constructed mainly to recover the waste of fruit and vegetables from industrial processing. This drying temperature varied between 40 and 58°C to eliminate the microorganism and reduce the biological activities by reducing moisture content by 84% compared to the initial moisture content. The new proposed model called "model 2" well described the greenhouse drying of tomato waste. Obtained diffusivity ranged between 3.2×10^{-9} and 4.7×10^{-9} m^2/s, with a 75.6 kJ/mol activation energy. Murugavelh et al (Murugavelh et al. 2019). developed a mixed mode solar tunnel dryer for drying 2 kg of tomato waste. The dryer reached the desired moisture level in 7 h, while the open sun drying took 15 h. Average diffusivity was estimated at 5.64×10^{-10} and 6.35×10^{-10} m^2/s via linear fit and slope methods, respectively. The average exergy efficiency of the developed dryer was 52.21% with a range of 38.53–67.58%. The highest exergy efficiency was observed during peak solar radiation. Wei (Wei, Robert, and Wei Wei 2020) extracted lycopene from tomato waste using different processing methods. The average quantity of extracted lycopene using freeze drying, ultrasonication method, and their combination were 104.10 ± 1.23, 45.51 ± 1.84, and 138.82 ± 6.64 μg/g, respectively. Otherwise, drying 1.5 mg of tomato waste by freeze drying reduced DPPH and ABTS free radicals by 49.64 ± 0.44 and 12.3 ± 0.11%, respectively. Albanese et al (Albanese et al. 2014). observed that hot air drying (HAD) at 50°C followed by freeze-drying is a suitable technique to preserve the highest amount of both carotenoids and antioxidants in dried tomato peels. Miccio et al (Miccio et al. 2020). investigated experimentally and by simulation the controlled microwave drying (MWD) of tomato residues. Mathematical equation describing the function of both moisture content and drying temperature was developed for the MWD of tomato residues. Compared with HAD, MWD was found to be more efficient in reducing drying time, i.e., 220 and 190 min at 50 and 70°C, respectively. Otherwise, HAD was estimated to be with low-cost industrial investment by 0.2 M€ compared with MWD at 0.9 M€. Fărcaș et al (Fărcaș et al. 2019). exploited the biologically active compounds of tomato waste (total phenolic content, flavonoids, lycopene, and antioxidant activity) to get new products used as an ingredient or for natural activities. Two processing methods were investigated: drying at 50°C for 6 h and freezing at −18°C. The highest total phenolic compound, flavonoids, lycopene, and antioxidant properties were 148–202.21 mg GAE/100 g dw, 41.54 mg QE/100 g dw, 42.18–70.03 mg/100 g dw, and 16.57–29.14%, respectively. They observed the same values for both drying and freezing methods. EL-Mesery and El-khawaga (EL-Mesery and El-Khawaga 2022) reported that the lowest energy consumption for drying of tomato waste biomass was recorded as 3.77 MJ/kg for microwave-hot air drying (MW-HAD) in comparison with HAD and infrared HAD (IR-HAD) with a reduced rate of 93 and 55% compared to HAD and IR-HAD, respectively. The increased drying temperature indicated a rise in thermal efficiency. In contrast, it decreased with increase of air velocity. The efficiency increased by 6% with a 20°C increase of drying temperature. Tomato peel was dried using both fluidized-bed and cabinet dryer at 50–70°C (Kaur et al. 2006). They observed that the drying process has occurred totally in the falling rate period. The Clausius-Clapeyron model was used to calculate the net isosteric heat of sorption at moisture content of 2.0 to 2.5% dry basis, which ranged from 0.74 to 23.23 kJ/mol. In comparison with cabinet dryer, fluidized bed dryer was found more efficient. Nour et al. (Nour et al. 2018) introduced the quality attributes of dried tomato waste. The obtained results showed that leucine amino acid was the most abundant at 10.7 g/kg. Among the total fatty acid content, linoleic acid represents the highest value at 51.91%, the highest concentration of mineral content was indicated for potassium at 30,301.7 mg/kg, total phenolic, flavonoids, lycopene, β-carotene, and antioxidant activities were found to be 1229.5 ± 55.5 mg GAE/kg, 415.3 ± 18.2 mg QE/kg, 510.6 ± 21.1 mg/kg, 95.6 ± 3.3 mg/kg, and 6.8 ± 0.2 mmol Trolox/kg, respectively. The authors concluded that tomato waste (peel and seeds) are highly nutritious and contain valuable minerals, fatty acids, and amino acids, which may have a significant future utility as animal feed. In addition to tomato waste valorization, drying means was used to determine the kinetics and valorize the

nutritional and biological properties of many food wastes, such as sweet lime peel (Suri et al. 2022), olive-waste cake (Pasten et al. 2019; Vega-Gálvez et al. 2010), olive mill (Gómez-de la Cruz et al. 2020), kiwifruit (Tylewicz et al. 2022), sardine (Bahammou et al. 2019), salade (Plazzotta, Calligaris, and Manzocco 2018), dragon fruit peel (Chua et al. 2020), pomegranate peel (Kumar et al. 2022), nectarine pomace (Milanović et al. 2021), and pepper (Zhang et al. 2021).

Currently, process optimization is critical in all industries to define the ideal operating parameters, produce better outcomes, and save costs and production time. This requires the use of mathematical and statistical approaches with scientific validity that aids in predicting the behavior of the variables of interest (Malekjani and Jafari 2020). Many different modeling and optimization approaches have been invented to date in order to understand the behavior of a system and determine a global optimal point in the system, particularly to design and operate a chemical process. Many researchers have employed three different modeling methods for this: (1) RSM, (2) mathematical modeling, and (3) artificial neural network. Each approach has its own set of advantages and disadvantages. The RSM is a powerful mathematical and statistical prediction tool that allows to assess the influence of independent factors on responses with a small number of experiments (Kim et al. 2019). Furthermore, when creating a mathematical equation, its validation can be checked. The use of RSM in food processes has been intensively researched in recent years, and despite promising results, improvement in this field remains a problem due to the complexity of the products.

The present study is a complementary investigation on drying tomato waste using controlled HAD temperatures (50 and 60°C) and air velocities (1 and 2 m s^{-1}). The main objectives of the study are given below:

- To investigate the effect of drying parameters on drying time, drying rate, and diffusional process.
- To propose a new mathematical model to describe the drying experiments.
- To analyze and discuss the biological properties of dried tomato waste, namely polyphenol (PPT), flavonoids (TFC), lycopene content (LYC), and ascorbic acid (AA).
- To optimize and model the experimental findings and analyze results using the RSM technique and obtain the optimum drying parameters.

The current investigation is in line with the 17 Sustainable Development Goals (SDGs) to reach the food security in the developing countries.

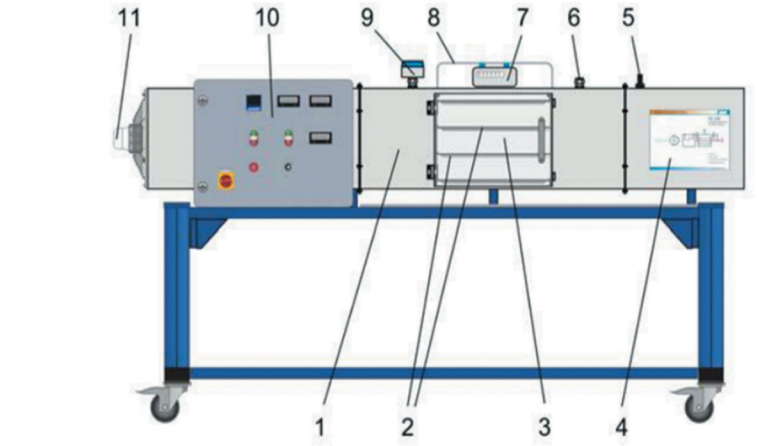
Materials and methods

Drying experiment

A tunnel-convective drier set up (see Figures 1–2) in the Mechanical Engineering Laboratory at Ouargla University in Algeria was used for drying tomato waste. It is a forced convection device made specifically to dry food and food waste. Four removable stainless-steel plates are positioned in the drying chamber with the dimensions of 250 cm × 50 cm × 50 cm. Four sheets of galvanized metal cover the outside of the drying chamber, which is thermally insulated with polystyrene of a thickness of 5 cm.

Waste of tomato peel and seeds deposited on a drying rack is exposed to airflow. Digital balance measures the mass loss of waste. A combined temperature and humidity sensor detects the temperature and relative humidity of the air (before and after the drying rack). An additional sensor serves to measure the air velocity. Electric resistance preheats the air coming from the surroundings. Hot air is circulated inside the drying chamber at an adjustable speed by a DC fan and passes through the waste while drying at a controlled drying temperature and exits at a lower value with high humidity.

Tomatoes used in this study are a commercial variety of the "rounded" type, belonging to the species *Licopersicum esculentum*, obtained from a local market in Ouargla, Algeria.



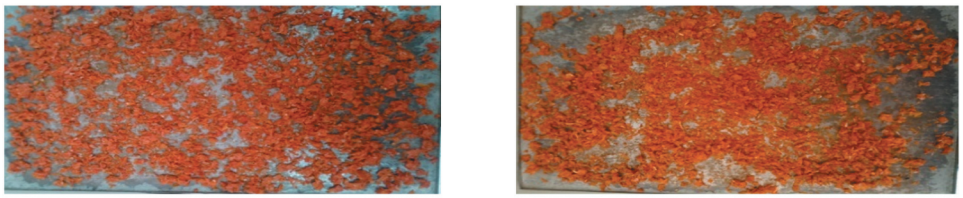
1- Drying channel 2- Drying sheet 3- Transparent door 4- Process diagram
5- Air speed sensor 6- Temperature and humidity sensors 7- Digital balance
8- Drying plate frame 9- Combined temperature and humidity sensor
10- Control cabinet 11- Fan

Figure 1. Experimental setup of HAD for drying tomato waste.

Figure 2. Real picture of the HAD loop.

The necessary steps for the preparation of the waste are as follows:

- Washing of fresh tomatoes to remove impurities, mud, insecticide residues, and other contaminations,
- Grinding the tomato using a grinder intended for vegetables,
- Separation of the three by-products: wet paste, seeds, and peel using a series of sieves from 2 mm to 1 mm in diameter,
- Draining the peels and seeds waste into a permeable bag,
- Calibration of the drained product in thin-layer form in a drying rack.



(a) (b)

Figure 3. (a) Fresh and (b) dried samples of the tomato waste.

Moisture content was defined in g water/g dry matter (g/g). Temperature and humidity measurements were made using a humidity-temperature probe. The air velocity in the dryer was measured using a propeller anemometer (type CA 1051).

The drying behavior and biological properties of the waste were investigated according to the design of the experimental method using StatGraphic software. A complete experimental design of 2^2 was used.

After the complete drying operation, the grinding of dried peel and seed waste was carried out using a grinder intended for dried products to achieve homogeneity, then stored at 25°C until use. A real picture of fresh and dried samples of the waste is shown in Figure 3.

Moisture measurement

The initial moisture content of tomato waste ($M_{c,0}$) was determined through the Association of Official Agricultural Chemists (AOAC) method (Naeemi et al. 1995). A laboratory-scale moisture analyzer was used to dry a sample of 3 g weight at 105°C temperature. Once the sample weight stabilizes, the analyzer shuts off automatically. The measurement of moisture content is reported as grams of water per gram of dry matter (g/g).

The $M_{c,0}$ was found to be 1.99 ± 0.02 g/g. Evolution of moisture content ($M_{c,t}$) can be expressed by Eq. 1:

$$M_{c,t} = \frac{m_t - m_d}{m_d} \tag{1}$$

Hence, the drying rate (DR) and moisture ratio(MR)were calculated, respectively, by Eq. 2-3 as follows:

$$DR = \frac{M_{c,t+\Delta t} - M_{c,t}}{\Delta t} \tag{2}$$

$$MR = \frac{M_{c,t} - M_{c,eq}}{M_{c,0} - M_{c,eq}} \tag{3}$$

$M_{c,eq}$ being the equilibrium moisture content.

Fitting of mathematical equations

The thin-layer of tomato wastes was fitted to six mathematical equations to describe the moisture content change during the HAD process. Based on suitable statistical parameters given by Eq. 4-6, the most appropriate mathematical equation is selected. The experimental MR curves fitted to mathematical equations are presented in Table 1.

The suitable mathematical equation describing the drying curve can be selected according to the regression parameters (highest correlation coefficient (R^2), lowest mean bias error (MBE), and

Table 1. Tested mathematical equations for drying curves fitting.

Name	Equation	Ref.
Lewis	$MR = \exp(-kt)$	(Bruce 1985)
Page	$MR = \exp(-kt^n)$	(Page 1949)
Modified page	$MR = \exp(-(kt)^n)$	(White, Ross, and Poneleit 1981)
Two term-exp	$MR = a\exp(-kt) + (1-a)\exp(-kat)$	(Sharaf-Eldeen, Blaisdell, and Hamdy 1980)
Wang and Singh	$MR = 1 + at + bt^2$	(Wang and Singh 1978)
Proposed model	$MR = (1-a)\exp(-(kt)^n) + \beta$	Current paper

Table 2. Full factorial design.

	Coded level	
Variables	−1	1
X_1	50	60
X_2	1	2

reduced chi-square (x^2)) (Akpinar and Bicer 2008). These parameters are given, respectively, by Eq. 4-6.

$$R^2 = \frac{\sum_{i=1}^{n}\left(MR_i - MR_{pre,i}\right).\sum_{i=1}^{n}\left(MR_i - MR_{\exp,i}\right)}{\sqrt{\left[\sum_{i=1}^{n}\left(MR_i - MR_{pre,i}\right)^2\right].\left[\sum_{i=1}^{n}\left(MR_i - MR_{\exp,i}\right)^2\right]}} \tag{4}$$

$$MBE = \left[\frac{\sum_{i=1}^{n}\left(MR_{pre,i} - MR_{\exp,i}\right)}{N}\right] \tag{5}$$

$$x^2 = \frac{\sum_{i=1}^{n}\left(MR_{\exp,i} - MR_{pre,i}\right)^2}{N-P} \tag{6}$$

Effective moisture diffusivity

The effective moisture diffusivity (D_{eff}, $m^2\ s^{-1}$) is a thermophysical property that describes all methods of transporting moisture from the core of the material toward the surface and helps characterize the moisture content change in the product with temperature (Oztop and Akpinar 2008). Usually, D_{eff} can be determined by applying the second simplified Fick's law (Guiné, Pinho, and Barroca 2011) as given by Eq. 7:

$$\frac{\partial MR}{\partial t} = \left[D_{eff}\left(\frac{M_{c,t} - M_{c,eq}}{M_{c,0} - M_{c,eq}}\right)\right] = D_{eff}\ {}^2\left(\frac{M_{c,t} - M_{c,eq}}{M_{c,0} - M_{c,eq}}\right) \tag{7}$$

For constant drying conditions, the $M_{c,eq}$ can be defined as the moisture loaded in the product at a controlled drying temperature and internal relative humidity. $M_{c,eq}$ of the tomato waste was calculated via a specific mathematical model as introduced by Sasongko et al. (Sasongko et al. 2020). Thus, $M_{c,eq}$ value can be determined utilizing the desorption isotherms method as defined in Eq. 8.

$$1 - RH = \exp\left(-A(T + C) + 100M_{c,eq}.B\right) \tag{8}$$

Here, the equation constants A, B, and, C are, respectively, 3.6×10^{-5}, 2.48, and 10.87 (Sasongko et al. 2020).

Assuming the waste HAD process is diffusional, the drying temperature and D_{eff} are constant, and the final shrinkage is negligible, Eq. 8 can be resolved based on Crank correlation (Crank 1979) as follows:

$$MR = \frac{8}{\pi^2} \sum_{n=1}^{\infty} \frac{1}{(2n+1)^2} \exp\left(-(2n+1)^2 \pi^2 \frac{D_{eff}.t}{4L^2}\right) \tag{9}$$

Otherwise, Eq. 9 can be simplified to the following equation:

$$MR = \frac{8}{\pi^2} \exp\left(\pi^2 \frac{D_{eff}.t}{4L^2}\right) \tag{10}$$

Experimental outputs of ln (MR) against drying time in (s) is plotted to determine D_{eff} which is the slope of the straight line, Eq. 11:

$$Slope = -\frac{\pi^2 D_{eff}.t}{4L^2} \tag{11}$$

Activation energy

The energy liberated through the drying process from the phenomenon of water transport from the inside to the surface of the material in (kJ mol^{-1}) is defined as the activation energy (E_a). The correlation between the water movement in the diffusional stage and the absolute temperature of the material during the drying process (T_{abs}) in K can be described using the Arrhenius relation (Mujumdar and Devahastin 2000):

$$D_{eff} = D_0 \exp\left(-\frac{E_a}{RT_{abs}}\right) \tag{12}$$

In the above equation, the constant D_0 in m^2 s^{-1} denotes the pre-exponential factor. E_a can be calculated by taking the natural logarithm of both sides as presented below:

$$\ln\left(D_{eff}\right) = \ln(D_0) - \frac{E_a}{R_g}\frac{1}{T_{abs}} \tag{13}$$

Biological properties

Antioxidants analysis

An extract was employed to measure the PPT and TFC content in the dried tomato wastes. This extract is obtained by maceration of 2 g of dried waste in methanol (Jimoh, Adedapo, and Afolayan 2010).

In order to quantify the total phenolic compounds, the Folin–Ciocalteu method proposed by Brune et al. (Brune, Hallberg, and Skånberg 1991) was employed, according to which the intensity of the blue color is proportional to the quantity of phenolic compounds present in the sample. This coloration is measured using a spectrophotometer (Agilent Technologies, Cary 60) at 765 nm taking gallic acid as a standard (Catalano et al. 1999).

Using a dose of 700 µl of extract, the TFC content was measured. The calorimetric method with sodium hydroxide and aluminum trichloride was used to determine the presence of TFC. The complex formed was analyzed using a spectrophotometer visible at 510 nm.

Extraction of the LYC was done by acetone–hexane–ethanol (1:1:1, v:v:v), and the LYC assay was carried out using the method described by Opiyo et al. (Opiyo and Ying 2005). The following formula was used to calculate the amount of LYC contained in the various dry samples:

$$Lycopene(mg/100g) = 312 \times A503 \tag{14}$$

where A_{503} is the absorbance at 503 nm.

Ascorbic acid analysis

In order to prepare the samples for analysis by HPLC (High-Performance Liquid Chromatography), an extraction method is used, as presented by Jorge et al. (Jorge et al. 2018). The sample to be analyzed was pushed by a mobile phase of the phosphoric acid H_3PO4 (0.003 M) in a column filled with a stationary phase (column C18 Luna Kintex of $4.6 \times 150 \times 5$ µm particle size); the column temperature was provided at 30°C and a UV–VIS spectrophotometer detects the ascorbic acid in a 345 nm wavelength. The injection volume was 20 µl. Then, 1 g of powder was mixed in 10 ml of phosphoric acid solution (0.003 M), stirred for 15 min, then refrigerated centrifuge for 60 min at 4°C; the extract was collected in a syringe and injected through a filter (0.45 mm).

RSM approach

The two variables examined in this study were temperature (*X1*) and velocity (*X2*). A full factorial design 2^2 was used to evaluate the impact of these factors' variations on the responses of drying time (Y_1), D_{eff} (Y_2), PPT (Y_3), TFC (Y_4), LYC (Y_5), and AA (Y_6). Four experimental runs in total were carried out. Each factor's range limits were assigned the values (−1) and (+1), respectively (see Table 2). The X_1 factor ranged between 50 and 60°C, and the X_2 factor ranged between 1 and 2 m s^{-1}. Statgraphics Centurion 18 software was used to analyze the experimental design responses. Analyses were carried out using the RMS and analysis of variance.

The design matrix shows the four experiments that resulted from different combinations of the *X(i)* parameters (namely, drying temperature and air velocity) and their corresponding RSM responses (i.e., the obtained results *Y(i)*). The general form of the regression equation for RSM optimization based on the first-order model, with two operating parameters and their interaction terms are defined using Eq. 15 as follows:

$$Y_i = b_0 + b_1X_{1i} + b_2X_{2i} + b_3X_{3i} + b_{12}X_{1i}X_{2i} + b_{13}X_{1i}X_{3i} + b_{23}X_{2i}X_{3i} \tag{15}$$

where Y_i is the RSM response or obtained results (Y_1, $Y_2 \ldots Y_6$), and b_0, b_i, and b_{ij} are, respectively, the constant of relation, linear constant, and interaction coefficient. X_i is the dimensionless drying parameter, where X_1 refers to the drying temperature in °C and X_2 refers to the air velocity in m s^{-1}.

Uncertainty analysis

Uncertainties and errors of the instruments in experimental measurements arise due to several factors, such as sensitivity, improper observation, handling, calibration of the instruments, etc. The overall errors and uncertainties in measurements of the HAD process can be given by taking the least amount value of instruments and uncertainty of the independent drying parameters (Murugavelh et al. 2019). Table 3 represents the estimated uncertainty in independent drying parameters. Moreover, the errors and uncertainties of the HAD process can be given in Eq. 16:

$$w_R = \left[\left(\frac{\partial R}{\partial x_1}w_1\right)^2 + \left(\frac{\partial R}{\partial x_2}w_2\right)^2 + \left(\frac{\partial R}{\partial x_3}w_3\right)^2 + \ldots. + \left(\frac{\partial R}{\partial x_n}w_n\right)^2\right]^{1/2} \tag{16}$$

Figure 4 represents a schematic diagram of the solution methodology.

Table 3. Errors and uncertainty of HAD drying parameters.

Parameter	Least count measurement	Error and Uncertainty
Temperature	0.8°C	±0.3
Air velocity	0.1 m s^{-1}	±0.08
Weight loss	1 g	±0.1
Moisture content	0.07 g/g	±0.018

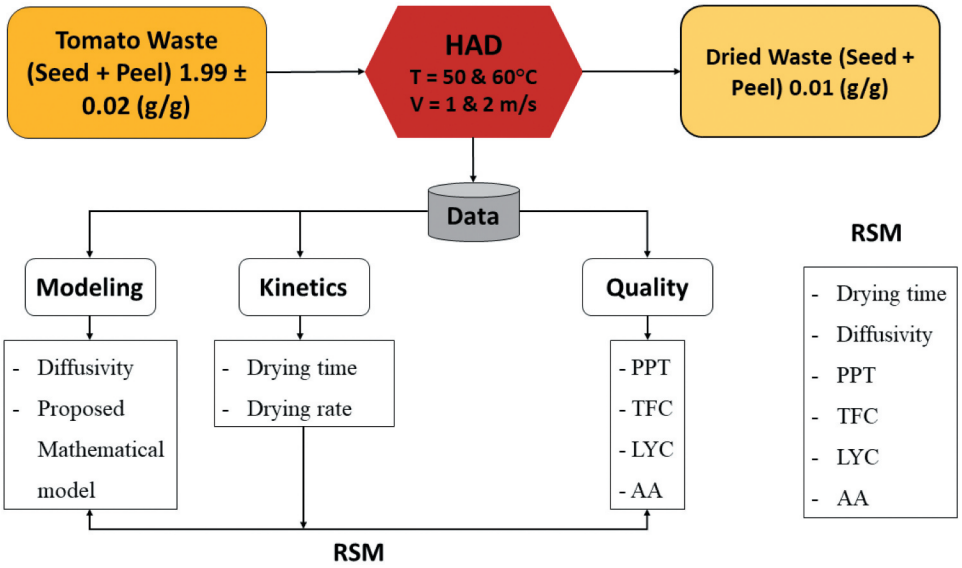


Figure 4. Experimental protocol of drying tomato waste by HAD.

Results and discussion

Drying kinetics

The evolution of the moisture content in the waste during the HAD process was plotted for various controlled drying temperature (Figure 5) and air velocity (Figure 6) values. The HAD with 1 m s^{-1} air velocity took 70 and 60 min to decrease the moisture content from 1.99 ± 0.02 g/g to the desired level at 0.01 g/g, under drying temperatures of 50 and 60°C, respectively. Otherwise, the drying process with 2 m s^{-1} air velocity was shortened by 5 and 20 min as 65 and 40 min, respectively, for drying temperatures of 50 and 60°C. Increasing the drying temperature systematically decreased the drying time. A remarkable influence of increasing air velocity on the drying time is observed. This phenomenon may be due to the high removal of water from the surface, as well as reduced external resistance of the waste, which accelerated the drying operation. This conclusion is in line with the findings of Allaf and Allaf (Allaf and Allaf 2016) for food drying processing. Figure 6 displays the controlled HAD speed of tomato waste in (g/g. min). The drying speed increased systematically with the elevating drying temperature from 50 to 60°C.

Moreover, increasing the air velocity from 1 to 2 m s^{-1} eliminated the water from the waste surface, which allowed water to move easily from the inside to the surface of tomato waste with low surface resistance, which led to an enhanced diffusion process. The reliance between moisture content evolution and drying speed is presented in Figure 7. Decreasing the moisture content reduced the drying speed of tomato waste. Evaporation of moisture content within time reduced the porosity of tomato waste, as well as increasing the internal resistance. The increase/decrease in the product porosity directly affects the increase/decrease speed of moisture evaporation. Hadibi et al. (Hadibi et al. 2022) observed similar characteristics for drying onion slices.

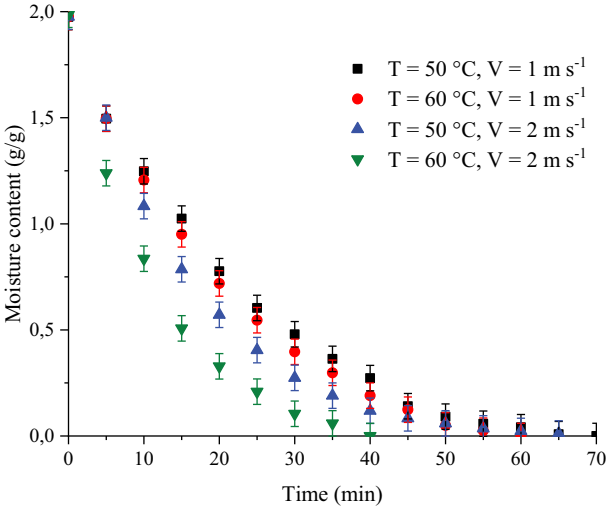


Figure 5. Evolution of tomato waste moisture content during HAD process.

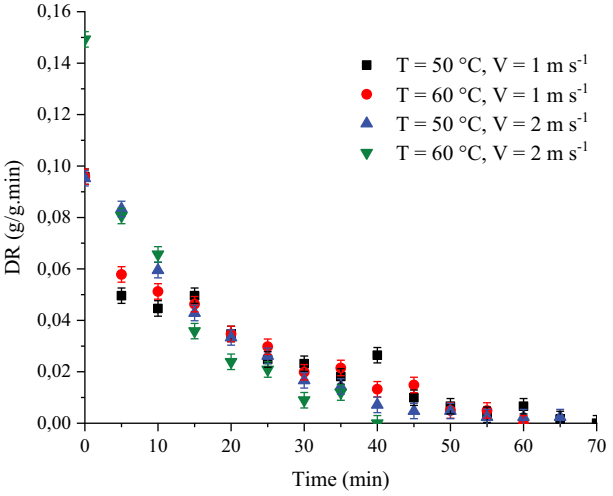


Figure 6. Drying speed of HAD tomato waste under constant drying conditions.

Mathematical modeling

Mathematical modeling of the controlled HAD process allows to predict the tomato waste behavior under controlled drying conditions. Statistical parameters R^2, *MBE*, and x^2 were adopted to determine the most appropriate mathematical equation describing the experiments. In addition to the proposed mathematical model, five other models were also tested. Among the six equations, the proposed mathematical model was identified to be the best model, which describes the drying behavior under the determined drying conditions with the highest $R^2 = 0.9997$, lowest $MBE = 0.0001$, and lowest $x^2 = 0.0091$. Figure 8 presents the alignment of the proposed mathematical model with the variation of moisture content against drying time. Tables 4 and 5 show the constant values of several tested models and statistical regression parameters, respectively. Various mathematical models were fitted to the drying curves of fruits and vegetable wastes. Tylewicz et al. (Tylewicz et al. 2022) reported the Weibull model as an appropriate model to describe PEF- and OD-assisted drying behavior of kiwifruit waste.

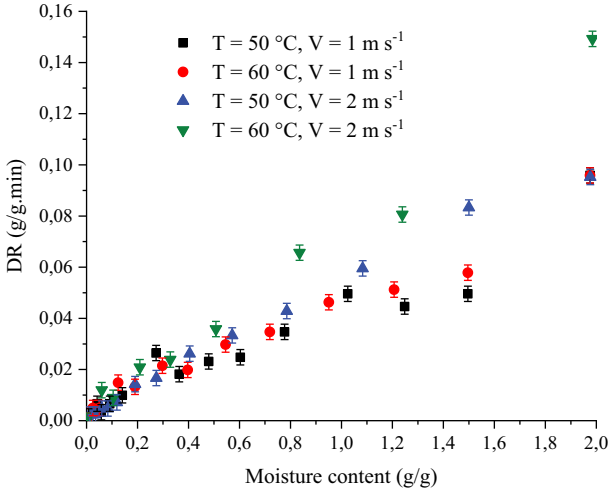


Figure 7. Drying rate against moisture content.

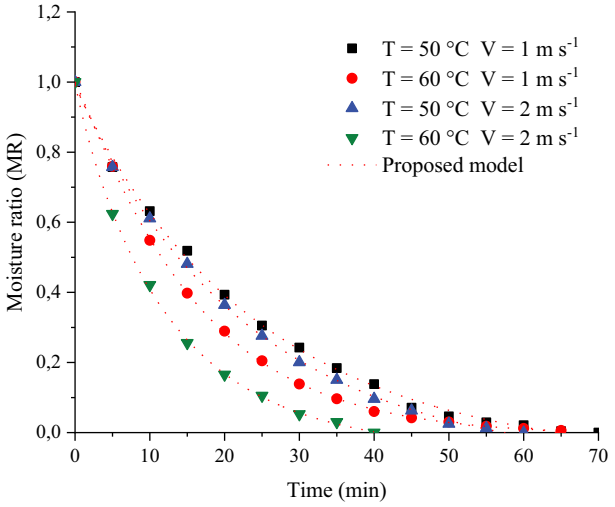


Figure 8. Drying curves fitting with the new mathematical equation.

Diffusivity and activation energy

The D_{eff} of tomato waste was calculated under constant temperatures of 50 and 60°C with 1 and 2 m/s air velocities, on one hand, and constant relative humidity, on the other hand. The diffusional process affected the $M_{c,eq}$ values, ranging between 3.26×10^{-4} and 3.47×10^{-4} (g/g). Considering $M_{c,eq}$, tomato waste diffusivity ranged between 1.24×10^{-6} and 1.67×10^{-6} m^2 s^{-1} (see Table 5). The highest D_{eff} was observed for HAD under the operating temperature of 60°C and 2 m s^{-1} air velocity. Increasing the drying temperature heated the inside of tomato waste and created a difference in partial pressure of moisture between the core and the surface, improving the diffusional process. A similar conclusion was reported by Hadibi et al. (Hadibi et al. 2021). A remarkable effect of air velocity reduced the external resistance and permitted a good moisture evacuation of water from the surface. The obtained D_{eff} values were higher than those reported by Zogzas et al. (Zogzas, Maroulis, and Marinos-Kouris 1996) in the range of 10^{-8}-10^{-12} m^2 s^{-1} for various agri-food products. This may be due to the low initial moisture of tomato waste and the high porosity of the waste.

Table 4. The constants of the mathematical equations for different drying conditions.

Model	*Cst*	*50℃ 1 m/s*	*60℃ 1 m/s*	*50℃ 2 m/s*	*60℃ 2 m/s*
Lewis	*k*	0.0494	0.0631	0.0536	0.0911
Page	*k*	0.0358	0.0468	0.0378	0.0859
	n	1.0993	1.1003	1.1106	1.0222
Modified page	*k*	0.0483	0.0619	0.0524	0.0906
	n	1.1015	1.1003	1.1126	1.0230
Two term exp	*a*	1.5685	1.5761	0.0001	1.3640
	k	0.0604	0.0777	0.7254	0.1006
Wang and singh	*a*	−0.0343	−0.0414	−0.0379	−0.0618
	b	0.0003	0.0004	0.0004	0.0009
Proposed model	*a*	−0.0766	−0.0076	−0.0994	−0.0501
	k	0.0403	0.0610	0.04264	0.0814
	n	0.9710	1.0812	0.96225	0.9259
	β	−0.0856	−0.0069	−0.10381	−0.0506

Table 5. Parameter regression of the tested mathematical equations.

Model	R^2	*MBE*	x^2
Lewis	0.9962	0.0004	0.0184
Page	0.9970	0.0003	0.0160
Modified page	0.9970	0.0003	0.0161
Two term exp	0.9964	0.0004	0.0173
Wang and Singh	0.9817	0.0020	0.0444
Proposed model	0.9997	0.0001	0.0091

Table 6. Tomato waste D_{eff} and E_a values for various drying conditions.

Run	T (℃)	V (m s^{-1})	H (%)	$M_{c,eq}$ (g/g)	D_{eff} (m^2 s^{-1})	E_a (kJ mol^{-1})
1	50	1	7.9	3.26×10^{-4}	1.24×10^{-6}	5.08
2	60	1	8.1	3.3×10^{-4}	1.3×10^{-6}	
3	50	2	9.5	3.97×10^{-4}	1.25×10^{-6}	3.15
4	60	2	8.4	3.47×10^{-4}	1.67×10^{-6}	

The drying experiments recorded low E_a values under both air velocities. The high values of D_{eff} reduce the internal resistance of waste, as well as permit transport of water from the inside to the surface with low delivered self-energy. Hadibi et al. (Hadibi et al. 2021) found a similar finding for the combined solar/geothermal drying of tomato paste. Table 6 presents the E_a and D_{eff} values based on the constant drying temperature and relative humidity.

Antioxydants

Figure 9 displays the final amount of PPT, TFC, LYC, and AA in dried waste under different drying conditions compared with fresh waste.

A significant increase in PPT contents was recorded at all temperature and air velocity values compared to the fresh waste of peel and tomato seeds. The amounts of PPT in various drying treatments increased by 285% to 312%. The trend of increase in the TFC, as observed in this study, agrees with the results on tomato powder resulting from freezing and hot drying (Jorge et al. 2018) and vacuum oven drying of tomato slices (Azeez et al. 2019). The increase in the PPT content during the drying process is justified by the ester decomposition between the TFC and cell wall. This decomposition causes the release of TFC compounds from the cell wall (Azeez et al. 2019).

Various processed wastes and drying operations increased TFC content relative to the fresh waste of peel and tomato seeds. TFC in all various drying treatments increases by 61% to 142%. The outcomes agree with those introduced by Changet al (Chang et al. 2006), who reported an increase

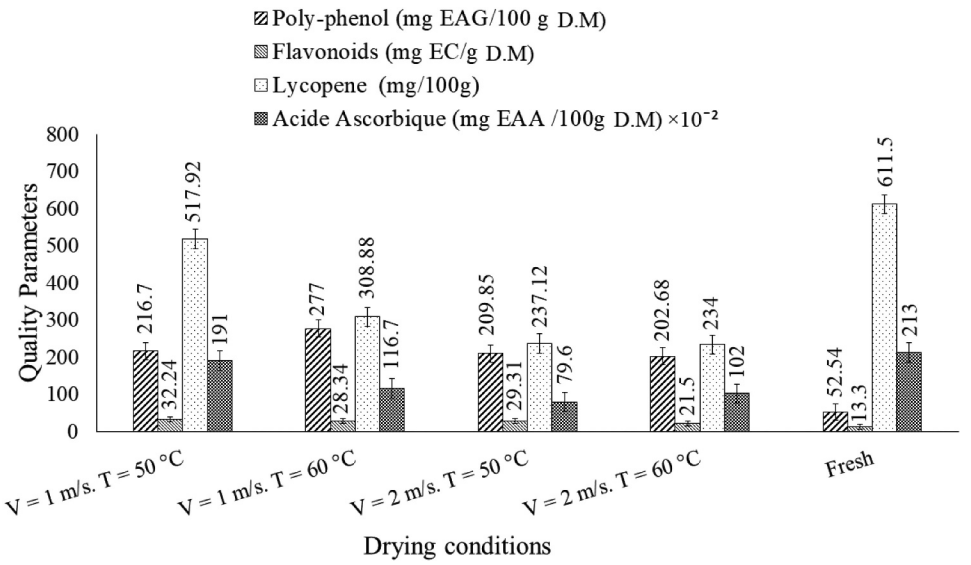


Figure 9. Quality parameters vs. drying conditions.

of 23% to 89% for two varieties of tomatoes treated by freezing and hot air drying. Many biochemical reactions might explain the increase in TFC in peels of tomatoes during the drying process.

The drying operation decreased the LYC content for all samples of the dried waste in comparison with the fresh one. The amounts of LYC in various drying treatments decreased by 15% to 62%. This finding is similar to that of Changrue (Changrue 2006), who found a reduction of 152% to 191% under HAD of two varieties of tomatoes. Other studies (Hossain, Amer, and Gottschalk 2008; Mwende, Owino, and Imathiu 2018) also show that heat treatment can degrade the LYC. This loss of LYC is based on isomerization and oxidation, reducing the fraction of trans-lycopene and causing loss of color.

The heat treatment by the drying process for different samples of tomato waste caused a decrease in AA content by 10%–92% compared to the fresh sample. This finding is similar to that reported by Hossain et al. (Hossain, Amer, and Gottschalk 2008), where a loss ranging from 71% to 77% was observed in tomato powder resulting from different drying methods. Yarahmadi et al. (Yarahmadi, Hojjatoleslamy, and Sedaghat Boroujeni 2020) reported that convective drying methods reduced the antioxidant capacity of dried fruits and vegetables.

Phenomenological modeling using RSM

According to the experimental design, the responses of the current RSM optimization are presented in Table 7. The RSM presented in Figure 10 reveals that air velocity negatively affects all the responses (i.e.,

Table 7. Experimentation matrix and results for the experimental design method.

	Independent Variables				Responses					
	Coded		Uncoded							
Run	X_1	X_2	T (°C)	V (m s^{-1})	Time (min)	D_{eff} (m^2 s^{-1})	PPT	TFC	LYC	AA
01	−1	−1	50	1	70	1.24×10^{-6}	216.7	32.24	517.92	1.91
02	+1	−1	60	1	60	1.3×10^{-6}	277	28.34	308.88	1,17
03	−1	+1	50	2	65	1.25×10^{-6}	209.85	29.31	237.12	0.796
04	+1	+1	60	2	40	1.67×10^{-6}	202.68	21.5	234	1.02

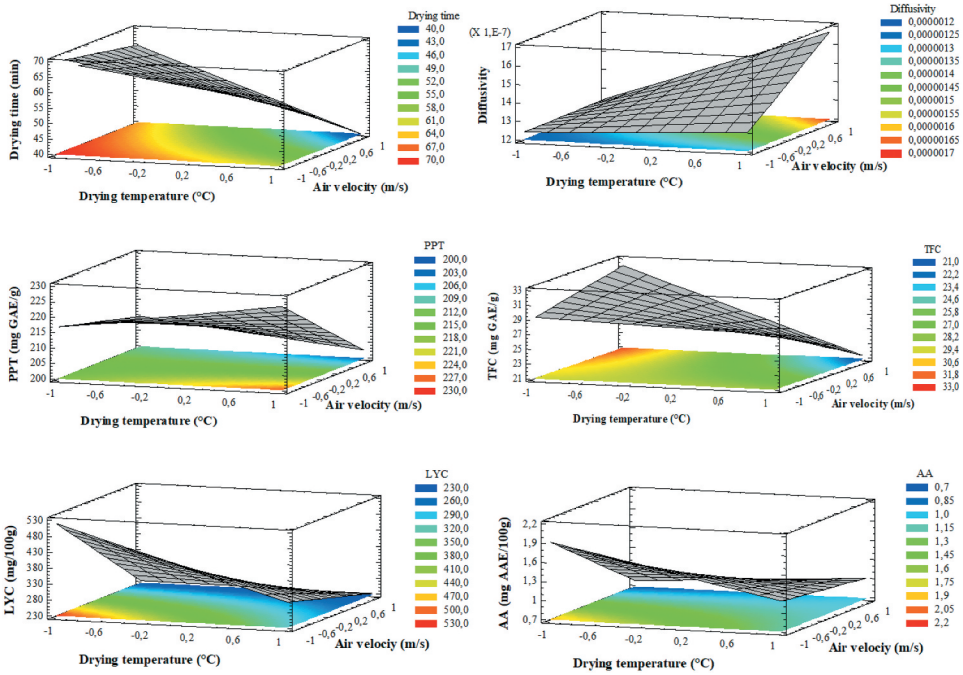


Figure 10. RSM of drying experiments.

PPT, TFC, LYC, and AA). However, drying temperature has an adverse influence on the drying time, D_{eff}, LYC, TFC, and AA content, and a positive effect on the PPT content, i.e., a rise in the temperature and air velocity causes losses in LYC, TFC, and AA contents. This result may be due to the increased drying temperature breakdown of the cellular walls and the chromoplast membranes (Parveen et al. 2016). It can be concluded that a rise in the temperature caused an increment in the PPT content. A similar remark was observed by Safaiee et al. (Safaiee et al. 2019). In contrast, Fărcaş et al. (Fărcaş et al. 2019) proved that the amount of PPT depends mainly on the initial moisture content, without significant effect of the conservation method. However, Wu et al. (Wu, Yu, and Pehrsson 2022) reported that increased losses of PPT are often associated with higher processing temperatures and longer processing durations.

Determination of the optimal conditions for temperature and air velocity is based on the optimization study of the set of responses (Multiple Response Optimization) via experimental design. The optimization objective is to minimize the drying time and maximize the rest of the responses (D_{eff}, PPT, TFC, LYC, and AA). The optimization results and phenomenological models describing the variation of drying time and quality parameters are given in Table 8 and Eq. 17-22, respectively. The optimal operating temperature and air velocity in the present HAD are 56.7°C and 1 m s^{-1}, respectively. The value of responses, namely, drying

Table 8. Different factors and responses for optimization.

Factor	Optimum
Temperature (°C)	56.7
Air velocity (m s^{-1})	1.0
Response	Optimum
Drying time (min)	44.6
Diffusivity (m^2/s)	1.67×10^{-6}
PPT (mg GAE/g)	227
TFC (mg GAE/g)	32.2
LYC (mg/100 g)	517.92
AA (mg AAE/100 g)	1.91

time, PPT, TFC, LYC, and AA are 44.6 h, 227 mg GAE/g, 32.2 mg GAE/g, 517.92 mg/100 g, and 1.91 mg AAE/100 g, respectively.

$$Dryingtime = 58.75 + 8.75 \times T + 6.25 \times V + 3.75 \times T \times V \quad (17)$$

$$Deff = 1.365 \times 10 - 6 + 1.2 \times 10 - 7 \times T + 9.5 \times 10 - 8 \times V + 9 \times 10 - 8 \times T \times V \quad (18)$$

$$PPT = 214.058 + 0.7825 \times T - 7.7925 \times V - 4.3675 \times T \times V \quad (19)$$

$$TFC = 27.835 - 2.915 \times T - 0.985 \times V - 2.435 \times T \times V \quad (20)$$

$$LYC = 324.48 - 53.04 \times T - 88.92 \times V + 51.48 \times T \times V \quad (21)$$

$$AA = 1.224 - 0.129 \times T - 0.316 \times V + 0.241 \times T \times V \quad (22)$$

Conclusions

The present study investigated and optimized (via RSM) the kinetics and biological properties of tomato waste drying using a controlled HAD considering different operating drying temperatures and air velocities. The proposed mathematical model was tested to predict the change in moisture content over time. The main findings of the current study are summarized as follows:

- Incrementing the drying temperature from 50 to 60°C consistently reduced the drying time by 5 and 20 min, respectively, for air velocities of 1 and 2 m s^{-1}.
- The proposed equation with the highest $R^2 = 0.9997$, lowest $MBE = 0.0001$, and the lowest $x^2 = 0.0091$ was found to be the most appropriate model to describe the HAD behavior of the tomato waste drying.
- The air velocity has a negative influence on all the responses. However, the temperature has a positive impact on both the drying time, LYC, TFC, and AA, and the PPT content.
- The optimal temperature and air velocity for the present dryer are 56.7°C and 1 m s^{-1}, and the different responses of the RSM to drying time, D_{eff}, PPT, TFC, LYC, and AA are 44.6 min, 1.67×10^{-6} m^2 s^{-1}, 227 mg GAE/g, 32.2 mg GAE/g, 517.92 mg/100 g, and 1.91 mg AAE/100 g, respectively.

Solar drying of food and vegetable waste using several thermal energy sources available in southern Algeria will be considered in the future work. The feasibility of the proposed hybrid dryer will be discussed based on energy, environmental, economic analysis, and quality attributes.

Acknowledgements

The first author would like to thank the Chemistry Lab, Faculty of Sciences and Technology, Adrar University, and Directorate General for Scientific Research and Technological Development (DGRSDT).

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Tarik Hadibi http://orcid.org/0000-0003-0117-9498

Abdelghani Boubekri http://orcid.org/0000-0002-1551-8000
Müslüm Arıcı http://orcid.org/0000-0002-3397-2215
Anil Kumar http://orcid.org/0000-0002-2846-4347
Wang Yunfeng http://orcid.org/0000-0002-2986-0464

References

Akpinar, E. K., and Y. Bicer. 2008. Mathematical modelling of thin layer drying process of long green pepper in solar dryer and under open sun. *Energy Conversion and Management* 49:1367–75. doi:10.1016/j.enconman.2008.01.004.

Albanese, D., G. Adiletta, M. D′ Acunto, L. Cinquanta, and M. Di Matteo. 2014. Tomato peel drying and carotenoids stability of the extracts. *International Journal of Food Science & Technology* 49:2458–63. doi:10.1111/ijfs.12602.

Allaf, T., and K. Allaf. 2016. *Instant controlled pressure drop (DIC) in food processing*. London: Springer.

Azeez, L., S. A. Adebisi, A. O. Oyedeji, R. O. Adetoro, and K. O. Tijani. 2019. Bioactive compounds' contents, drying kinetics and mathematical modelling of tomato slices influenced by drying temperatures and time. *Journal of the Saudi Society of Agricultural Sciences* 18:120–26. doi:10.1016/j.jssas.2017.03.002.

Badaoui, O., S. Hanini, A. Djebli, B. Haddad, and A. Benhamou. 2019. Experimental and modelling study of tomato pomace waste drying in a new solar greenhouse: Evaluation of new drying models. *Renew Energy* 133:144–55. doi:10.1016/j.renene.2018.10.020.

Bahammou, Y., H. Lamsyehe, M. Kouhila, A. Lamharrar, A. Idlimam, and N. Abdenouri. 2019. Valorization of co-products of sardine waste by physical treatment under natural and forced convection solar drying. *Renew Energy* 142:110–22. doi:10.1016/j.renene.2019.04.012.

Bruce, D. M. 1985. Exposed-layer barley drying: Three models fitted to new data up to 150°C. *Journal of Agricultural Engineering Research* 32:337–48. doi:https://doi.org/10.1016/0021-8634(85)90098-8.

Brune, M., L. Hallberg, and A. Skånberg. 1991. Determination of iron□binding phenolic groups in foods. *Journal of Food Science* 56:128–31. doi:10.1111/j.1365-2621.1991.tb07992.x.

Catalano, L., I. Franco, M. de Nobili, and L. Leita. 1999. Polyphenols in olive mill waste waters and their depuration plant effluents. A comparison of the Folin-Ciocalteau and HPLC methods. *Agrochim* 43:193–205.

Chang, C. -H., H. -Y. Lin, C. -Y. Chang, and Y. -C. Liu. 2006. Comparisons on the antioxidant properties of fresh, freeze-dried and hot-air-dried tomatoes. *Journal of Food Engineering* 77:478–85. doi:10.1016/j.jfoodeng.2005.06.061.

Changrue, V. 2006. *Hybrid (osmotic, microwave-vacuum) drying of strawberries and carrots*. Diss. Ph. D thesis. Canada: McGill University.

Chua, B. L., S. F. Tang, A. Ali, and Y. H. Chow. 2020. Optimisation of pectin production from dragon fruit peels waste: Drying, extraction and characterisation studies. *SN Applied Sciences* 2:1–13. doi:10.1007/s42452-020-2415-y.

Crank, J. 1979. *The mathematics of diffusion*. England: Oxford university press.

EL-Mesery, H. S., and S. E. El-Khawaga. 2022. Drying process on biomass: Evaluation of the drying performance and energy analysis of different dryers. *Case Studies in Thermal Engineering* 33:101953. doi:10.1016/j.csite.2022.101953.

Fărcaș, A. C., S. A. Socaci, D. Michiu, S. Biriș, and M. Tofană. 2019. Tomato waste as a source of biologically active compounds, Bull. *UASVM Food Science and Technology* 76:85. doi:10.15835/buasvmcn-fst:2019.0014.

Gómez-de la Cruz, F. J., J. M. Palomar-Carnicero, Q. Hernández-Escobedo, and F. Cruz-Peragón. 2020. Determination of the drying rate and effective diffusivity coefficients during convective drying of two-phase olive mill waste at rotary dryers drying conditions for their application. *Renew Energy* 153:900–10. doi:10.1016/j.renene.2020.02.062.

Guiné, R. P. F., S. Pinho, and M. J. Barroca. 2011. Study of the convective drying of pumpkin (Cucurbita maxima). *Food and Bioproducts Processing* 89:422–28. doi:10.1016/j.fbp.2010.09.001.

Hadibi, T., A. Boubekri, D. Mennouche, A. Benhamza, C. Besombes, and K. Allaf. 2021. Solar–geothermal drying/instant controlled pressure drop□swell drying of mechanically dewatered tomato paste. *Journal of Food Process Engineering* 44: e13811.

Hadibi, T., A. Boubekri, D. Mennouche, A. Benhamza, A. Kumar, C. Bensaci, and H. -W. Xiao. 2021. Effect of ventilated solar-geothermal drying on 3E (exergy, energy, and economic analysis), and quality attributes of tomato paste. *Energy* 243:122764. doi:10.1016/j.energy.2021.122764.

Hadibi, T., A. Boubekri, D. Mennouche, A. Benhamza, K. Mazouzi, A. Kumar, H. -W. Xiao, and K. Allaf. 2022. Energy, environmental, economic, and color analysis of geo-exchange energy assisted-insulated north wall solar dryer for onion slices under relatively cloudy and rainy conditions. *Solar Energy* 236:1–16. doi:10.1016/j.solener.2022.02.037.

Hossain, M. A., B. M. A. Amer, and K. Gottschalk. 2008. Hybrid solar dryer for quality dried tomato. *Drying Technology* 26:1591–601. doi:10.1080/07373930802467466.

Jimoh, F. O., A. A. Adedapo, and A. J. Afolayan. 2010. Comparison of the nutritional value and biological activities of the acetone, methanol and water extracts of the leaves of Solanum nigrum and Leonotis leonorus. *Food and Chemical Toxicology* 48:964–71. doi:10.1016/j.fct.2010.01.007.

Jorge, A., E. Sauer Leal, R. Sequinel, T. Sequinel, E. T. Kubaski, and S. M. Tebcherani. 2018. Changes in the composition of tomato powder (Lycopersicon esculentum Mill) resulting from different drying methods. *Journal of Food Processing and Preservation* 42:e13595. doi:10.1111/jfpp.13595.

Kaur, D., A. A. Wani, D. S. Sogi, and U. S. Shivhare. 2006. Sorption isotherms and drying characteristics of tomato peel isolated from tomato pomace. *Drying Technology* 24:1515–20. doi:10.1080/07373930600961371.
Kim, Z., Y. Shin, J. Yu, G. Kim, and S. Hwang. 2019. Development of NOx removal process for LNG evaporation system: Comparative assessment between response surface methodology (RSM) and artificial neural network (ANN). *Journal of Industrial and Engineering Chemistry* 74:136–47. doi:10.1016/j.jiec.2019.02.020.
Kumar, N., A. H. Sami, A. A. M. Khojah, R. Aljahani, E. Al-Mushhin, A. H. Aljahani, and A. A. M. Al-Mushhin. 2022. Effects of drying methods and solvent extraction on quantification of major bioactive compounds in pomegranate peel waste using HPLC. *Scientific Reports* 12:1–9. doi:10.1038/s41598-022-11881-7.
Malekjani, N., and S. M. Jafari. 2020. Food process modeling and optimization by response surface methodology (RSM). *Mathematics and Statistics Apply Food Engineering* CRC Press. 181–203.
Miccio, M., R. Pierri, G. Cuccurullo, A. Metallo, and P. Brachi. 2020. Process intensification of tomato residues drying by microwave heating experiments and simulation. *Chemical Engineering and Processing-Process Intensification* 156:108082. doi:10.1016/j.cep.2020.108082.
Milanović, M. P., M. S. Komatina, I. J. Zlatanović, N. G. Manić, and D. L. Antonijević. 2021. Kinetic parameters identification of conductive enhanced hot air drying process of food waste. *Thermal Science* 25:1785–807. doi:10.2298/TSCI200312223M.
Mujumdar, A. S., and S. Devahastin. 2000. Fundamental principles of drying. *Exergex, Brossard, Canada* 1:1–22.
Murugavelh, S., B. Anand, K. Midhun Prasad, R. Nagarajan, and S. Azariah Pravin Kumar. 2019. Exergy analysis and kinetic study of tomato waste drying in a mixed mode solar tunnel dryer, energy sources, part a recover. *Utilization, and Environmental Effects* 1–17. doi:10.1080/15567036.2019.1679289.
Mwende, R., W. Owino, and S. Imathiu. 2018. Effects of pretreatment during drying on the antioxidant properties and color of selected tomato varieties. *Food Science & Nutrition* 6:503–11. doi:10.1002/fsn3.581.
Naeemi, E. D., N. Ahmad, T. K. Al-Sharrah, and M. Behbahani. 1995. Rapid and simple method for determination of cholesterol in processed food. *Journal of AOAC International* 78:1522–24. doi:10.1093/jaoac/78.6.1522.
Nour, V., T. D. Panaite, M. Ropota, R. Turcu, I. Trandafir, and A. R. Corbu. 2018. Nutritional and bioactive compounds in dried tomato processing waste. *CyTa-Journal Food* 16:222–29. doi:10.1080/19476337.2017.1383514.
Opiyo, A. M., and T. Ying. 2005. The effects of 1□methylcyclopropene treatment on the shelf life and quality of cherry tomato (Lycopersicon esculentum var. cerasiforme) fruit. *International Journal of Food Science and Technology* 40:665–73. doi:10.1111/j.1365-2621.2005.00977.x.
Oztop, H. F., and E. K. Akpinar. 2008. Numerical and experimental analysis of moisture transfer for convective drying of some products. *International Communications in Heat and Mass Transfer* 35:169–77. doi:10.1016/j.icheatmasstransfer.2007.06.005.
Page, G. E. 1949. Factors influencing the maximum rates of air drying shelled corn in thin layer. Masters Thesis Purdue Univ.
Parveen, J., A. Iqrah, Z. Yumi, and J. Irwandi. 2016. Process development for maximum lycopene production from selected fruit waste and its antioxidant and antiradical activity. *Journal of Food Processing & Technology* 7:576.
Pasten, A., E. Uribe, K. Stucken, A. Rodríguez, and A. Vega-Gálvez. 2019. Influence of drying on the recoverable high-value products from olive (cv. Arbequina) waste cake. *Waste and Biomass Valorization* 10:1627–38. doi:10.1007/s12649-017-0187-4.
Plazzotta, S., S. Calligaris, and L. Manzocco. 2018. Application of different drying techniques to fresh-cut salad waste to obtain food ingredients rich in antioxidants and with high solvent loading capacity. *LWT* 89:276–83. doi:10.1016/j.lwt.2017.10.056.
Riggi, E., C. Patané, and G. Ruberto. 2008. Content of carotenoids at different ripening stages in processing tomato in relation to soil water availability. *Australian Journal of Agricultural Research* 59:348–53. doi:10.1071/AR07215.
Safaiee, P., A. Taghipour, F. Vahdatkhoram, and K. Movagharnejad. 2019. Extraction of phenolic compounds from Mentha aquatica: The effects of sonication time, temperature and drying method. *Chemical Papers* 73:3067–73. doi:10.1007/s11696-019-00843-5.
Sasongko, S. B., H. Hadiyanto, M. Djaeni, A. M. Perdanianti, and F. D. Utari. 2020. Effects of drying temperature and relative humidity on the quality of dried onion slice. *Heliyon* 6:e04338. doi:10.1016/j.heliyon.2020.e04338.
Savatović, S. M., G. S. Ćetković, J. M. Čanadanović-Brunet, and S. M. Đilas. 2010. Utilization of tomato waste as a source of polyphenolic antioxidants. *Acta Periodica Technologica* 187–94. doi:10.2298/APT1041187S.
Sharaf-Eldeen, Y. I., J. L. Blaisdell, and M. Y. Hamdy. 1980. A model for ear corn drying. *Transactions of the ASAE* 23:1261–65. doi:10.13031/2013.34757.
Singh, D., S. K. Patel, and D. Singh. 2021. Computational studies of drying characteristics in thin-layer microwave-heated Solanum tuberosum. *Chemical Papers* 75:2727–35. doi:10.1007/s11696-021-01510-4.
Suri, S., A. Singh, P. K. Nema, S. Malakar, and V. K. Arora. 2022. Sweet lime (Citrus limetta) peel waste drying approaches and effect on quality attributes, phytochemical and functional properties. *Food Bioscience* 48:101789. doi:https://doi.org/10.1016/j.fbio.2022.101789.
Tylewicz, U., C. Mannozzi, J. M. Castagnini, J. Genovese, S. Romani, P. Rocculi, and M. Dalla Rosa. 2022. Application of PEF-and OD-assisted drying for kiwifruit waste valorisation. *Innovative Food Science & Emerging Technologies* 77:102952. doi:10.1016/j.ifset.2022.102952.

Vega-Gálvez, A., M. Miranda, L. P. Díaz, L. Lopez, K. Rodriguez, and K. Di Scala. 2010. Effective moisture diffusivity determination and mathematical modelling of the drying curves of the olive-waste cake. *Bioresource Technology* 101:7265–70. doi:10.1016/j.biortech.2010.04.040.

Wang, C. Y., and R. P. Singh. 1978. Use of variable equilibrium moisture content in modeling rice drying. *Transactions American Society Agricultural Engineering* 11:668–72.

Wei, W., P. Robert, and T. Wei Wei. 2020. Improving extraction of lycopene from tomato waste by-products using ultrasonication and freeze drying. *World Journal of Advanced Research and Reviews* 5:177–85. doi:10.30574/wjarr.2020.5.2.0044.

White, G. M., I. J. Ross, and C. G. Poneleit. 1981. Fully-exposed drying of popcorn. *Transactions of the ASAE* 24:466–68. doi:10.13031/2013.34276.

Wu, X., L. Yu, and P. R. Pehrsson. 2022. Are processed tomato products as nutritious as fresh tomatoes? Scoping review on the effects of industrial processing on nutrients and bioactive compounds in tomatoes. *Advances in Nutrition* 13:138–51. doi:10.1093/advances/nmab109.

Yarahmadi, N., M. Hojjatoleslamy, and L. Sedaghat Boroujeni. 2020. Different drying methods of Pistacia Atlantica seeds: Impact on drying kinetics and selected quality properties. *Food Science & Nutrition* 8:3225–33. doi:10.1002/fsn3.1582.

Zhang, H. S., M. Q. Chen, B. A. Fu, and Q. H. Li. 2021. Evaluation on microwave drying of waste paper towel with multi-magnetron and mode stirrer. *Drying Technology* 39:882–95. doi:10.1080/07373937.2020.1728306.

Zogzas, N. P., Z. B. Maroulis, and D. Marinos-Kouris. 1996. Moisture diffusivity data compilation in foodstuffs. *Drying Technology* 14:2225–53. doi:10.1080/07373939608917205.

FedEmo: A Privacy-Preserving Framework for Emotion Recognition using EEG Physiological Data

Mohd Ayaan Anwar
Manan Agrawal
Dept. of Computer Science and Engineering
Delhi Technological University
Delhi, India
mohdayaananwar_2k19co232@dtu.ac.in
mananagrawal_2k19co210@dtu.ac.in

Neha Gahlan
Dept. of Software Engineering
Delhi Technological University
Delhi, India
nehagahlan_2k21phdcs02@dtu.ac.in

Divyashikha Sethia
Dept. of Software Engineering
Delhi Technological University
Delhi, India
divyashikha@dtu.ac.in

Gaurav Kumar Singh
Dept. of Software Engineering
Delhi Technological University
Delhi, India
gauravkumarsingh_2k21dsc06@dtu.ac.in

Rishabh Chaurasia
Dept. of Computer Science and Engineering
Delhi Technological University
Delhi, India
rishabhchaurasia_2k18co291@dtu.ac.in

2023 15th International Conference on COMmunication Systems & NETworkS (COMSNETS) | 978-1-6654-7706-2/23/$31.00 ©2023 IEEE | DOI: 10.1109/COMSNETS56262.2023.10041308

***Abstract*—Emotions are intricate mental states triggered by neurophysiological adjustments linked to ideas, sensations, behavioral reactions, and a level of pleasure or annoyance. These changes are best traced with the physiological signal Electroencephalogram (EEG), as it records the direct sensations sent by the brain. Recent research on emotion classification methods employs conventional machine learning classifiers to access human emotions and perform automatic emotion recognition tasks. However, they lack in securing users' privacy and sensitive information because they need access to all data. A newly introduced framework Federated Learning (FL), can resolve this problem. It is an approach that aims to create a global model classifier without requiring access to users' local data. This study proposes a novel FL framework, Federated learning for Emotion recognition *(FedEmo)*, for emotion state classification from physiological signal EEG while preserving users' data privacy. It uses Artificial Neural Network (ANN) as a baseline model for classifying emotional states: Arousal, Valence, and Dominance. Adding the concept of federated learning to build a framework *FedEmo* prevents loss of privacy as it enables the local training on the client's end with an updated model from the global server without compromising privacy. The proposed *FedEmo* framework approach achieves accuracies of 63.3%, 56.7%, and 52.2% for Valence, Arousal, and Dominance, respectively, using the well-known DREAMER dataset. These results are comparable to the basic centralized ANN model with the additional development of privacy preservation.**

***Index Terms*—Emotion Recognition, Physiological signals, Federated Learning, Data Privacy**

I. Introduction

Emotions are interpersonal states that have an impact on the behavior and thoughts of a person. They originate from physical and physiological reactions and internal or external stimuli. Face expressions, behavioral patterns, and physiological reactions can all be used to discriminate between various emotions [1]. Out of these, physiological signals are the most preferred parameters as they map the exact emotions of a person at any instance. Recording such signals is performed by the different smart wearables [2]. Recent technology developments have made it easy to access physiological signal monitoring via wearable devices, including smartphones, smartwatches (Empatica E4), and many others. Emotional well-being can be better cared for by creating new techniques to extract emotional states from physiological measurements utilizing mobile devices often, which will have a favorable impact on physical health. The Electroencephalogram (EEG) is a popular and effective physiological signal for identifying different emotions. These signals record the current emotions of a person by sending the signals directly to the brain [3].

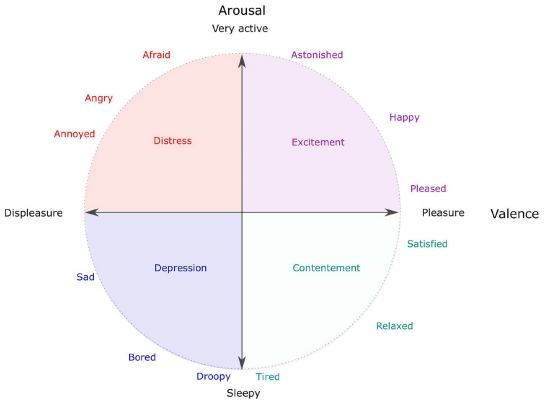


Fig. 1. Russell's circumplex model of affect [4]

Relative to this, Russell [4] proposed a two-dimensional model of emotions. This model categorizes emotions based on valence and arousal (as shown in Fig.1). Mehrabian and Russell [5], [6] extended the two-dimensional model into a three-

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 02,2023 at 06:43:00 UTC from IEEE Xplore. Restrictions apply.

dimensional model called VAD (Valence-Arousal-Dominance) model. This model introduces a new dimension representing dominance (as shown in Fig. 2), which ranges from submissive to dominant. Anger and fear are easily distinguished in this dimension because anger lies on the dominant axis, and fear lies on the submissive axis. Hence, this paper adopts the VAD model by Mehrabian and Russell [5], [6] for the proposed framework.

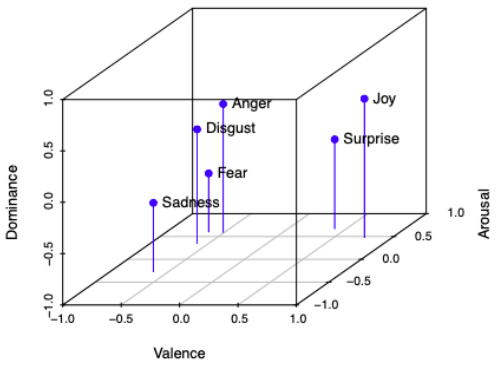


Fig. 2. 3D Model of emotions [3]

The physiological measurements by wearable devices are used as input to the automated artificial intelligence systems to classify emotions. Deep learning and machine learning models make it possible to automatically define correlations between measurements made under various situations, analyze vast volumes of data and classify the different emotional states. Support Vector Machine (SVM) [7], Decision Trees (DT), K-Nearest Neighbor (KNN) [8], Recurrent Neural Network (RNN) [9], Convolutional Neural Network (CNN) [10], and Long Short-Term Memory (LSTM) [11] are a few of the algorithms used in traditional automated emotion classification systems.

These algorithms hence proved to be accurate in terms of accuracy. However, concerning the privacy of the users' data, they have a severe disadvantage when used for emotion recognition. These algorithms employ the user's complete raw data as a training set for a model whose ability to preserve the user's privacy drastically declines, leading to a severe loss of privacy protection for the user's sensitive data.

To solve this problem, Federated Learning (FL) is the best method for preventing the loss of users' private data because it allows for local training using the most recent model from a global server without compromising privacy [12].

A. Background for Federated Learning

Federated Learning is a distributed paradigm for machine learning systems [13], [14] that creates an overall model utilizing training data scattered among several clients. The clients' mobile devices serve as nodes that process the local data they have recorded to update a global model. It transmits only the local model to the server instead of the entire raw data. After combining them using an aggregation approach called FedAvg [15], the server delivers these back to the client to produce a more accurate and private classification model. FedAvg is introduced by Google and relies on an optimization technique - Stochastic Gradient Descent (SGD). In this algorithm, the training process begins at the server machine by sharing the variables from the global model with the local models at the clients' end. This approach generates the results without compromising user privacy or data sharing.

Why privacy is important?

Privacy concerns are considered one of the significant challenges in Physiological signal-based Emotion Recognition (PER) applications. Traditional PER systems involve the complete sharing of sensitive, physiological data, which can bring threatening data attacks and leakage of private data.

For instance, the PER systems when used for analyzing employees in a work-from-home environment. The employees had different emotions and responses that they did not want to share with the other employees. In such a case, emotional information (EEG recordings) needs to be preserved.

Recent cases involving data leakage from emotion detection algorithms have occurred often. Two million users' biometric information and face recognition records got leaked in 2019 [16]. Their face recognition and fingerprint scan data became available to the public. Therefore, protecting such delicate information is crucial to prevent third parties' attacks or comments on people's emotional states.

Also, the EEG-based biometric authentication systems take the individual difference in the EEG signals as the only corresponding identity and perform authentication [12]. The leakage of such sensitive EEG data will ruin the authentication systems.

B. Related Work

Federated learning has been used for the task of emotion recognition in recent literature as summarised in Table I. Nandi and Xhafa [17] developed Fed-ReMECS, a federated learning system for real-time emotion state classification that leverages multimodal physiological data from Electrodermal Activity (EDA) and Respiration (RESP) [17] wearable sensors. Hassani [18] on the other hand uses Galvanic Skin Response (GSR) [18] for recognizing emotions. However, this study is pioneering the use of EEG signals instead of EDA, RB, and GSR in the federated learning paradigm for emotion recognition.

The authors classify nine emotional states from the DEAP [19] dataset using two domains of Arousal and Valence. Tara [18] employed the physiological signal GSR on the CASE [20] dataset to train a federated CNN-based model for emotion recognition. According to the author's findings, the proposed federated CNN architecture performs similarly to a centralized CNN architecture in terms of accuracy. However, its federated learning nature allows it to protect users' sensitive information.

Over the past ten years, the emotions from physiological data have grown across various industries. Traditional deep

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 02,2023 at 06:43:00 UTC from IEEE Xplore. Restrictions apply.

TABLE I
CURRENT RESEARCH WORK ON FEDERATED LEARNING FOR EMOTION RECOGNITION USING PHYSIOLOGICAL SIGNALS

Reference	Dataset	Physiological Signal	Classification Model	Average Accuracy	Tool	Aggregation Algorithm
Nandi et al. [17]	DEAP [19]	EDA, RB	FFNN	81.92 %	TFF	FedAvg
Hassani [18]	CASE [20]	GSR	CNN	79%	TFF	FedAvg

learning and machine learning algorithms are progressively used in most practical approaches to achieve high accuracy. However, these methods lack preserving the data privacy of a person's emotional information.

Dar et al. [21] proposed a computational framework of 2D Convolutional Neural Network (CNN) architecture for the arrangement of 14 EEG channels from the DREAMER [22] dataset and achieved an average accuracy of 90.8%. Tao et al. [23] proposed an Attention-based Convolutional Recurrent Neural Network (ACRNN) model to extract more discriminative features from EEG signals of the DREAMER [22] dataset. Furthermore, it achieved an average accuracy of 97.93% for arousal, valence, and dominance, which outperformed the state-of-the-art methods but failed to preserve the users' emotional data. Cheng et al. [24] proposed a multi-channel EEG-based emotion recognition method using the deep forest. The author considered the effect of baseline signal to preprocess the raw artifact-eliminated EEG signal with baseline removal. This method achieved 89.03%, 90.41%, and 89.89% accuracies for valence, arousal, and dominance, respectively, for the DREAMER [22] dataset.

C. Contributions

The following are the paper's primary research contributions:

- This paper proposes a novel Federated Learning framework *FedEmo* for Emotion recognition that utilizes an Artificial Neural Network (ANN) as a baseline model for classifying emotional states. It uses arousal, valence, and dominance scores using the EEG physiological signal. The proposed *FedEmo* framework uses the multimodal dataset DREAMER [22].
- The performance comparisons between the results of the baseline ANN model and the proposed *FedEmo* model indicate that the *FedEmo* model has comparable accuracies with the baseline ANN model.
- The results prove that the emotion recognition task can be achieved without sacrificing users' sensitive, emotional information with the proposed *FedEmo* framework.

The main goal of the paper is to create a distributed EEG physiological data-based emotion classification system that is effective and scalable, in which the global classifier is created without access to (privacy protection) the users' local data.

D. Roadmap

The rest of the paper comprises Section II describing the proposed work's methodology. Section III evaluates and discusses the results obtained. We finally conclude the paper with Section IV elaborating on the conclusion and future work.

II. METHODOLOGY

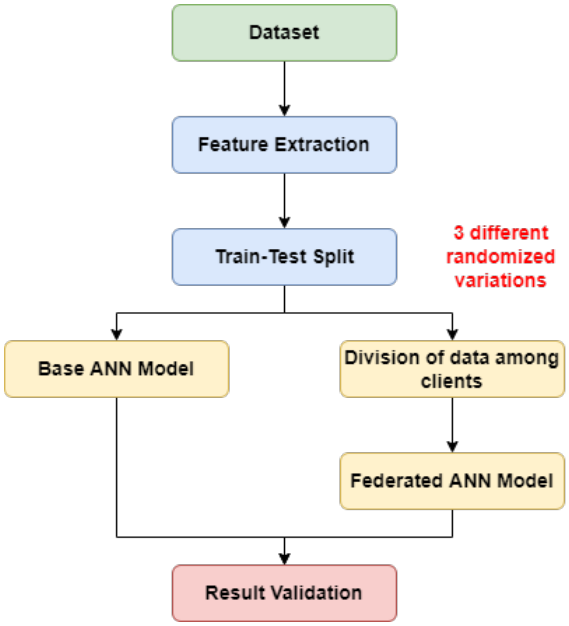


Fig. 3. Federated Emotion Classification Framework

A. Dataset Description

The DREAMER dataset [22] consists of EEG and ECG recordings of 23 subjects consisting of 14 males and 9 females, obtained from the individuals through 14 EEG and 2 ECG electrodes, respectively. The subjects watch 18 movie clips during the experiment. Each movie clip lasts between 65 and 393 seconds. The average length of a movie clip is 199 seconds. Before the data collection, a neutral movie clip is shown to the subjects to induce a neutral emotional state, which also serves as a baseline signal. Sixty seconds of baseline data were gathered before each affected movie clip. Subjects rated their arousal, valence, and dominance levels from 1 to 5 after watching each clip. Low/high valence, low/high dominance, and low/high arousal are the three binary states given to all the EEG signals.

B. Feature Extraction

Before the feature extraction process, firstly, the last 60 seconds of each video clip's EEG data is trimmed before decomposing them into the θ (4-8 Hz), α (8-13 Hz), and β (13-20 Hz) frequency bands.

- The 60-second EEG signals split into a series of 59 blocks for each frequency band using a sliding window with a width of 256 EEG points, where there is a half overlap between two succeeding blocks.
- Using Welch's overlapping segment averaging estimator, each EEG band's Power Spectral Density (PSD) is calculated using a 256-sample window with a 128-sample overlap.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 02,2023 at 06:43:00 UTC from IEEE Xplore. Restrictions apply.

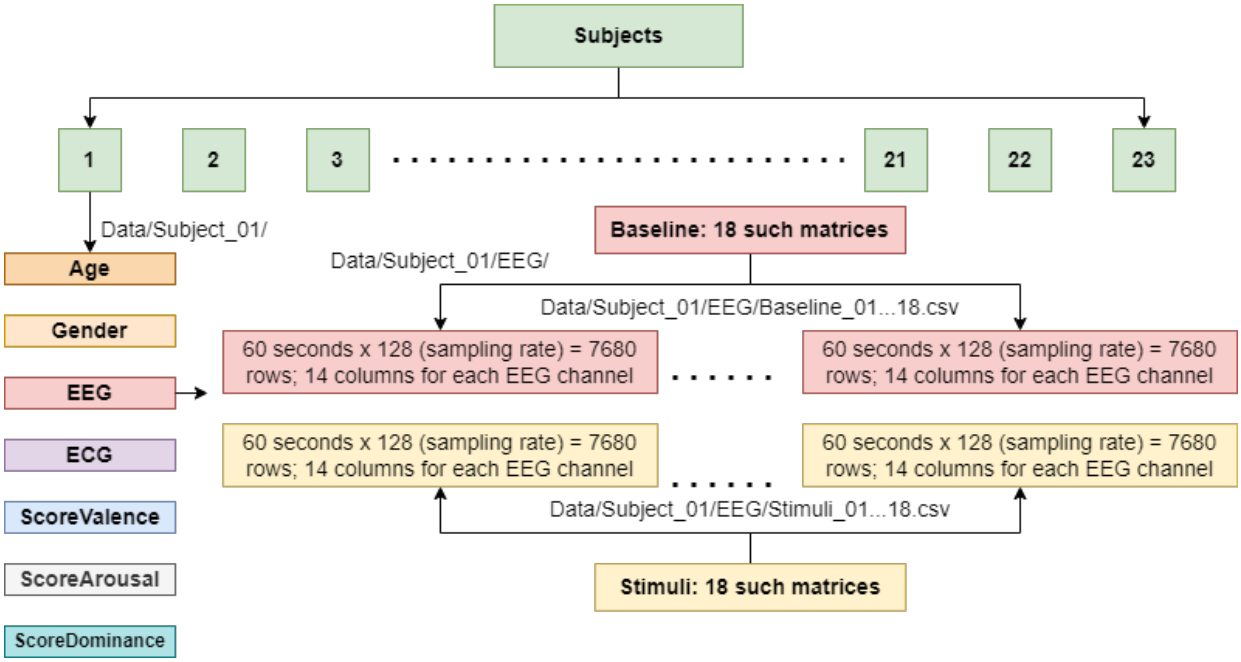


Fig. 4. DREAMER dataset overview

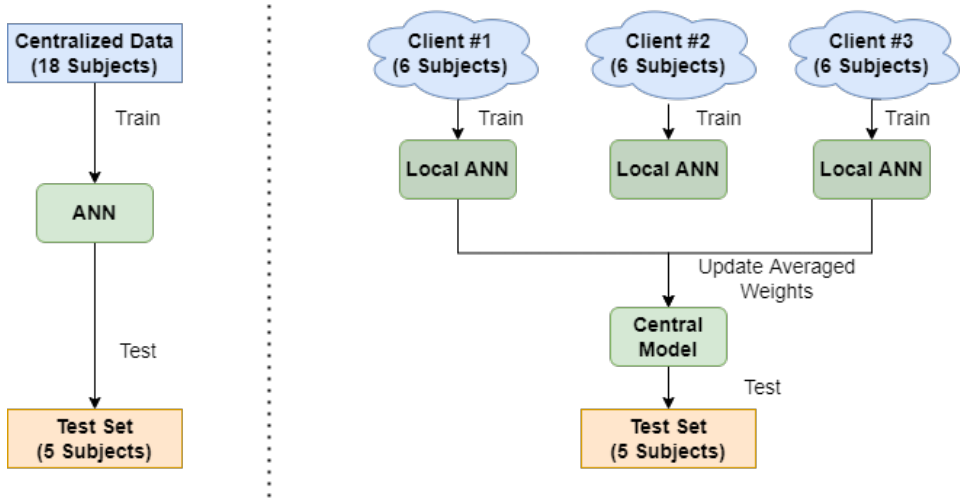


Fig. 5. Training Process

- The 14 electrodes of the signal and logarithms of PSD from bands are extracted as a feature [25], [26].
- Forty-two features are generated as a consequence (3 for each of the 14 electrodes). The final feature vector is created by concatenating all the features.

Let $P_{i\theta}, P_{i\alpha}, P_{i\beta}$ are the log of PSD for the signal of an electrode where i = 1, 2, ..., 14 of the θ, α, and β bands, respectively.

Then, the final feature vector will contain $14 \times 3 = 42$ columns as shown in (1).

$$P_{EEG} = [P_{1\theta}, P_{1\alpha}, P_{1\beta}, \ldots, P_{14\theta}, P_{14\alpha}, P_{14\beta}] \quad (1)$$

The proposed framework does binary classification for the labels: valence, arousal, and dominance. As part of the data preprocessing, step values greater than three are labeled as "high", and values less than or equal to three as "low."

C. Federated Data Division

For the proposed *FedEmo* framework, a subject-wise dataset division is done as EEG data is person specific. For the proposed framework, there is an 18-5 split for the 23 subjects, i.e., the training set comprises the data of 18 subjects, and the remaining five subjects contribute to the testing set. The train-test division is done randomly. The experiment is repeated for three different random seeds for the train-test split, and the best result is selected.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 02,2023 at 06:43:00 UTC from IEEE Xplore. Restrictions apply.

1) Base Model: The data for the 18 trained subjects is directly input into the Artificial Neural Network (ANN) model described in the following section. We then evaluate the trained model using the 5 test subjects.

2) Federated Model: The data for the 18 trained subjects were divided into three clients, each getting data from 6 subjects. After training, the model is evaluated centrally with the 5 test subjects.

D. Architecture of Proposed FedEmo Framework

This work uses a Sequential ANN model as the baseline with and without Federated learning. We employ the ANN architecture as it tends to work well with the Federated Averaging algorithm (*FedAvg*) [15].

- Sequential ANN model with 42 input features which are the PSD features (14 channels * 3 (alpha, beta, theta))
- The model comprises 4 Dense Layers as hidden layers; chronologically, the units are- 32, 16, 8, and 4.
- A Dropout layer follows every hidden layer with a rate of 0.8.
- The 5th and final hidden layers have two units.
- Each Dense layer has the Rectified Linear Activation Unit ('ReLU') as the activation function.
- Finally, the output layer has one output and uses the sigmoid activation function.

E. Federated Learning Environment

The proposed framework uses TensorFlow Federated (TFF) for the federated learning environment. It takes advantage of the horizontal federated learning approach [27]. After the dataset division, the framework distributes the data to the three virtual clients. It simulates a federated learning environment by creating a federated averaging process using TFF. The framework uses the standard federated averaging algorithm [15]. During each iteration of backpropagation, it uses the global weight update, as given in (2). For n clients and the i'th iteration:

$$w^i_{global} = \frac{1}{n}\sum_{j=1}^{n} w^i_j \tag{2}$$

where w^i_{global} is the global weight of the i'th iteration and w^i_j is the local weight of the jth client.

F. Evaluation Metrics

For both models, we use the following evaluation metrics:

- *Loss function*: Binary Crossentropy
- *Metrics*: Binary Accuracy

The framework uses the Adam optimizer [28] in the federated and standard environment for the baseline model. Since we can set different learning rates for the client and server optimization, we settle on the learning rates given in Table II after hyperparameter tuning.

TABLE II
LEARNING RATES FOR ADAM

Base Model	Federated Model	
	Server	Client
0.001	0.02	0.05

III. RESULTS AND DISCUSSION

The novelty put forward by this paper is the introduction of the Federated Learning paradigm to the field of emotion classification and EEG analysis. The emotion of a subject is classified using the three-dimensional model proposed by Mehrabian and Russell. We perform three binary classification tasks- each subject can have a low or high valence, low or high arousal, and low or high dominance values.

This work uses the same ANN architecture and compares the results achieved in a standard deep learning environment and a federated learning environment. It is a direct comparison between the usual approach, which works at the cost of data privacy, versus a more privacy-preserving approach. The aim is to have the federated paradigm achieve parity with the standard deep learning paradigm. We believe it will open up many new opportunities for hospitals to share their computations instead of confidential patient data to train meta-models.

Table III presents the binary classification performance of the standard and the federated models. Despite following a completely different training process and not having central access to the whole dataset, we observe that the federated model performs nearly as well as the standard deep learning model.

- For the valence label, we achieve the same accuracy of 63.33%.
- For the arousal label, FedEmo has a marginally higher score of 56.7%.
- For the dominance label, we achieve nearly the same accuracy score of 52.2%.

TABLE III
RESULT COMPARISON: BASE MODEL VS. FEDERATED MODEL

Label	Base Model	Federated
Valence	63.33%	63.33%
Arousal	56.6%	56.7%
Dominance	52.22%	52.2%

IV. CONCLUSION

This work aims to bring data privacy and emotion recognition together into a synonymous simultaneous system that predicts human emotions using technology without the expense of handing out proprietary and private information of any individual. Federated Learning techniques emerge as a future advancement in data privacy that allows collaborative learning more than ever without losing private information.

The proposed *FedEmo* framework proposes a privacy-preserving framework for emotion recognition using EEG data

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 02,2023 at 06:43:00 UTC from IEEE Xplore. Restrictions apply.

from the DREAMER database. The end-to-end deep learning model yields comparable results in and out of the federated environment. We process the EEG signals to extract the PSD of the θ, α, and β bands for all 14 EEG channels. In the federated learning setup, we split the data of the 23 subjects in the DREAMER dataset by allocating six subjects each to the three virtual clients for training. The remaining five subjects are then used to test the proposed model. We simulate the federated averaging process using TensorFlow Federated. The proposed sequential ANN model achieves an accuracy of 63.33 % for the Valence label, 56.7% for the Arousal label, and 52.22% for the Dominance label. *FedEmo* achieves these accuracies at par with the Base Model while providing the essential data privacy for the EEG data in the DREAMER dataset. *FedEmo* can bridge the gap between the security of emotional information and its usability in a distributed and decentralized manner for research and commercial purposes.

A. Future Work

Human emotion recognition is still in the nascent research and development stage. Future research can utilize larger datasets and robust emotion-capturing sensors, which may lead to greater accuracy metrics and the ability to cover a more complex set of emotions. Developing multimodal emotion recognition systems that combine other physiological signals like GSR, ECG, voice patterns, and/or facial expressions can also be an interesting research directive in the future.

In the future, we may try to combine one or more publicly available physiological datasets (such as AMIGOS [29] and DEAP [26]) to develop a generalized and privacy-preserving emotion classification model. We can also explore the performance of Federated Learning with other deep learning architectures, including LSTM, Neural Networks, and a new class of models based on RNN [30].

Further studies could try to tune *FedEmo* so that it can assist in the deployment of emotion recognition systems in the real world that would help tackle several problems: such as in the healthcare industries, provide valuable insights on marketing campaigns, aid in the improvement of consumer experiences, and better emotional care and psychiatric treatment of people dealing with emotional distress.

REFERENCES

[1] P. Ekman, "An argument for basic emotions," *Cognition & emotion*, vol. 6, no. 3-4, pp. 169–200, 1992.

[2] Y. Athavale and S. Krishnan, "Biosignal monitoring using wearables: Observations and opportunities," *Biomedical Signal Processing and Control*, vol. 38, pp. 22–33, 2017.

[3] B. O. et al., "Emotion classification based on biophysical signals and machine learning techniques," *Symmetry*, vol. 12, no. 1, p. 21, 2019.

[4] P. J. Lang, "The emotion probe: Studies of motivation and attention." *American psychologist*, vol. 50, no. 5, p. 372, 1995.

[5] J. A. Russell and A. Mehrabian, "Evidence for a three-factor theory of emotions," *Journal of research in Personality*, vol. 11, no. 3, pp. 273–294, 1977.

[6] A. Mehrabian and J. A. Russell, *An approach to environmental psychology.* the MIT Press, 1974.

[7] T. Tuncer, S. Dogan, and A. Subasi, "Ledpatnet19: Automated emotion recognition model based on nonlinear led pattern feature extraction function using eeg signals," *Cognitive Neurodynamics*, pp. 1–12, 2021.

[8] F. Galvão, S. M. Alarcão, and M. J. Fonseca, "Predicting exact valence and arousal values from eeg," *Sensors*, vol. 21, no. 10, p. 3414, 2021.

[9] J. Cheng, M. Chen, C. Li, Y. Liu, R. Song, A. Liu, and X. Chen, "Emotion recognition from multi-channel eeg via deep forest," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 2, pp. 453–464, 2020.

[10] H. P. Martinez, Y. Bengio, and G. N. Yannakakis, "Learning deep physiological models of affect," *IEEE Computational Intelligence Magazine*, vol. 8, no. 2, pp. 20–33, 2013.

[11] H. Tang, W. Liu, W.-L. Zheng, and B.-L. Lu, "Multimodal emotion recognition using deep neural networks," in *International Conference on Neural Information Processing*. Springer, 2017, pp. 811–819.

[12] S. Zhang, L. Sun, X. Mao, C. Hu, and P. Liu, "Review on eeg-based authentication technology," *Computational Intelligence and Neuroscience*, vol. 2021, 2021.

[13] J. Konečnỳ, H. B. McMahan, F. X. Yu, P. Richtárik, A. T. Suresh, and D. Bacon, "Federated learning: Strategies for improving communication efficiency," *arXiv preprint arXiv:1610.05492*, 2016.

[14] H. Brendan McMahan, E. Moore, D. Ramage, S. Hampson, and B. Agüera y Arcas, "Communication-efficient learning of deep networks from decentralized data," *arXiv e-prints*, pp. arXiv–1602, 2016.

[15] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y. Arcas, "Communication-efficient learning of deep networks from decentralized data," 2016. [Online]. Available: https://arxiv.org/abs/1602.05629

[16] https://www.indiatoday.in/technology/news/story/new-data-breach-exposes-fingerprints-and-face-recognition-records-of-millions-of-users-1581182-2019-08-15, Last accessed on August 27, 2022.

[17] A. Nandi and F. Xhafa, "A Federated Learning Method for Real-Time Emotion State Classification from Multi-Modal Streaming," *Methods*, 2022.

[18] T. Hassani, "Federated Emotion Recognition with Physiological Signals-GSR," 2021.

[19] Koelstra et al., "DEAP: A Database for Emotion Analysis; using Physiological Signals," *IEEE transactions on Affective Computing*, vol. 3, no. 1, pp. 18–31, 2011.

[20] Sharma et al., "A Dataset of Continuous Affect Annotations and Physiological Signals for Emotion Analysis," *Scientific Data*, vol. 6, no. 1, pp. 1–13, 2019.

[21] M. N. Dar, M. U. Akram, S. G. Khawaja, and A. N. Pujari, "Cnn and lstm-based emotion charting using physiological signals," *Sensors*, vol. 20, no. 16, p. 4551, 2020.

[22] Katsigiannis et al., "DREAMER: A Database for Emotion Recognition Through EEG and ECG Signals From Wireless Low-cost Off-the-Shelf Devices," *IEEE Journal of Biomedical and Health Informatics*, vol. 22, no. 1, pp. 98–107, 2017.

[23] W. Tao, C. Li, R. Song, J. Cheng, Y. Liu, F. Wan, and X. Chen, "Eeg-based emotion recognition via channel-wise attention and self attention," *IEEE Transactions on Affective Computing*, 2020.

[24] J. Cheng, M. Chen, C. Li, Y. Liu, R. Song, A. Liu, and X. Chen, "Emotion recognition from multi-channel eeg via deep forest," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 2, pp. 453–464, 2020.

[25] M. Soleymani, S. Asghari-Esfeden, Y. Fu, and M. Pantic, "Analysis of eeg signals and facial expressions for continuous emotion detection," *IEEE Transactions on Affective Computing*, vol. 7, no. 1, pp. 17–28, 2016.

[26] S. Koelstra, C. Muhl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, "Deap: A database for emotion analysis ;using physiological signals," *IEEE Transactions on Affective Computing*, vol. 3, no. 1, pp. 18–31, 2012.

[27] H. Zhu, H. Zhang, and Y. Jin, "From federated learning to federated neural architecture search: a survey," *Complex & Intelligent Systems*, vol. 7, no. 2, pp. 639–657, 2021.

[28] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014. [Online]. Available: https://arxiv.org/abs/1412.6980

[29] J. A. Miranda-Correa, M. K. Abadi, N. Sebe, and I. Patras, "Amigos: A Dataset for Affect, Personality and Mood Research on Individuals and Groups," *IEEE Transactions on Affective Computing*, vol. 12, no. 2, pp. 479–493, 2018.

[30] A. Rahim, A. Sagheer, K. Nadeem, M. N. Dar, A. Rahim, and U. Akram, "Emotion charting using real-time monitoring of physiological signals," in *2019 International Conference on Robotics and Automation in Industry (ICRAI)*. IEEE, 2019, pp. 1–5.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 02,2023 at 06:43:00 UTC from IEEE Xplore. Restrictions apply.

Physica Scripta

ACCEPTED MANUSCRIPT

Flexible organic solar cell to power modern cardiac pacemakers: Versatile for all age groups, skin types and genders

To cite this article before publication: Paritosh Chamola *et al* 2023 *Phys. Scr.* in press https://doi.org/10.1088/1402-4896/acbac8

Manuscript version: Accepted Manuscript

Accepted Manuscript is "the version of the article accepted for publication including all changes made as a result of the peer review process, and which may also include the addition to the article by IOP Publishing of a header, an article ID, a cover sheet and/or an 'Accepted Manuscript' watermark, but excluding any other editing, typesetting or other changes made by IOP Publishing and/or its licensors"

This Accepted Manuscript is **© 2023 IOP Publishing Ltd**.

During the embargo period (the 12 month period from the publication of the Version of Record of this article), the Accepted Manuscript is fully protected by copyright and cannot be reused or reposted elsewhere.
As the Version of Record of this article is going to be / has been published on a subscription basis, this Accepted Manuscript is available for reuse under a CC BY-NC-ND 3.0 licence after the 12 month embargo period.

After the embargo period, everyone is permitted to use copy and redistribute this article for non-commercial purposes only, provided that they adhere to all the terms of the licence https://creativecommons.org/licences/by-nc-nd/3.0

Although reasonable endeavours have been taken to obtain all necessary permissions from third parties to include their copyrighted content within this article, their full citation and copyright line may not be present in this Accepted Manuscript version. Before using any content from this article, please refer to the Version of Record on IOPscience once published for full citation and copyright details, as permissions will likely be required. All third party content is fully copyright protected, unless specifically stated otherwise in the figure caption in the Version of Record.

View the article online for updates and enhancements.

This content was downloaded from IP address 14.139.251.98 on 16/02/2023 at 06:27

Flexible Organic Solar Cell to Power Modern Cardiac Pacemakers: Versatile for All Age Groups, Skin Types and Genders

Paritosh Chamola[1, a] and Poornima Mittal[1, b, *],
[1]Department of Electronics & Communication Engineering, Delhi Technological University, New Delhi, 110042, India
e-mail: [a]paritoshchamola@gmail.com, [b]poornimamittal@dtu.ac.in
[*]Corresponding Author: poornimamittal@dtu.ac.in

Abstract **–** Bio-medical electronic components execute an vital part in medical services. Powering these devices is a task. Thus, the biomedical electronic devices which are able to self-harvest and store power are in huge demand. Present pacemakers are powered by batteries which have limited volume for energy packing and are compulsory to be changed. This needs a surgical intervention and is costly, with attachment of complications and risk. The objective of this paper is to validate if a subdermal PPV-PCBM [poly (2-methoxy-5-{3',7'-dimethyloctyloxy}-p-phenylene vinylene) and {6,6}-phenyl C_{61} – butyric acid methyl ester] active layer bulk heterojunction (BHJ) organic photo-voltaic (OPV) device could power a cardiac pacemaker. Power yield of 0.05 milliWatts (mW), 0.45 milliWatts & 2.1 milliWatts for African, Asian & Caucasian skin tones are gained at 2-millimeter implementation depth, acceptable to operate cardiac pacemaker demanding approximate power of 10 microWatts. Additionally, results correspondingly display higher output power is generated if the skin is thinner and brighter.

Keywords **-** Organic photovoltaic cell, Sub-dermal, Biocompatible, Implantable, Pacemaker.

1. Introduction

Electronic implants are widely utilized for therapeutic treatment. Best example are cardiac pacemakers. Pacemaker work by setting the heart by simulation induced electrically. Power required for a pacemaker is 10 µW [1]. Primary batteries have a lifespan of about six to seven years [2] with replacement rate of upto 25% [3]. Thus, organic PV cells are best suitable for stated purpose because of their suitable power production, lower fabrication price and bio-compatibility forming them best suitable technique of onsite electricity production inside the body [4-6]. Recently, organic photo-voltaic (OPV) devices have appealed significant attention for their commercial application primarily due to their mechanical flexibility, lower fabrication cost and semi-transparent character. However, the operational lifetime of the devices presents a key concern. Moreover, most of the research papers reporting the best stable OPVs devices degraded within few years [7-11], which created an impression that organic photovoltaic devices have shorter operational lifespan and thus have a lesser chance of being commercially attractive in comparison to their silicon counterparts, in spite of the quick and successful acceptance of the organic light emitting devices (OLEDs) having lifespan of million hours [12]. This impression or myth was put to rest by Burlingame *et. al.* [13] by constructing a organic photovoltaic cell with T_{80} of 4.9×10^7 hours which is equal to 27,000 years of outdoor. T_{80} is time required for power conversion efficiency to reach 80% of its initial performance. Similarly, Xiang Xu *et. al.* [14] fabricated a OPV device with T_{80} of 34,000 hours which equals to a lifespan of 22 years. In 2018, X. Du *et al.* [15] investigated a polymer solar cell device with a T_{80} of 11,000 hours which is equivalent to 10 years of lifespan. For OPVs an estimated lifespan of 20 years can be considered as commercially competitive and suitable for long term applications.

Though majority of the reported literature on PPV:PCBM materials are on increasing the efficiency of the PPV:PCBM OPV cells, Krens *et. al.* [16] systematically studied the life time of PPV:PCBM bulk heterojunction OPV in respect of electrodes, atmospheric conditions and device preparation. Schuller *et. al.* [17] applied Temperature accelerated lifetime measurements to PPV:PCBM bulk heterojunction solar cell. Further, X. Du et. al. [15] analysed T_{80} lifetime for PCBM to be 7400 hours. Additionally, Burlingame *et. al.* [13] in their paper found that the OPV cells have potential to have a long operational lifespans, only if pure materials, right fabrication process and architecture of OPV devices are used. In the present article an organic photovoltaic cell is considered for subdermal implantation so as to power a cardiac pacemaker. The different skin's optical properties are considered from various literature to compute and analyze the effect on the power produced by the organic cell under different type of skin tones. Further, effect due to age & gender on sustainability of the organic PV device is also investigated.

The present paper is divided in seven different sections with introduction as the first section. The second section describe the basic idea and benefits behind using organic PV cell to be used for solar energy harvesting. The third section give an insight into the skin's optical properties. In the fourth section the tissue geometry for the implementation of the photovoltaic (PV) is described. The characteristic analysis of the cell is performed in the fifth section. Subsequently, result and analysis are performed in the sixth section. Lastly, digest of all the significant results obtained is presented in conclusion.

Accepted Manuscript

2. Flexible Organic Solar Cell for Subdermal Implant

The idea of combining different types of electronically engineered devices with animate organic life so as to accomplish desired appropriate medicinal result is defined as bioelectronic medicine [18]. This accomplishment of correct medicinal or therapeutic result requires an extremely unfailing, steady and faithful interface between a living organic, potent biotic tissue surrounding and electronically dynamic inanimate device. Materials that could optimize the functionality of these electronic devices requires diverse aspects of design into contemplation as presented in Table I [19]. The living tissue comprises of complex, active mix of biomolecules, lipids and proteins etc. and majority of tissues are enormously soft. In comparison all inorganic materials are hard and dense. As presented in Table I, organic polymers due to their structure, composition, mechanical flexibility and charge transport mechanism are exceptional materials to be considered for biotic living tissues and abiotic electronic devices interfaces required for biomedical electronic implants or devices.

Table I. Characteristic of organic photovoltaic cell & in-organic inserts in evaluation to biological tissue.

Characteristics	**Inorganic Implants**	**Organic photovoltaic Cell**	**Tissue**
Composition	Inorganic semiconductors & metals	Polymers & Organic molecules	Protein, lipids & nucleic acid
Young's modulus	100 GPa	20kPa - 3 GPa	10 kPa
Charge transport carriers	Holes and electrons	Ions, holes & electrons	Ions
Physical solid state	Hard	Soft	Soft

Photovoltaic cells harvest energy from external sources [20]. The PV cells convert incoming irradiance energy to electric current within body [21]. In BHJ solar cell, charge accepter & charge donor materials both are blended simultaneously. The architecture for an BHJ cell is shown in Figure 1.

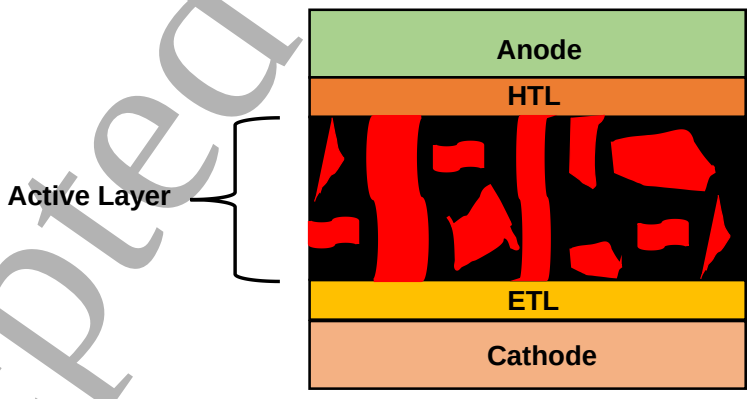


Fig. 1. BHJ (Bulk Heterojunction) architecture of organic solar cell.

The subdermal organic solar have a lifespan of upto 25 years [22]. Various harvesting methods [23, 24] are considered to use energy sources inside the body for instance, electrochemical responses [25], wireless energy [26], motion [27] etc. This paper tries to see the potentials of organic solar cell's use as a power supply source for implants to over-come restrictions of lithium batteries such as low harvested energy [28, 29], reactions [29], invasive implants [29], external power source requirement [28]. Only small quantity of daylight is capable to breach skin, predominantly near Infrared spectrum displaying decent penetrance of human skin [30]. Table II represents diverse energy harvesting practices.

Accepted Manuscript

Table II. Energy harvesting methods & their benefits & detriments.

Energy Harvesting Method	Advantages	Disadvantages	References
Piezoelectricity	Greater power production	Not easily biocompatible Implant location issues	[31]
Electrostatic	Greater power yield	Higher output impedance	[32]
Lithium batteries	Compatibility with modern bio-medical microelectronic systems	Dimensions Material toxicity	[33]
Electromagnetic	Unlimited implantable locations	Complex fabrication	[34]
Photo-charging	High power output	Large dimension	[35]
Thermoelectricity	Unlimited lifetime	Low power output	[36]
Nuclear battery	Extensive lifespan Continuous power production	High cost Radioactive Risks	[33]
Bio-fuel cell	Biocompatibility with tissues	Lifespan Lower energy production	[37]

In present paper, a complete analysis is made to estimate the opportunity of PPV-PCBM organic PV devices for harvesting solar energy for diverse tones of human skins at diverse implantation depth.

3. Optical Properties of Skin:

The primary function skin is to guard inner tissues. Skin is a complicated structure with various parameters & optical properties. The thicknesses & refractive index of skin are represented in Figure 2.

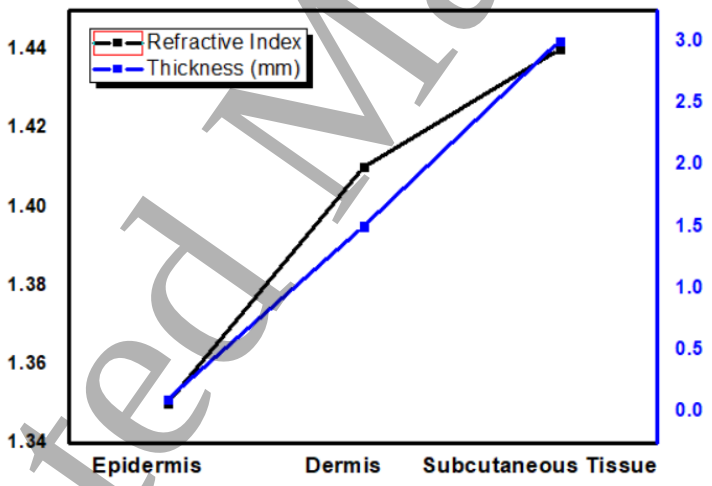


Figure 2. Refractive index & thickness of skin's different layers.

Skin encompasses of 3 layers (i) epidermis (ii) dermis & (iii) sub-dermis or subcutaneous tissue. Optical properties of skin include reduced scattering coefficient (μ_s), anisotropy factor (g), refractive index (n) and absorption coefficient (μ_a). Each of the 3 layer have distinct optical properties reliant on wavelength range. Optical properties are stated in several works [38-40]. Human skin is not transparent only a percentage of daylight from 700 nm to 1000 nm can breach the skin upto a implementation depth of 4 mm [41].

Absorption coefficient (μ_a) is calculated in accordance with the equation (1)[40].

$$\mu_a(\lambda) = 519\left(\frac{\lambda}{500}\right)^{-3} \quad (1)$$

The coefficient of absorption of blood [42], water [43], and fat [44] are stated in literature. Skin's coefficient of absorption is calculated via summing component's of individual absorption coefficient in harmony with volume fractions (equation 2, Figure 3) [40,1].

Accepted Manuscript

$$\mu_a = \sum VF_k \mu_{a,k} \quad (2)$$

Jacques [40] presented equation for coefficients of reduced scattering for skin and subcutaneous tissue.

$$\mu_s' = a\left(\frac{\lambda}{500}\right)^{-b} \quad (3)$$

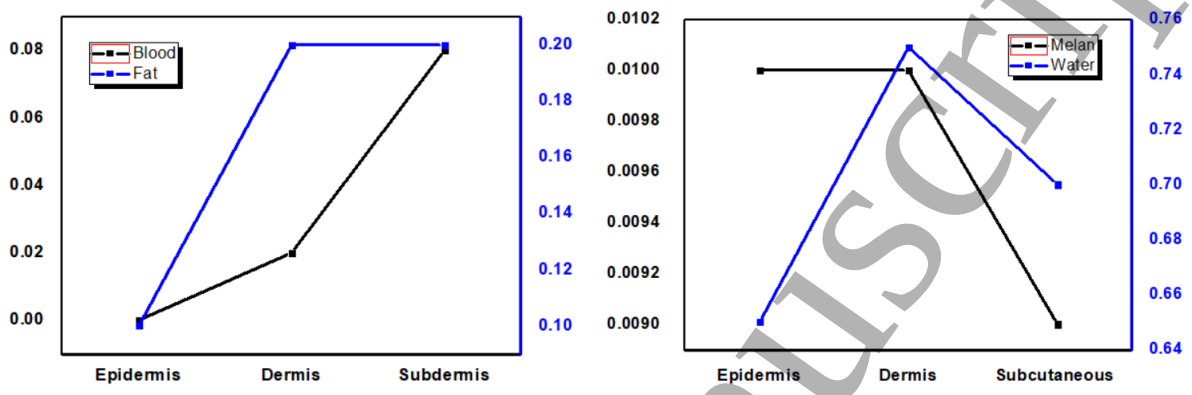


Figure 3. Volume fractions of skin components (a) Blood and Fat (b) Melan and Water in the three layers [45, 46]. (Melanosome abbreviated as Melan)

The a & b coefficients are the scattering power scale and coefficients are represented Figure 4.

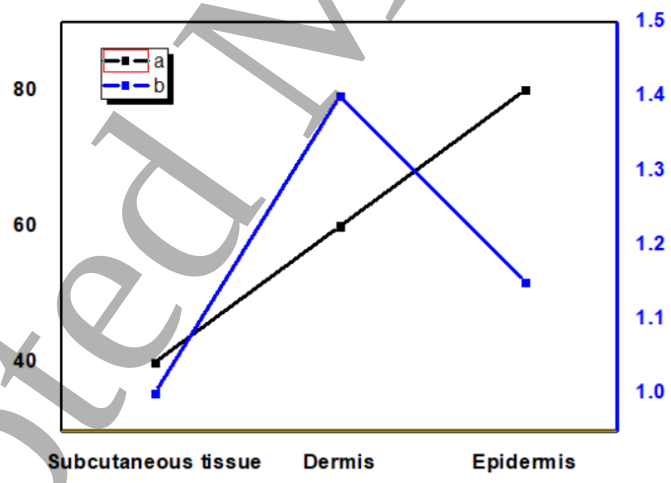


Figure 4. Skin's scattering coefficients (a & b) [40].

The factor of anisotropy measured is 0.9 for whole skin layers. Further, Refractive index is 1.41 for dermis & subcutaneous tissue & 1.34 for epidermis [47].

4. Tissue Model Geometry for Photovoltaic Cell Implantation

Tissue geometry is signified in Figure. 5. Tissue model is realized by dividing it into three layers (a) epidermis (b) dermis & (c) subcutaneous tissue. Human skin can be viewed as multilayered assembly with dissimilar layers categorized by optical properties of skin [46]. Thickness of Epidermis layer is considered to be 0.1 millimeter & of dermis is 1.4 millimeter. The rest of the model is subcutaneous tissue layer, which is site for photovoltaic device placement [48].

Accepted Manuscript

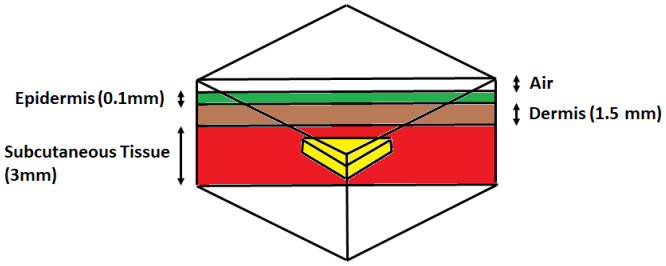


Figure. 5. View of tissue model geometry demonstrating organic solar cell implantation site inside skin. The organic solar cell IS presented by yellow colour. The distinct skin layers i.e., epidermis, dermis and subcutaneous tissue are shown by green, brown and red colour respectively.

The volume of sunlight absorbed by solar cell conditional on the placement of organic PV cell (ς) & wavelength. Computed using equation (4) [49,1].

$$\varsigma = \frac{N_{ABS}\,(Incident\ photons\ absorbed) \Big/ A_{TS}\,(Top\ surface\ area)}{N_{TOT}\,(Total\ projected\ photons) \Big/ A_{MS}\,(Surface\ area)}, \tag{4}$$

Where, N_{ABS} is sum of incoming photons absorbed by PV device, A_{TS} is top surface area of PV device. N_{TOT} is sum of overall photons projected & A_{MS} is area of the model surface. The pure atmosphere, irradiance is defined by global standard spectrum Air Mass 1.5 (AM 1.5) [50]. Irradiance power associated with Air Mass 1.5 is 100 mW/cm^2. Photon flux (ϕ) is calculated by dividing irradiance by energy of photon. Each PV cell have their individual External Quantum Efficiency (EQE). EQE dependent on wavelength & represents spectrum where organic PV cell show its finest performance. Furthermore, individual PV cell is denoted by its open-circuit voltage (V_{OC}). Short-circuit current (I_{SC}) for a PV cell is stated as [1]

$$I_{SC} = \int_{\lambda=400\ nm}^{750\ nm} \phi \times \varsigma \times EQE \times Q \times d\lambda \tag{5}$$

The Q represents charge. Power output (P_{OUT}) of a photo-voltaic device with area of A_{TS} is stated by equation (6). The parameter fill factor is associated with PV cells maximum power output [1].

$$P_{OUT} = FF \times V_{OC} \times I_{SC} \times A_{TS}, \tag{6}$$

A modern Cardiac pacemaker necessitates a normal operational power of roughly ten microwatts.

5. Characteristics Analysis for implanted organic solar cell

Study of BHJ organic solar cell arrangement is executed using software simulator Silvaco-Atlas-TCAD. The Silvaco Atlas integrates models and mathematical equations for thorough & comprehensive scrutiny of device's internal physics [51, 52]. Further, by defining the material properties new materials can be defined. The conditions of biasing can be utilized to attain various device parameters such as density of current, concentration of electron & hole, rate of recombination etc. The device dimensions are defined precisely with particular material regions specified by defining exact material properties such as energy levels, work function etc. Simulators built on computational logical models have recognized their significant and critical role to determine truthful behavior of the device and to improve device's performance. The precision of organic PV cell in Atlas TCAD simulation software initiates from accessibility of numerous optical transmission models of robust character.

In present paper Indium Tin Oxide (ITO); poly(3,4 ethylenedioxythiophene):poly(styrenesulfonate)(PEDOT:PSS); PPV/PCBM; Lithium Fluoride (LiF); Aluminium (Al) BHJ organic solar cell is utilized. PV cell's length is considered to be 1 micrometer. Computational modeling based logical device simulators are well-known to envisage and improve performance of a device beforehand its real-time application. Systematic simulation of organic PV cell computes elementary arithmetical and statistical equations with objective of discovering organic PV cell's performance and internal device physics. These basic computational mathematical statements define charge transportation. Atlas simulator could demonstrate processes which administers charge transportation, charge photo-generation and recombination in BHJ organic PV cell. TCAD simulator has integral AM 1.5 and AM 0 solar spectrum available in open public domain easily accessible by BEAM statement [53, 54].

The Figure. 6. represents design of a sub-dermal PV cell competent of harvesting the inbound sunlight of the sun. The projected solar power harvesting design includes of BHJ organic PV cell made up of ITO/PEDOT:PSS/PPV-PCBM/LiF/Al. The interior examination of device offers understanding about functioning of device internal physics. The layer thickness of PV cell is considered as to be 10 nm for Al, 0.6 nm for LiF, 80nm for PPV/PCBM, 20 nm for PEDOT:PSS and 10 nm for ITO respectively,

Accepted Manuscript

Further, human skin is considered to be 3 mm thick. To imitate appropriate human-skin under spectrum of solar illumination extra 3 layers of aggregate 2 millimeter (depth) with particular optical properties of the skin i.e., absorption (μ_a) & scattering (μ_s) are included at the topmost of PPV-PCBM organic PV cell as shown in Figure 6 This entire arrangement functions as area of the skin which covers organic PV cell irradiated by daylight and after that software simulation is performed. The average output power in range of milliwatts (mW) is attained, in contrast to that the total power spent by a modern cardiac pacemaker is roughly about 10 microwatts [1].

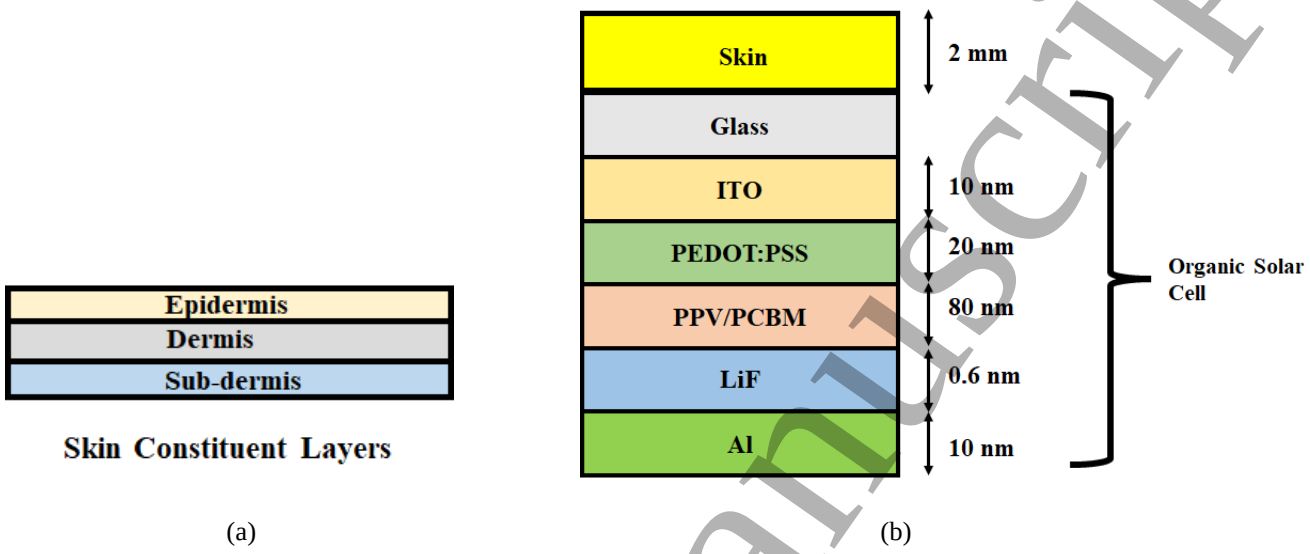


Figure 6. (a) Top layer imitating skin with constituent layer optical properties. (b) Device assembly of organic solar cell when implanted beneath human skin.

The performance of proposed and planned organic PV cell can be authenticated via measuring the characteristic I-V curve & other solar PV parameter for instance open-circuit voltage (V_{OC}). The PPV/PCBM BHJ organic PV cell is simulated using Silvaco Atlas [55, 56] an industry standard device software simulator. Further, the photovoltaic cell is investigated with the objective to authenticate its performance with an similar experimentally investigated device [57]. The device simulation parameters are considered as the electron mobility is taken to be 2.5×10^{-7} m^2 V^{-1} s^{-1}, mobility of hole is taken as 0.3×10^{-7} m^2 V^{-1} s^{-1}, energy bandgap taken as 1.34 eV, dielectric constant taken as 3×10^{-11} Fm^{-1} and density of state is taken to be as 2.5×10^{25}.

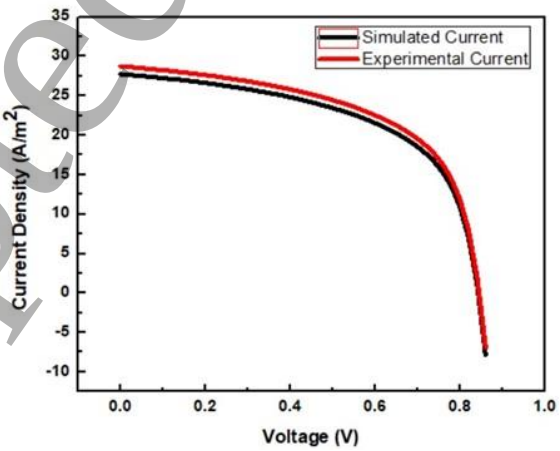


Fig. 7. I-V characteristics plots of the solar cells the red colour curve is experimental cell's current & black colour curve is simulated cell's current.

The Figure 7 represents current-voltage curve of simulated (black curve) and experimental (red curve) of OCC-PPV/PCBM organic PV cell. The extreme value of short circuit current density (J_{SC}) obtained is 28 A/m^2. Further, experimentally stated magnitude of open-circuited voltage (V_{OC}) is 0.90 volt & simulated magnitude of open circuit voltage is achieved as 0.84 volt which is in near approximation to experimentally gained magnitude. The performance parameters obtained from the organic PV cell are reported

Accepted Manuscript

as Fill Factor of 52.5%, efficiency of 1.22%, J_{SC} of 28 A/m^2 and V_{OC} of 0.8 Volt. This outstanding agreement in-between the simulated and experimental current-voltage characteristic curves allow us to advance with our investigation of PPV-PCBM organic PV cell to check if it's appropriate for our intended application of sub-dermal PV cell which can be used to satisfactorily harvest incoming solar energy for running a cardiac pacemaker.

6. Results and Analysis

6.1 Absorption and Scattering Coefficients

As one particular literature didn't provide all absorption and scattering coefficients for the three types of skin i.e., Caucasian, African and Asian in the range of 400 to 700 nm, multiple literatures are considered, combined and inferred to obtain the coefficients to be used for the simulation so as to obtain accurate results. For instance, A. N. Bashkatov *et al.* [39] is referred for absorption coefficients and S. L. Jacques [40] is referred for reduced scattering coefficients of the skin for the range of 400 to 750 nanometer and 500 - 750 nm respectively. Additionally, reduced scattering coefficients from A. N. Bashkatov *et al.* [39] are used the range of 400 to 500 nm as these coefficients' ranges were not included from the S. L. Jacques [40] literature. Similarly, C. R. Simpson [45] *et al.* literature is used to determine the absorption and scattering coefficients of Caucasian & African skin type in the range of 600 to 750 nanometer & for the missing coefficients from 400 to 600 nm data from literatures such as A. N. Bashkatov *et al.* [39], Van Veen *et al.* [44] etc. is considered. In the same manner some literature measured the coefficient data for dermis layer but not for subdermal layer or subcutaneous tissues. Therefore, this complicated issue of not having all the parameters related to the three skin types in one specific literature lead us to combine all the best accepted and available literatures so as to form a consolidated data for simulations so as to achieve valid and accurate results.

A. N. Bashkatov *et al.*[39] measured the absorption and scattering coefficients of subcutaneous adipose tissue and mucous in range of 400-2000 nm. The measurements were done invitro with the samples of skin collected from post-mortem & fresh subcutaneous adipose tissue which were taken from patient peritoneum area while surgery. For experimental data processing and tissue optical property determination method in Prahl *et al.* [58] is utilized which allowed to determine absorption (μ_a) and scattering (μ_s) coefficients. Chan *et al.* [38] measured the in vitro optical properties of the tissue for photomedicine within the visible and IR range i.e. from 400 to 1800 nm. For the measurements specimens of Caucasian and Hispanic donor skin were used from the skin bank. Harvesting of tissues was done within 24 hours of postmortem. The three specimens taken were male Caucasian donor of age 53, male Hispanic donor age 35 and female Caucasian donor age 72. Both the above papers have been widely accepted and used in research papers associated with the biomedical and therapeutic treatment studies. Moreover, In 2021 Finlayson *et al.* [59] used Monte Carlo simulations for light propagation via skin for a range of 200 to 1000 nm. The deviation in data obtained from the [38, 39] is minimal.

Additionally, Finlayson *et al.* used the equations developed by [39] for their simulations which confirms that though the reference [38, 39] are not the most recent available ones but are accepted worldwide by the people for their research. Similarly, Kono *et al.* [60] also studied the in vivo scattering and absorption coefficients of 198 Japanese volunteers. They further analysed the effects of gender and age. Though the absorption and scattering coefficients vary with the gender and age the variations in optical properties do not present any trend of high deviation. In 2016 Bereuter *et al.* [49] studied the subcutaneous photovoltaic energy harvesting and concluded that persons above the age of 65 years achieved more power in comparison to the younger population due to more leisure time and spending more time under the sun. As transfer of power to biomedical implant via wires piercing the skin might put the patient at the hazard of infection. Thus, producing power underneath the human skin could evade the risk. The organic photovoltaic cell can be implanted underneath human skin of hands (at inner forearm, inner or dorsal surface of hands and area in-between the thumb & forefinger) or neck region. The fixation of implant location in the body primarily depends on the amount of incident daylight received by the body's area.

In 2010 Firooz *et al.* [61] studied the variation of skin parameters with respect to age and gender by taking fifty healthy persons both males and females in five age groups from 10 years to 60 years in Iran and reported that the gender is independent of melanin index, as melanin is a dominant aspect in determination of absorption coefficients it can be stated that gender and absorption coefficients are unrelated, but in 2019 Kono *et al.* [60] reported that there is a slight difference (~ 0.4) in the absorption coefficients for Japanese males and females (considering 198 volunteers), but as our solar cell is capable enough to produce sufficient power for the African skin types i.e., 50 microwatts (µw) with maximum absorption coefficients even if we consider this slight variation as standard the power produced my our photovoltaic cell is approximately 5 times more than the operational power (of ~ 10 µw) required by modern cardiac pacemakers. Thus, it can be concluded that even if there are minor variations in the absorption coefficients due to gender our solar cell can sustain those variations and can sufficiently power the pacemaker. Similarly, for age Mayes *et al.* [62] in accordance with the study conducted in China reported that no connection between pigmentation of skin and age was found. However, Kono *et al.* [60] reported slight variation of absorption coefficient with age which can be sufficiently managed and taken care by our implanted solar cell. Thus, the subdermal solar cell is absolutely suitable for the old age patients and even for immobile or disabled people by artificial lighting.

Accepted Manuscript

Figure 8 (a) represents the literature data for coefficients of absorption (μ_a) of top (epidermis) human skin layer for Caucasian, Asian and African type of skin as a function of wavelength ranging from 400 - 750 nanometers (visible light band). Likewise, absorption coefficients of middle (dermis) & bottom skin layer (sub-dermis tissue) as a function of wavelength are represented in Figure. 8 (b) & Fig 8 (c). The figures represent plot of optical properties related to the Caucasian, Asian and African epidermis layer together for comparison [45, 63]. The absorption coefficients show the amount of daylight absorbed at any definite wavelength. Absorption data for Caucasian, Asian and African skin tones are plotted as blue, red and black colour correspondingly. From the plots it could be inferred that value of absorption coefficients for African skin tone are higher in comparison to Asian and Caucasian skin tones, in the same manner magnitude of absorption coefficient for Asian skin tone are higher than Caucasian skin tone in visible spectrum and decrease gradually as wavelength increases [45]. The plots of absorption spectrum in Figure. 8 brings a precise assessment on sunlight absorbed by human skin and this can be utilized for approximation & examination of share of daylight captured by sub-dermal organic PV device.

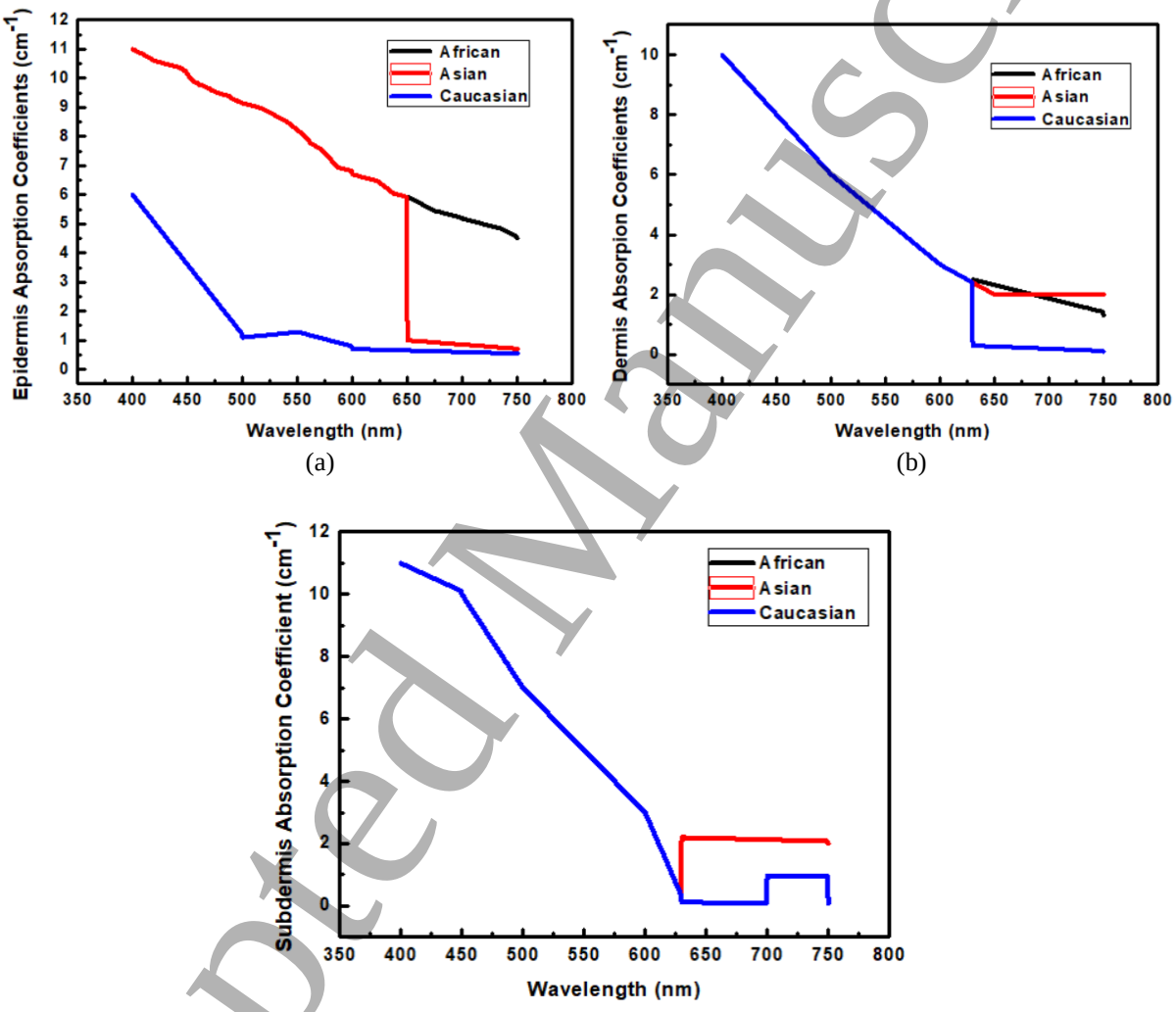


Figure 8. Coefficients of absorption (μ_a) of (a) epidermis (b) dermis (c) subcutaneous tissue verse wavelength ranges from 400 to 750nm.

Accepted Manuscript

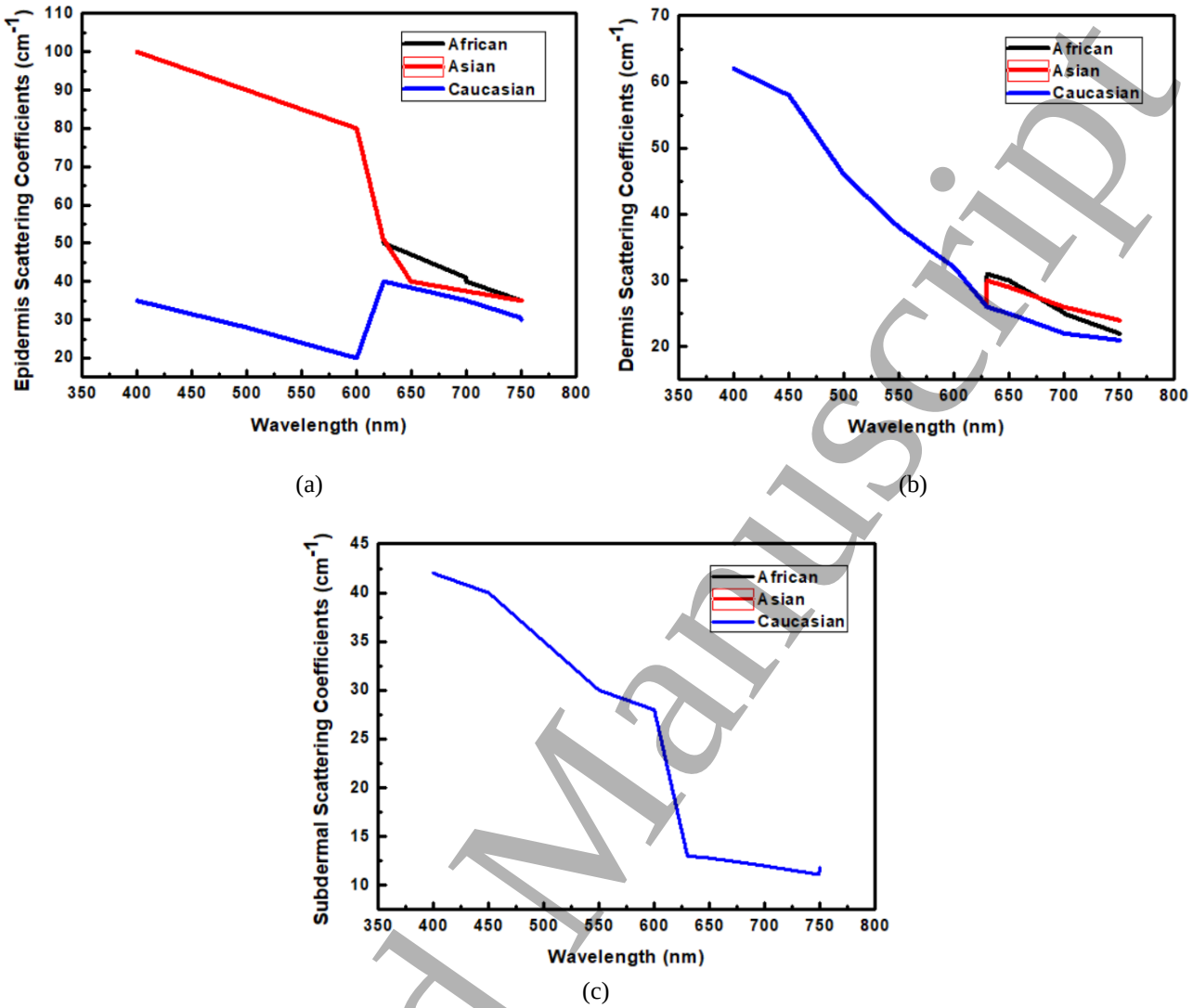


Figure 9. Coefficients of reduced scattering (μ_s') of (a) epidermis (b) dermis (c) subcutaneous tissue as a function of wavelength range from 400 to 750 nm.

Figure 9 (a) represents literature data for coefficients of reduced scattering ($\mu_s^{'}$) of epidermis layer for Caucasian, Asian and African skin tones versus (vs) wavelength in 400 to 750 nm band i.e., in visible light spectrum. Likewise, scattering coefficients for middle (dermis) and bottom (sub-dermis tissue) layer vs wavelength are shown in Figure 9 (b) and Figure 9 (c) [63]. The absorption data of Caucasian, Asian and African are plotted in blue, red & black colour correspondingly. The scattering coefficient illustrate that there are no substantial variances amongst the African, Asian and Caucasian skin types [45]. Scattering spectrum characterizes a sturdy decline with increasing wavelength. In summary, though optical properties of the skin very with age, gender and different regions of the world but these all variations are very minor and our organic solar cell is capable of sustaining these minor changes and can provide continues power to the pacemakers.

6.2 Skin Penetration Depth

Sunlight penetration depth of human skin is determined by the equation (7) [64]. The assessment of depth of penetration is utmost significant parameter for defining quantity of inbound solar illumination captured by human skin.

$$\delta = \frac{1}{\sqrt{3\mu_a(\mu_a+\mu_s')}} \tag{7}$$

The Figure 10 represents mathematically computed penetration depth of light achieved by means of reduced scattering & absorption coefficients. The plot in Figure 7 represents depth of light penetration vs wavelength and is for visible band (400 to 750 nanometers). From the Figure 10 it can be inferred that penetration depth rises with rise in wavelength. Highest depth of penetration of 2.1 millimeter is observed at wavelength of 750 nanometer as shown in Figure 10. Therefore, firmly representing that incident

Accepted Manuscript

sunlight of greater energy (of lower wavelength) has a lesser penetration depth in contrast to incident sunlight of lessened energy. This happens owing to the fact that lesser energy sunlight or photons aren't absorbed easily & thus have greater depth of penetration.

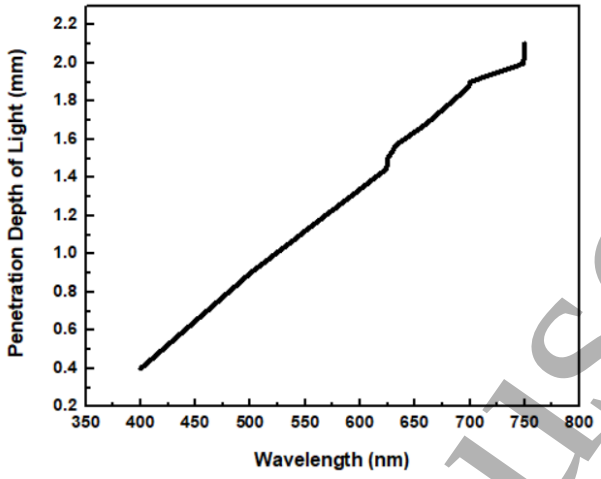


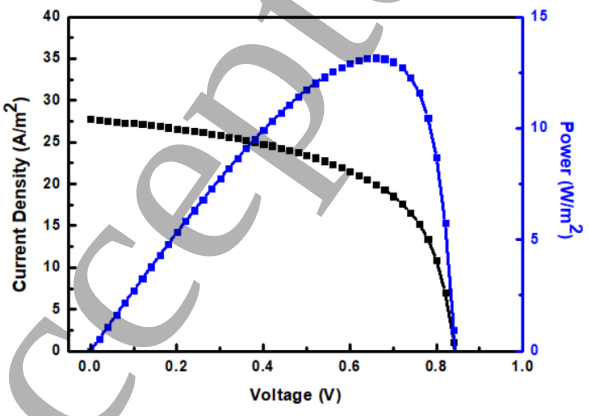
Figure 10. Sunlight penetration depth inside the skin for wavelengths of 400 - 750 nanometers.

To comprehend notion of depth of light penetration, the blue color light consisting greater energy & shorter wavelength would-be absorbed at top surface, but red color light of lesser energy & greater wavelength is absorbed in lesser amount. For example, in a photovoltaic cell build of silicon material blue color light will be absorbed in few mm however red color light will take few hundred mm before it is absorbed completely [65]. Additionally, from Figure 10 it is clear and apparent that penetration depth is dependent on wavelength. The penetration depth is highest in spectrum of 650 to 750 nanometers & sunlight can penetrate till the penetration depth of 2.1 millimeter which exceeds epidermis & dermis layer's thickness and then gets as far as to subcutaneous tissue.

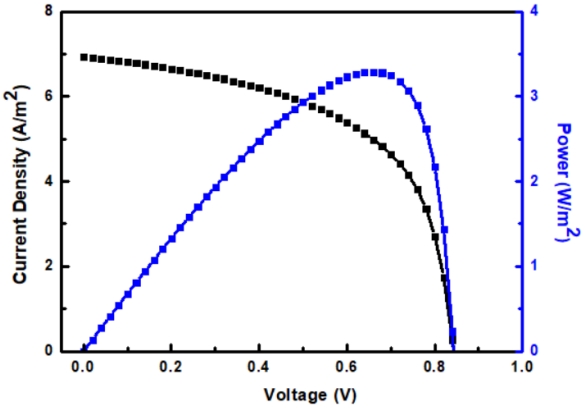
As transfer of power to biomedical implant via wires piercing the skin might put the patient at the hazard of infection. Thus, producing power underneath the human skin could evade the risk. The photovoltaic device can be implanted underneath human skin of hands (at inner forearm, inner or dorsal surface of hands and area in-between the thumb & forefinger) or neck region. The fixation of implant location in the body primarily depends on the amount of incident daylight received by the body's area.

6.3 Solar cell placement for all Skin Types

In subcutaneous implanted organic PV cell, energy harvested is restricted by the amount of light absorbed & scattered. The characteristic of sub-dermally implanted organic PV cell is calculated & noticed underneath the normal daylight illumination. To validate performance of sub-dermally implanted organic PV cell, numerous characteristic graphs are plotted for the measurement of characteristics of organic PV cells in varied settings such as current as a function of voltage (I-V) & power as a function of voltage (P-V) graphs.



(a)



(b)

Accepted Manuscript

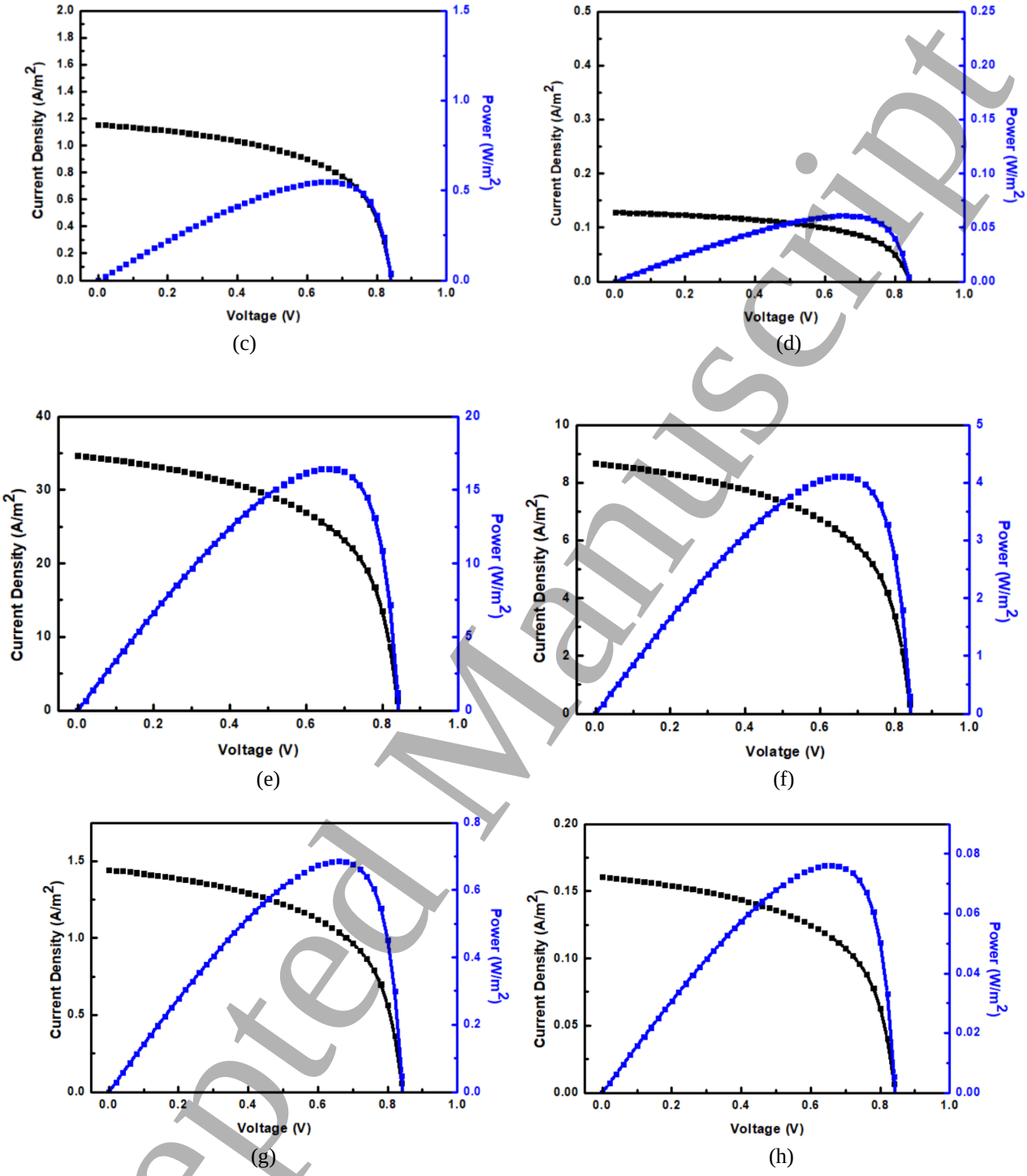


Figure 11. Characteristics I-V & P-V curves of organic solar device for (a) Uncovered solar cell (b) Caucasian (c) Asian (d) African skin tones attained for AM 1.5 irradiance condition, (e) Uncovered solar cell (f) Caucasian (g) Asian (h) African skin tones attained for AM 1.0 irradiance condition.

Figure 11 represents the simulated I-V plot attained from PPV-PCBM BHJ organic PV cell. The Current-Voltage plots presented in Figure 11 present comparisons of different photocurrents attained. The highest value of photo-current density gained for uncovered PV device equals to 28.1 A/m^2 & 6.8 A/m^2 for enclosed (underneath the skin) PV cell (for Caucasian tone of skin) at 0 volt. Figure 11 (a) show the I-V & P-V graphs of organic solar cell when subjected to AM 1.5 illumination (1000 W/m^2). The plots in Figures 11 show the characteristics of organic photovoltaic cell covered and uncovered. In Figure 11 (a) double Y plot the black color arc show current-voltage characteristic plot and blue color arc present the power-voltage characteristic curve if organic PV cell is set underneath direct incoming sunlight (uncovered PV cell). Similarly, Figure 11 (b), Figure 11 (c) & Figure 11 (d) represent the double Y plot (black color arc for current-voltage and blue for power-voltage characteristics) for Caucasian, Asian and African

Accepted Manuscript

skin tones once underneath the human skin at 2 mm depth or strictly speaking when PV cell is located in the subcutaneous layer beneath epidermis layer and dermis skin layers.

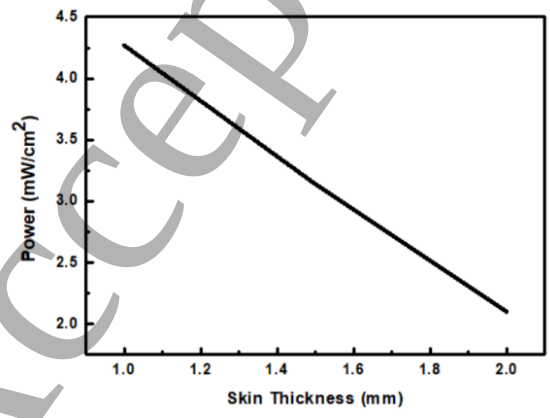
In Figure 11 (a) it can be noted that current is having a maximum magnitude of 28 A/m^2 if organic PV device is placed under direct daylight (uncovered solar cell). The magnitude of photocurrent declines by around six folds to magnitude of 6.8 A/m^2 just as the organic PV cell placed underneath Caucasian skin tone. The current's magnitude further declines to 1.2 A/m^2 if that organic PV cell is placed underneath Asian tone of skin, decrease in magnitude becomes more substantial if organic PV cell in placed underneath African skin tone. The main rational logic for aforesaid trend in I-V plot is due to coefficients of absorption of skin layers for African skin tone are maximum & greater in contrast to that of Asian and Caucasian tones of skin. Same can be said for Asian tone of skin compared to Caucasian skin tone. The P-V plots for skin tone likewise demonstrates same pattern i.e., P-V plot reaches peak for the uncovered organic PV cell & decreases for Caucasian skin tone & decreases moreover for the Asian & African skin tones because of the different absorption coefficients. As larger absorption coefficients absorb the incoming photons easily therefore, skin layers absorb much of the incoming light hence, leaving reduced quantity of photons to reach up to the implanted photovoltaic device to get absorbed and produce electric current.

In identical manner organic PV cell is simulated with AM0 illumination settings (1376 W/m^2). The Figure 11 (e), 11 (f), 11 (g) & 11 (h) present the I-V & P-V curve. The current produced by organic photovoltaic cell for AM 0 illumination is notably greater than that of AM 1.5 illumination hence, confirming the statistic that greater irradiance power results in greater generation of electric current in organic solar cell due to the greater amount of photon being absorbed, this can be congruently confirmed by linking the P-V curves for AM 1.5 and AM 0 illumination.

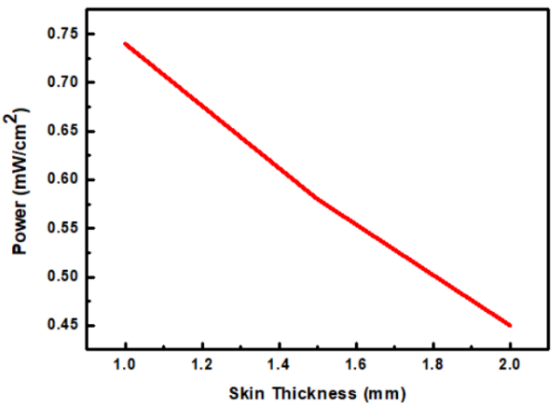
6.4 Performance Analysis on Skin Thickness Variation

Figure 12 presents relationship between power output of the implanted organic photovoltaic cell (covered) & skin thickness when varied from 2 to 1 millimeter. The black color, red color & blue color plots represent to the Caucasian, Asian & the African skin tones respectively. Moreover, form above curves represented in Figure 12 it can be inferred that at skin thickness or at the skin depth of 1 millimeter, the produced power is greatest for all three tones of skin & starts to decrease as the implementation depth or thickness rises from 1 - 2 millimeters. This establishes that power-density of sub-dermally implanted organic PV device is contrariwise related to thickness of the skin & that power-density depend on measure of light reaching the subdermal PV cell through human skin. The photovoltaic device generates more output power when kept in bright sunlight. Additionally, from graphs in Figure 12 it can as well inferred that different human skin tones of identical thickness can have different densities of power. The greatest achieved power of the Caucasian tone of skin is 4.28 milliwatts at an thickness of 1 millimeter & lowest power achieved is 2.2 milliwatts at 2 millimeter of thickness. Similar way greatest output power obtained for Asian & African tones of skin are 0.73 mW or 730 microwatts (µW) & 0.085 mW or 85 µW at thickness of 1 millimeter and lowest power attained is 0.44 milliwatt or 440 microwatt and 0.051 milliwatt or 52 microwatt at 2 millimeter thickness when put at AM 1.5 standard global spectrum settings.

Thus, from Figures. 10 to 12 it can be deduced that volume of power generated by sub-dermally implanted organic PV cell relies on volume of sunlight acquired by it, & volume of light acquired by PV cell through human skin relies on skin tone & skin thickness underneath which it's positioned. The darker tones of skin absorb higher incident sunlight hence, distributing reduced number of photons for the sub-dermally implanted PV cell which in turn generates lesser power even with the similar skin thickness. The dispersal of melanin pigmentation conspicuously regulates skin tone. The melanin pigmentation existing in skin's epidermis layer absorbs visible & ultraviolet spectrum band of light [66].

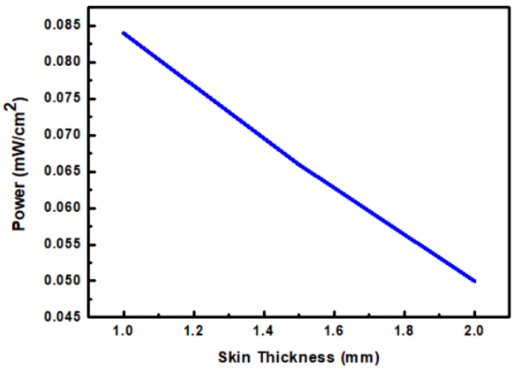


(a)



(b)

Accepted Manuscript

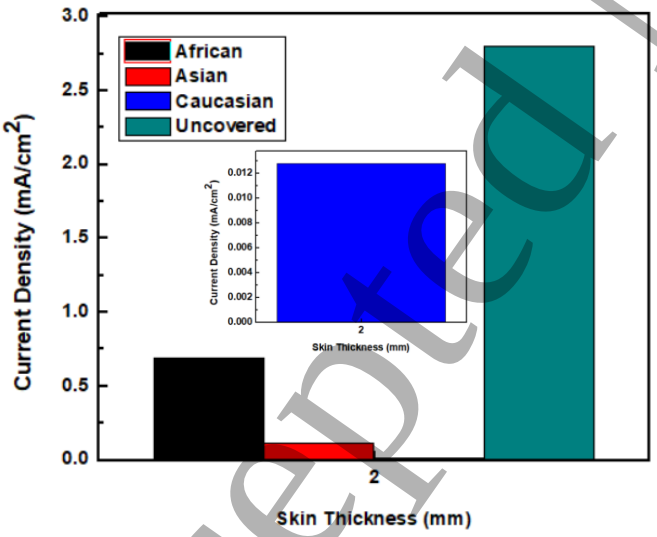


(c)

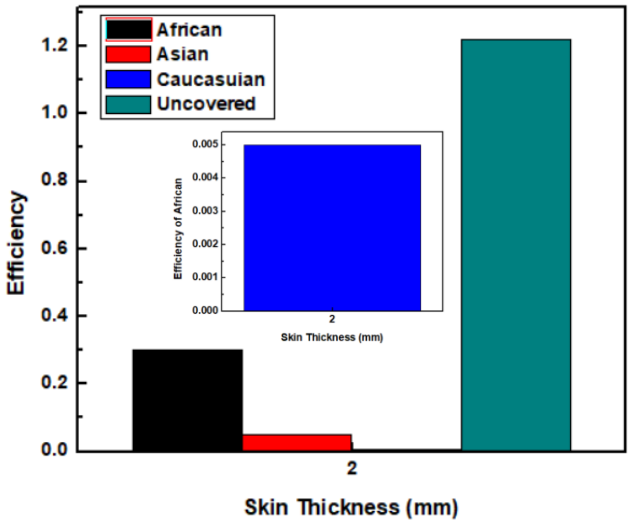
Figure 12. Power output verses thickness skin for (a) Caucasian (b) Asian and (d) African skin tones.

Darker skin tone have distribution of melanin pigmentation on higher side which ultimately drops sunlight carried to the sub-dermally implemented organic PV cell. Thicker skin scatter incident sunlight more thus decreasing the sunlight transported to the sub-dermally implanted PV cell which leads to lesser power production. The power produced by sub-dermal PV cell is equivalent to the prior reported works using piezoelectric cells (0.12-0.19 μW/cm^2) [67] or glucose-based bio-fuel cell (0.2-1.2 mW/cm^2) [68] & is acceptable enough to operate a cardiac pacemaker demanding an operative power of 10 to 40 μW.

Figure 13 shows the performance parameter characteristics of sub-dermally implanted organic PV device when put directly in daylight i.e., when its uncovered (presented with sea green color) & when its covered and kept in AM 1.5 irradiance, black color corresponds to Caucasian, red color corresponds to Asian & blue corresponds to the African skin tones. The Figure 13 (a) shows short-circuit current density (J_{SC}) of PV device at an implementation skin depth of 2 millimeter. Similarly, Figure 13 (b) displays the efficiency of PV device. The short-circuit current density & efficiency (η) of subdermal implanted PV cell declines from 1.21 % to 0.3 % (Caucasian skin tone) & from 2.9 mA/cm^2 to 0.698 mA/cm^2 (Caucasian skin tone) correspondingly.

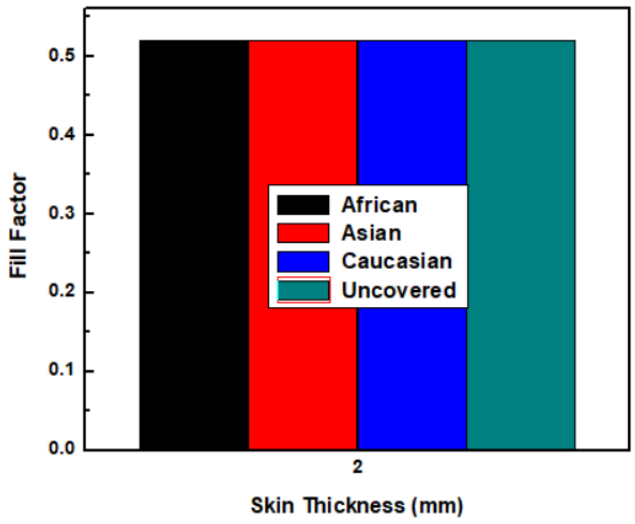


(a)

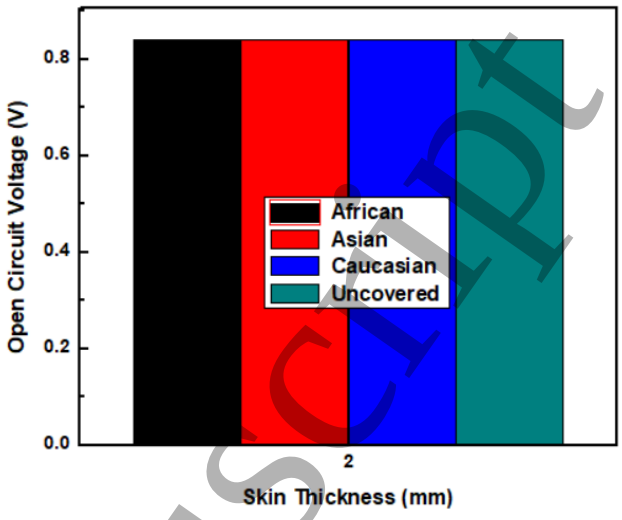


(b)

Accepted Manuscript



(c)



(d)

Figure 13. The graphical representation of (a) Circuit density (b) efficiency (c) Fill Factor & (d) Open circuit voltage at the sub-dermal skin depth of 2 millimeters.

The Figure 13 (c) represents Fill Factor (FF) of implanted PV device at an implementation skin depth of 2 millimeters. Likewise, Figure 13 (d) presents V_{OC} at 2 millimeters skin implementation depth. V_{OC} is maximum voltage generated by solar cell without exterior load connected. V_{OC} is primarily function of materials energy level, contacts & interface. Open-circuit voltage at 0.84 volts and fill factor at 52 % do not show any key deviation after implementation underneath the human skin as both are determined by properties of material of the solar cell.

6.5 Charging Time Analysis under different Illumination Conditions

The power requirement for functioning of a cardiac pacemaker is approximately of 10 µW for a period of twenty-four (24) hours it comes out to be 0.25 milliwatt hour (mWh) or 0.86 Joules, agreed by means of the relationship among power & energy in accordance with equation (8).

$$Eenergy = Power \times time \tag{8}$$

Maximum density of power attained for implanted sub-dermal organic PV device at skin depth of 2 millimeter is 2.2 mW/cm^2 (Caucasian tone), in view of area of solar cell to be equal to 1 cm^2 & energy needed to function cardiac pacemaker for twenty-four hours is equal to 0.24 milliwatt hour thus, by means of equation (8) time period required to completely charge a battery would be roughly around (six) 6 minutes. Therefore, it can be inferred that implemented organic PV device can adequately charge a re-chargeable battery of appropriate size for a pacemaker which can operate a cardiac pacemaker for twenty-four (24) hours devoid of any light at an skin depth of 2 millimeter & thus is capable of operating in the time duration of lesser illumination or dimness. Moreover, its implied that implanted sub-dermal organic PV device can by itself straightaway power cardiac pacemaker when placed in standard circumstances of illumination. Moreover, artificial lighting can as well be utilized to power cardiac pacemaker indoors as indoor lighting systems such as LEDs and halogen bulbs can generate substantial sum of power. For instance, a indoor halogen bulb emanates a spectrum from 400 to beyond 1100 nm & a LED light's spectrum is from 400 to 700 nm [69], & as implanted organic PV cell presented in this paper functions in spectral range of 400 to 750 nm since it has a bandgap energy of 1.34 eV, the photovoltaic cell can produce adequate power to function for twenty-four (24) hours through such artificial lighting sources. Thus, discussion in present section equitably states that implanted subdermal PV cell could effortlessly function an modern cardiac pacemaker for upto twenty five years as the lifespan of organic PV device is past twenty five years, this long-lasting lifespan of organic PV device correspondingly validates usage of sub-dermal implanted organic PV cells are finest technique to operate & power modern cardiac pacemaker when compared with other approaches of supplying power to cardiac pacemakers.

Conclusion

The determination of preset paper is to find & represent possible circumstances for PVV-PCBM sub-dermal PV cell for daylight harvesting of the solar energy, attained via using optical properties of the sikn available in works & bearing in mind the skin thickness chosen as 2 millimeters. In present paper PVV-PCBM based organic PV cell is proficient to provide an power output of 2.1 milliwatts (mW) at the depth of 2 millimeters. Furthermore, method utilized in present paper allows to guess sub-dermal PV cell's power yield thus making them an appropriate choice for bioelectronic implant such as pacemaker. Results gained and

Accepted Manuscript

conveyed in present article carry substantial statistics and data for developing & realizing a real-time real-world sub-dermal implantable photovoltaic cell which can be placed underneath skin & can sufficiently source workable power to bio-medical implant (pacemaker). The photovoltaic cell presented in present article permits for appropriate harvesting of energy to electrically charge an energy storing accumulator to power an medical pacemaker for a whole duration of twenty-four (24) hours by simply charging it in six (6) minutes when exposed under daylight. The learning in this research paper offers promising and favorable results, further authenticates applicability of sub-dermal PV device for powering cardiac pacemaker. The sub-dermally implanted organic PV cell could increase the lifespan of cardiac pacemaker therefore, dropping the number of complex surgeries essential for replacement of the implants, thereby reducing the size of bio-medical implant and increasing comfort level of patients. Additionally, such organic photovoltaic cell can be used to power up other bio-medical similar types of implants.

References

1. M.V. Tholl, H.G. Akarçay, H. Tanner, T. Niederhauser, A. Zurbuchen, M. Frenz, A. Haeberlin, "Subdermal solar energy harvesting – A new way to power autonomous electric implants", Appl Energy 269, 114948 (2020).
2. D. Katz, T. Akiyama, "Pacemaker longevity: The world's longest lasting VVI pacemaker", Ann Noninvasive Electrocardiol 12, 223 (2007).
3. H. G. Mond, A. Proclemer, "The 11th world survey of cardiac pacing and implant cardioverter-defibrillators: Calendar year 2009-a world society of arrhythmia's project" Pacing Clin Electrophysiol 34, 1013 (2011).
4. S. Arora et al., "Role of donor–acceptor domain formation and interface states in initial degradation of P3HT:PCBM-based solar cells", Phys. Scr. **83** 035804 (2011).
5. D. Qin et al., "Electron injection via doped p–n heterojunction for inverted organic light emitting diodes", Phys. Scr. **89** 015802 (2014).
6. H. Cheng and N. Yi "Dissolvable tattoo sensors: from science fiction to a viable technology" Phys. Scr. **92** 013001 (2017).
7. Peters, C. H. et al. High efficiency polymer solar cells with long operating lifetimes. Adv. Energy Mater. **1**, 491–494 (2011).
8. D. Baran et al. Reducing the efficiency-stability-cost gap of organic photovoltaics with highly efficient and stable small molecule acceptor ternary solar cells. Nat. Mater. **16**, 363 369 (2017).
9. N. Gasparini et al. Burn-in free nonfullerene-based organic solar cells. Adv. Energy Mater. **7**, 1700770 (2017).
10. Q. Burlingame et al. Reliability of small molecule organic photovoltaics with electron-filtering compound buffer layers. Adv. Energy Mater. **6**, 1601094 (2016).
11. S. B. Sapkota, A. Spies, B. Zimmermann, I. Dürr and U. Würfel, "Promising long-term stability of encapsulated ITO-free bulk-heterojunction organic solar cells under different aging conditions." Sol. Energy Mater. Sol. Cells **130**, 144–150 (2014).
12. P. E. Burrows & S. R. Forrest, "Operating lifetime of phosphorescent organic light emitting devices." Appl. Phys. Lett. **76**, 2493 (2000).
13. Q. Burlingame, X. Huang, X. Liu et al. "Intrinsically stable organic solar cells under high-intensity illumination.", Nature **573**, 394–397 (2019).
14. X. Xu, J. Xiao, G. Zhang, L. Wei, X. Jiao, H-L. Yip, Y. Cao, "Interface-enhanced organic solar cells with extrapolated T 80 lifetimes of over 20 years", Science Bulletin, 65, 208-216 (2019).
15. X. Du et al., "Efficient Polymer Solar Cells Based on Non-fullerene Acceptors with Potential Device Lifetime Approaching 10 Years", Joule, 3, 215-226 (2018).
16. F. C. Krebs, J. E. Carlé, N. Cruys-Bagger, M. Andersen, M. R. Lilliedal, M. A. Hammond, S. Hvidt,, "Lifetimes of organic photovoltaics: photochemistry, atmosphere effects and barrier layers in ITO-MEHPPV:PCBM-aluminium devices", Solar Energy Materials and Solar Cells, 86, 499-516 (2005).
17. S. Schuller, P. Schilinsky, J. Hauch et al. "Determination of the degradation constant of bulk heterojunction solar cells by accelerated lifetime measurements." Appl. Phys. A **79**, 37–40 (2004).
18. K. Birmingham, V. Gardinaru, P. Anikeeva, W.M. Grill, V. Pikov, B. McLaughlin, P. Pasricha, D. Weber, K.A. Ludwig, and K. Famm, 'Bioelectronic medicines: a research roadmap." Nat. Rev. Drug Discov. 13, 399–400 (2014).
19. D.C. Martin, "Molecular design, synthesis, and characterization of conjugated polymers for interfacing electronic biomedical devices with living tissue." MRS Communications, **5**, 131–152 (2015).
20. S. C. Chang, Y. J. Hsiao, T. S. Li, "Improving the Power Conversion Efficiency of Organic Solar Cell by Blending with CdSe/ZnS Core–Shell Quantum Dots.", J. Electron. Mater. 43, 3077 (2014).
21. M. Scholdt, H. Do, J. Lang, A. Gall, A. Colsmann, U. Lemmer, J. D. Koenig, M. Winkler, H. Boettner, "Organic Semiconductors for Thermoelectric Applications." J. Electron. Mater. 39, 1589 (2010).
22. D. C. Jordan, S. R. Kurtz, "Photovoltaic degradation rates - an analytical review" Prog Photovoltaics Res Appl 21, 12 (2011).

Accepted Manuscript

23. K.T. Settaluri, H. Lo, R.J. Ram, "Thin Thermoelectric Generator System for Body Energy Harvesting.", J. Electron. Mater. 41, 984 (2012).
24. M. Thielen, G. Kara, I. Unkovic, et al. "Thermal Harvesting Potential of the Human Body." J. Electron. Mater. 47, 3307 (2018).
25. C. Agnes, M. Holzinger, A. Le Goff, B. Reuillard, K. Elouarzaki, S. Tingry, S. Cosnier, "Supercapacitor/biofuel cell hybrids based on wired enzymes on carbon nanotube matrices: autonomous reloading after high power pulses in neutral buffered glucose solutions", Energy Environ. Sci. 7, 1884 (2014).
26. S. Kim, J. S. Ho, L. Y. Chen, A. S. Poon, "Wireless power transfer to a cardiac implant", Appl. Phys. Lett. 101, 073701 (2012).
27. P. Bai, G. Zhu, Z. H. Lin, Q. Jing, J. Chen, G. Zhang, J. Ma, Z. L. Wang, "Integrated multilayered triboelectric nanogenerator for harvesting biomechanical energy from human motions", ACS Nano 7, 3713 (2013).
28. E. Miyako, C. Hosokawa, M. Kojima, M. Yudasaka, R. Funahashi, I. Oishi, Y. Hagihara, M. Shichiri, M. Takashima, K. Nishio, Y. Yoshida, "A photo-thermal-electrical converter based on carbon nanotubes for bioelectronic applications", Angew Chem Int Ed Engl 50, 12266 (2011).
29. P. Roberts, G. Stanley, J. M. Morgan, "Harvesting the energy of cardiac motion to power a pacemaker" Circulation 118, 679 (2008).
30. International Commission on Non-Ionizing Radiation Protection, "ICNIRP statement on far infrared radiation exposure", Health Phys 91, 630 (2006).
31. P. Niu, P. Chapman, R. Riemer, X. Zhang, "Evaluation of Motions and Actuation Methods for Biomechanical Energy Harvesting" In Proceedings of the IEEE 35th Annual Power Electronics Specialists Conference, Aachen, Germany, 2100, (2004).
32. R. Tashiro, N. Kabei, K. Katayama, Y. Ishizuka, F. Tsuboi, K. Tsuchiya, "Development of an electrostatic generator that harnesses the motion of a living body (Use of a resonant phenomenon)", JSME Int. J. Ser. C 43, 916 (2000).
33. X. Wei, J. Liu, "Power sources and electrical recharging strategies for implantable medical devices" Front. Energy Power Eng. China 2, 1 (2008).
34. A. Irani, M. Bianco, D. Tran, P. D. Deyoung, M. L. R. Wyld, T. H. Li, "Energy Generating Systems for Implanted Medical Devices" U.S. Patent US20090171404 A1, July 2, 2009.
35. K. Goto, T. Nakagawa, O. Nakamura, S. Kawata, "An implantable power supply with an optically rechargeable lithium battery", IEEE Trans. Biomed. Eng. 48, 830 (2001).
36. I. Stark, M. Stordeur, "New Micro Thermoelectric Devices Based on Bismuth Telluride-Type Thin Solid Films." In Proceedings of the Eighteenth International Conference on Thermoelectrics, Piscataway, NJ, USA, (1999), p. 465.
37. N. Mano, F. Mao, A. Heller, "Characteristics of a miniature compartment-less glucose-O2 biofuel cell and its operation in a living plant" J. Am. Chem. Soc. 125, 6588 (2003).
38. E. Chan, B. Sorg, D. Protsenko, M. O'Neil, M. Motamedi, A. Welch "Effects of compression on soft tissue optical properties", IEEE J Sel Top Quantum Electron 2, 943 (1996).
39. A. N. Bashkatov, E. A. Genina, V. I. Kochubey, V. V. Tuchin "Optical properties of human skin, subcutaneous and mucous tissues in the wavelength range from 400 to 2000nm", J Phys D: Appl Phys 38, 2543 (2005).
40. S. L. Jacques, "Optical properties of biological tissues: a review.", Phys Med Biol 58, R37 (2013).
41. D. Barolet, "Light-emitting diodes (LEDs) in dermatology." Semin. Cutan. Med. Surg. 27, 227 (2008).
42. N. Bosschaart, G. J. Edelman, M. C. G. Aalders, T. G. van Leeuwen, D. J. Faber, "A literature review and novel theoretical approach on the optical properties of whole blood." Lasers Med Sci 29, 453 (2013).
43. G. Yoon, A. Welch, M. Motamedi, M. Gemert, "Development and application of three-dimensional light distribution model for laser irradiated tissue." IEEE J Quantum Electron 23, 1721 (1987).
44. R. L. Van Veen, H. Sterenborg, A. Pifferi, A. Torricelli, R. Cubeddu, "Determination of VIS- NIR absorption coefficients of mammalian fat, with time- and spatially resolved diffuse reflectance and transmission spectroscopy." Biomedical topical meeting OSA, 2004.
45. C. R. Simpson, M. Kohl, M. Essenpreis, M. Cope, "Near-infrared optical properties of exvivo human skin and subcutaneous tissues measured using the monte carlo inversion technique." Phys Med Biol 43, 2465 (1998).
46. I. V. Meglinski, S. J. Matcher, "Quantitative assessment of skin layers absorption and skin reflectance spectra simulation in the visible and near-infrared spectral regions.", Physiol Meas 23, 741 (2002).
47. G. J. Tearney, M. E. Brezinski, J. F. Southern, B. E. Bouma, M. R. Hee, J. G. Fujimoto, "Determination of the refractive index of highly scattering human tissue by optical coherence tomography.", Opt Lett 20, 2258 (1995).
48. Y. Lee, K. Hwang,"Skin thickness of korean adults.", Surg Radiol Anat 24, 183 (2002).
49. Bereuter, L., Williner, S., Pianezzi, F. et al. Energy Harvesting by Subcutaneous Solar Cells: A Long-Term Study on Achievable Energy Output. Ann Biomed Eng 45, 1172–1180 (2017).

Accepted Manuscript

50. ATLAS, ATLAS User's Manual Device Simulation Software, Silvaco International Ltd. Santa clara, 2014.
51. P. Mittal, Y. S. Negi, R. K. Singh, "A Depth Analysis for Different Structures of Organic Thin Film Transistors: Modeling of Performance Limiting Issues" Microelectronics Engineering150, 7 (2016).
52. S. Negi, P. Mittal and B. Kumar, "Analytical Modelling and Parameters Extraction of Multilayered OLED", IET Circuits, Devices and Systems 13, 1255 (2019).
53. P. Mittal, Y. S. Negi, R. K. Singh, "An analytical approach for parameter extraction in linear and saturation regions of top and bottom contact organic transistors." Journal of Computational Electronics 14, 828 (2015).
54. S. Negi, P. Mittal and B. Kumar "Numerical Modeling and Parameters Extraction of Novel Triple Hole Block layer based Organic Light Emitting Diode for Display", Journal of the Society for Information Display 22, 956 (2020).
55. S. Negi, P. Mittal, and B. Kumar, "In-Depth analysis of structures, materials, models, parameters and applications of organic light emitting didoes", Jou Of Ele Mat, 49, 4610 (2020).
56. S. Negi, P. Mittal and B. Kumar, "Modeling and Analysis of High-Performance Triple Hole Block Layer Organic LED Based Light Sensor for Detection of Ovarian Cancer," in IEEE Trans. Circuits Syst. I: Reg. Papers 68, 3254 (2021).
57. V. D. Mihailletchi, P. W. M. Blom, J. C. Hummelen, M. T. Rispens, "Cathode dependence of the open circuit voltage of polymer: fullerene bulk heterojunction solar cells", Journal of Applied Physics 94, 6849 (2003).
58. S. A. Prahl, M. J. C. Van Gemert and A. J. Welch, 1993 Determining the optical properties of turbid media by using the adding–doubling method Appl. Opt. **32** 559–68.
59. L. Finlayson, I. R. M. Barnard, L. McMillan, S. H. Ibbotson, C. T. A. Brown, E. Eadie, K. Wood, "Depth Penetration of Light into Skin as a Function of Wavelength from 200 to 1000 nm." Photochem. Photobiol. **2021**.
60. Kono, T., Yamada, J. In Vivo Measurement of Optical Properties of Human Skin for 450–800 nm and 950–1600 nm Wavelengths. Int J Thermophys **40**, 51 (2019).
61. Alireza Firooz, Bardia Sadr, Shahab Babakoohi, Maryam Sarraf-Yazdy, Ferial Fanian, Ali Kazerouni-Timsar, Mansour Nassiri-Kashani, Mohammad Mehdi Naghizadeh, Yahya Dowlati, "Variation of Biophysical Parameters of the Skin with Age, Gender, and Body Region", The Scientific World Journal, vol. 2012, Article ID 386936, 5 pages, 2012.
62. A. E. Mayes, P. G. Murray, D. A. Gunn et al., "Ageing appearance in China: biophysical profile of facial skin and its relationship to perceived age," Journal of the European Academy of Dermatology and Venereology, vol. 24, no. 3, pp. 341–348, 2010.
63. S.H. Tseng, A. Grant, A. J. Durkin, "In vivo determination of skin near-infrared optical properties using diffuse optical spectroscopy." J Biomed Opt 13, 014016 (2008).
64. J. P. Ritz, A. Roggan, C. Isbert, G. Muller, H. Bhur, C. T. Germer "Optical properties of native and coagulated porcine liver tissues between 400 and 2400 nm", Lasers Surg. Med. 29, 205 (2001).
65. H. R. Philip, E. A. Taft, "Optical constants of silicon in the region 1 to 10 eV," Phys. Rev. 120, 37 (1960).
66. T. Igarashi, K. Nishino, S. K. Nayar, "The appearance of human skin: A survey." Found. Trends Comput. Graph. Vis. 3, 1 (2007).
67. C. Dagdeviren, B. D. Yang, Y. Su, P. L. Tran, P. Joe, E. Anderson, J. Xia, V. Doraiswamy, B. Dehdashti, X. Feng, B. Lu, R. Poston, Z. Khalpey, R. Ghaffari, Y. Huang, M. J. Slepian, J. A. Rogers, "Conformal piezoelectric energy harvesting and storage from motions of the heart, lung, and diaphragm.", Proc. Natl. Acad. Sci. 111, 1927 (2014).
68. A. Zebda, S. Cosnier, J. P. Alcaraz, M. Holzinger, A. Le Goff, C. Gondran, F. Boucher, F. Giroud, K. Gorgy, H. Lamraoui, P. Cinquin, "Single glucose biofuel cells implanted in rats power electronic devices.", Sci. Rep. 3, 1 (2013).
69. G. Grandi, A. Ienina, M. Bardhi, "Effective low-cost hybrid LED-halogen solar simulator.", IEEE Trans Ind Appl 50, 3055 (2014).

Accepted Manuscript

Fourier Analysis Based Respiration Rate Estimation Using Corrupted Photoplethysmogram Signal

2022 5th International Conference on Multimedia, Signal Processing and Communication Technologies (IMPACT) | 978-1-6654-7647-8/22/$31.00 ©2022 IEEE | DOI: 10.1109/IMPACT55510.2022.10029065

Pankaj
Department of Electronics and Communication Engineering
Bennett University
Uttar Pradesh, India
er.pankaj08@gmail.com

Ashish Kumar
School of Electronics Engineering
Vellore Institute of Technology
Chennai, India
akumar.1june@gmail.com

Rama Komaragiri
Department of Electronics and Communication Engineering
Bennett University
Uttar Pradesh, India
rama.komaragiri@bennett.edu.in

Manjeet Kumar
Department of Electronics and Communication Engineering,
Delhi Technological University (DTU)
Rohini, Delhi.
manjeetchhillar@gmail.com

Abstract **-** This work proposes a Fourier series-based Fourier decomposition method (FDM) technique to analyze the photoplethysmography (PPG) signals to estimate respiration rate (RR). The proposed method consists of data acquisition, preprocessing, signal decomposition into Fourier intrinsic band function (FIBFs), Superposition of clean FIBFs, fast Fourier transform, spectral peak identification, and estimation of RR. The efficacy of the proposed FDM-based RR estimation method is evaluated using the CapnoBase database, a benchmark RR estimation database. The proposed method improves estimation accuracy with a mean absolute error (MAE) value of 1.02 breaths per minute (BPM) and a root mean square error value (RMSE) of 1.22. The proposed method can accurately estimate vital physiological parameter RR from the PPG signal.

Index Terms - Photoplethysmography (PPG), Fourier decomposition method (FDM), Respiratory signal, Fast Fourier transform (FFT), Respiration rate (RR).

I. INTRODUCTION

Continuous monitoring of respiratory rate (RR) provides early indications of forthcoming illnesses and saves lives. Respiratory rate is the number of breaths a person takes per minute (breaths per minute). Deviations in normal RR indicate a subject's health status, such as pneumonia and COVID-19. Currently, electrocardiogram (ECG) and Photoplethysmogram (PPG) sensors [1] are used to estimate RR. Therefore, accurate and reliable measurement of RR using a wearable PPG sensor provides vital information to healthcare professionals. Photoplethysmogram (PPG) is an optical technique used to measure blood volume variation in tissue[2]. The widespread wearable health monitoring devices don't have a feature to monitor RR, which can be used for prognosis and diagnosis [3]. Capnography is a gold-standard technique used to estimate the RR by estimating the fluctuations in the concentration of carbon dioxide with the respiratory signal. Capnography is used mainly during intensive care and anesthesia, as the process is bulky and cumbersome.

Previous research has shown that abnormal RR detection is much more critical than measuring blood pressure or heart rate (HR) changes. Therefore real-time, continuous monitoring of RR is more likely to be a significant predictor of analyzing the patient at risk. During the COVID-19 pandemic, PPG sensor-based pulse oximetry became a household technique to monitor SpO2 and HR [4].

The signal measured by a PPG device consists of the RR information, as the optical intensity measured by a PPG signal depends on the oxygen levels in the tissue [5]. During a respiratory cycle, changes in intra-thoracic pressure triggers a respiratory-induced intensity variation (RIIV) component in the PPG signal. Variation in the PPG signal due to respiration differs for subjects with different health statuses. Thus, hospitals can use a PPG sensor-based RR estimation technique for real-time continuous RR monitoring and spot-checking.

To estimate RR from a PPG signal, several methods are proposed in the literature using signal processing techniques. The algorithm proposed in the literature for RR estimation from PPG sensor signal have certain limitations. Some of these techniques are independent component analysis [6], short-time Fourier transform[7], and continuous wavelet transform [8]. The literature on RR estimation reported by different researchers is presented in [9], [10]. Initially, digital signal processing filter-are used to filter out the noise from the signal to estimate RR. The performance of the digital filter depends on the selection of their cutoff frequency. Moreover, the presence of motion artifacts within the desired estimation range makes these approaches unreliable. In [11] autoregressive model time-frequency (TF) domain-based approach is proposed to estimate RR.

In [12], the authors presented the RR estimation technique using empirical mode decomposition (EMD). But the mode mixing problem and a manual selection of intrinsic mode function make this technique less familiar for RR estimation. In [13], the authors proposed a method by combining ensemble EMD and principal component analysis (PCA) techniques to estimate RR. However, the proposed works have significantly contributed to PPG-based RR estimation but faced mode mixing problems and high computational complexity. Thus in [14], the Fourier–Bessel series expansion-based empirical wavelet transform is proposed to estimate RR.

Further in [15], a smart fusion method based on EEMD is proposed to estimate RR. Recently a machine learning (ML) approach to estimate RR from motion artifacts affected PPG signal has been proposed [16]. The raw PPG signal is decomposed for motion artifact suppression using the

978-1-6654-7647-8/22/$31.00 ©2022 IEEE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on February 06,2023 at 05:14:55 UTC from IEEE Xplore. Restrictions apply.

variational mode decomposition technique. Then the features are extracted to train the ML model for RR estimation.

In [17], variable-frequency complex demodulation (VFCDM) approach is proposed to estimate RR. However, VFCDM provides a better TF resolution than STFT and CWT. Still, the need for a minimum of sixty-second data and high computational complexity makes this approach unworkable in real-time RR monitoring wearable devices.

Thus, a computationally efficient approach for real-time RR estimation based on time-frequency analysis of PPG signal is still in great demand.

This paper proposes a computationally efficient Fourier decomposition method (FDM) to estimate accurate and reliable RR from a PPG signal. In literature, FDM has demonstrated reliable results while analyzing non-stationary and non-linear signals [18]. Moreover, compared to EMD and its variants, FDM provides a clear time-frequency representation while analyzing the non-stationary signals. Hence, FDM can potentially decompose PPG signals into a set of Fourier intrinsic band functions (FIBFs) with desired cutoff frequency [19]. The superposition of clean FIBFs is the only step required in the proposed method to estimate RR from a PPG signal. Hence, FDM is a technique with reduced computational complexity to estimate the RR.

The rest of the paper is organized as follows. The material and the methodology used to estimate RR is briefly describe in Section II. Section III provides Results and discussions, and Section IV concludes the work.

II. MATERIALS AND METHODS

A. Datasets

The CapnoBase [20] is a publicly available PPG RR benchmark dataset comprising concurrently recorded PPG and the reference respiratory signals. To evalute the performance of the proposed methodology signal from the CapnoBase dataset are used. The CapnoBase dataset is collected from 59 pediatrics aged between 1 to 17 years with an average age of 9 years and 35 adult subjects undergoing routine anesthesia between 26 to 76 years with an average age of 52 years. The sampling frequency of the capnography reference respiratory signal and the PPG signal is 300 Hz. Out of the randomly selected 42 subjects, 29 are pediatric, and 13 are adult subjects. Each signal in the Capnobase dataset is recorded for an eight-minute duration. A research expert manually labeled the simultaneously recorded gold standard reference Capnogram signal waveform from each breath cycle to calculate the reference RR value.

B. Proposed Approaches

The block diagram of the proposed FDM-based RR estimation method is shown in Fig. 1. As shown in fig. 1, the proposed approach is segmented into four main stages. Stage-1 preprocess the acquired PPG signal. Stage-2 decomposes the PPG signal into Fourier intrinsic band functions (FIBFs). The superposition of clean FIBFs and estimation of RR are performed in stage 3. Stage 4, is used to evaluate the performance of the proposed framework. The following subsections details each stage.

B.1 Preprocessing stage

The acquired PPG signal from the CapnoBase dataset is partitioned using an overlapping windowing scheme for a thirty-two-second window. The number of samples (N) in each window is 9600.

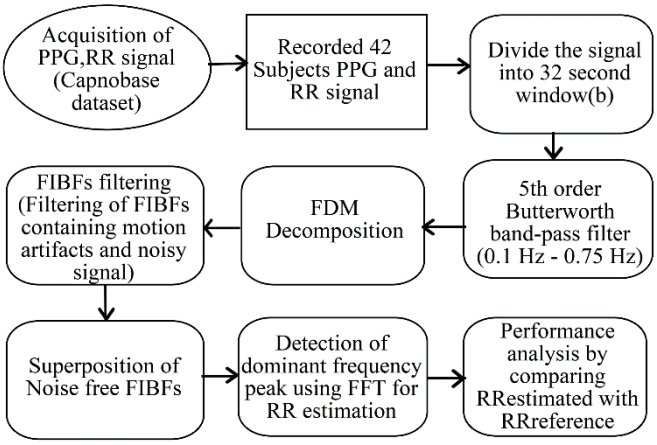


Fig. 1 Block representation of the proposed method

The first window (b = 1) consists of samples numbered from 1 to 9600, as shown in Fig. 2, which are considered to estimate the RR. After the first window, each consecutive window overlaps the immediate prior window by (N-w) samples. Here w is the step size of length three-second. Thus, 8700 samples from the first window are coupled with 900 new samples to obtain 9600 new samples in the next window. Therefore, RR estimation includes new data for three seconds in any window.

Fig. 2 shows the sample number assignment. This work considers N = 9600 and w = 900 to match the reference RR available in the CapnoBase dataset.

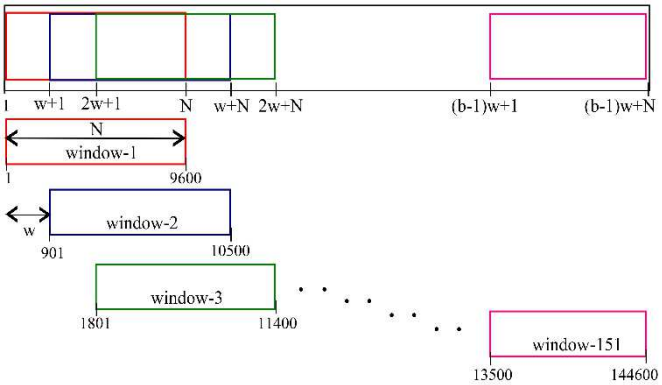


Fig. 2. Pictorial representation of consecutive window's sample number assignment

In the next step, the N second segmented window is passed through a 5th-order Butterworth band-pass filter [0.05-0.75 Hz] to filter out undesired components lies outside the desired range.

B.2 Decomposition and Denoising stage

The FDM technique is an adaptive signal decomposition method that maps the Fourier series basis function into a number of FIBFs, which are band-limited and finite. The PPG signal $P[n]$ decomposes into an R number of FIBFs using eq. (1). In eq. (1), C_0 is the average of signal $P[n]$. $S_1[n]$, $S_2[n]$, $\cdots$ $S_R[n]$ are amplitude and frequency modulated FIBFs, which are orthogonal or linearly independent non-orthogonal yet energy preserving (LINOEP) and adaptive.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on February 06,2023 at 05:14:55 UTC from IEEE Xplore. Restrictions apply.

$$P[n] = C_0 + \sum_{l=1}^{R} S_l[n] \quad (1)$$

The preprocessed PPG signal decomposes into frequency bands to eliminate motion artifacts and cardiac information from the PPG signal, as shown in Fig. 3.

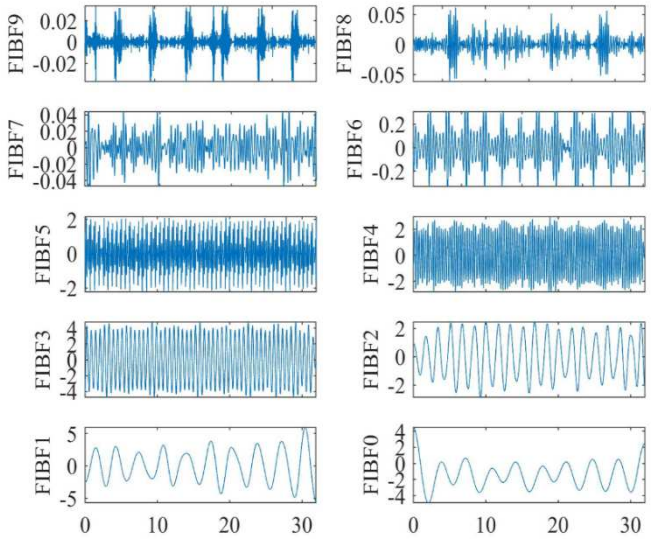


Fig. 3 Decomposition of PPG signal into FIBFs

The frequency bands corresponding to motion artifacts and cardiac components are removed. The rest of the frequency band is used to retrieve the clean PPG signal. In this proposed work, FDM decompose the signal *P[n]* into ten FIBFs (FIBF0–FIBF9), as shown in Fig. 3.

B.3 RR Estimation stage

The cutoff frequency corresponding to desired RR signal of 3 to 45 BPM (0.05 to 0.75 Hz) is selected to estimate the RR. By superposition of rest FIBF signal components, a clean PPG signal (*FIBFclean*) is obtained at the output, as shown in Fig. 4. For further processing of RR estimation, *FIBFclean* is used. Since the *FIBFclean* preserved the respiratory activity, a fast Fourier transform is employed on the *FIBFclean* to estimate the respiratory frequency (f_{RR}). *U*sing eq. 2, the RR is estimated ($RR_{estimated}$) in BPM.

$$RR_{estimated} = f_{RR} \times 60 \quad (2)$$

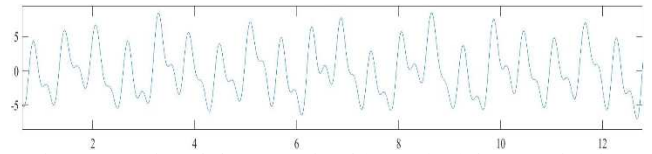


Fig. 4. Extraction of clean PPG signal (*FIBFclean*) for RR estimation

B.4 Performance Analysis stage

The mean absolute error (MAE) and root mean square error (RMSE) of the $RR_{estimated}$ are computed to evaluate the performance of the proposed FDM-based RR estimation method. For each signal in the Capnobase dataset, eq. (3) and eq. (4) are used to compute MAE and RMSE, respectively.

$$MAE = \frac{1}{W}\sum_{b=1}^{W}\left|RR_{ref}(b) - RR_{estimated}(b)\right| \quad (3)$$

$$RMSE_{RR} = \sqrt{\frac{1}{W}\sum_{b=1}^{W}\left(RR_{ref}(b) - RR_{estimated}(b)\right)^2} \quad (4)$$

Here, $RR_{ref}(b)$ and $RR_{estimated}(b)$ are the reference RR and PPG-derived RR for the b^{th} observation, and W is the total number of windows.

Table I and Table II compare MAE and RMSE with various methods found in the literature. The Bland-Altman plot and Pearson correlation coefficient are used to assess the performance of the FDM-based RR estimation method from the PPG signal. Pearson's correlation plot demonstrates the linear relation between reference RR and estimated RR.

The (RR_{ref} - $RR_{estimated}$) versus (RR_{ref} + $RR_{estimated}$)/2 is plotted in the Bland-Altman plot. The bias is estimated by computing the mean value of (RR_{ref} - $RR_{estimated}$). The 95% LOAs must fall within bias ± 1.95 standard deviations (SD).

III. RESULTS AND DISCUSSION

Overall, data from 4563 windows are analyzed from the CapnoBase PPG data to estimate RR. The MAE and RMSE values of the proposed work are 1.02 BPM and 1.22 BPM, respectively. Both performance matrices indicate that the proposed work outperforms the previously reported methods in Table I and Table II.

Table I Comparison of RMSE of the proposed method with existing methods

Work	RMSE	Work	RMSE
Zhang 2016 [4]	3.25	Motion [13]	2.77
Garde 2014 [9]	3.18	Zhang 2020 [21]	1.4
Zhang 2017 [10]	2.93	Karlen [22]	1.56
Zhang 2017 [10]	3	Dehkordi [23]	1.8
Garde 2013 [12]	3.5	**This Work**	**1.22**

Table II Comparison of MAE of the proposed method with existing methods

Work	MAE (BPM)	Work	MAE (BPM)
Zhang [10]	0.82	Zhu2018 [24]	1.89
Zhang [10]	0.83	Motion [25]	2.4
Motion [13]	3.5	Khreis [26]	2.1
Karlen [22]	1.2	Pimentel [27]	1.5
Zhu 2019 [28]	1.46	Zhu 2021 [29]	1.18
Birrenkott [30]	1.23	This Work	1.02

The methodology used to estimate RR proposed in the literature has inherent limitations. The ICA-based approach only provides accurate results if the noise and PPG signal are statistically independent, which is impossible if the estimation is done in real-time. Some of the work used a wavelet-based approach to remove motion artifacts. However, the approach provides reliable results, however, the need to empirically set the values of wavelet threshold, limits its usage. Another methodology used in the literature used the EMD technique,

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on February 06,2023 at 05:14:55 UTC from IEEE Xplore. Restrictions apply.

similar to wavelets. The EMD is an efficient feature extractor approach. However empirical and heuristic procedure based EMD algorithm makes the mathematical modelling of EMD and its variant more complex.

Moreover, EMD also suffers from aliasing, detrend uncertainty and mode mixing. To overcome these issues the FDM approach is used. FDM provides enhance TF estimation and can accurately filter out motion artifacts and noise components by separating the signal into FIBFs. The PPG signal decomposes into I orthogonal FIBFs with desired frequency band, which is not possible in other TF-based approaches. Table 1 and Table 2 shows the comparative analysis of result obtained with the FDM approach with the state-of-the-art approach present in the literature to analyze the potential of the FDM technique. The results show that the proposed RR estimation framework perform better than the state-of-art methods.

The correlation between RR_{ref} and $RR_{estimated}$ for all the subjects in the Capnobase dataset is plotted in the Bland-Altman plot as shown in Fig. 5.

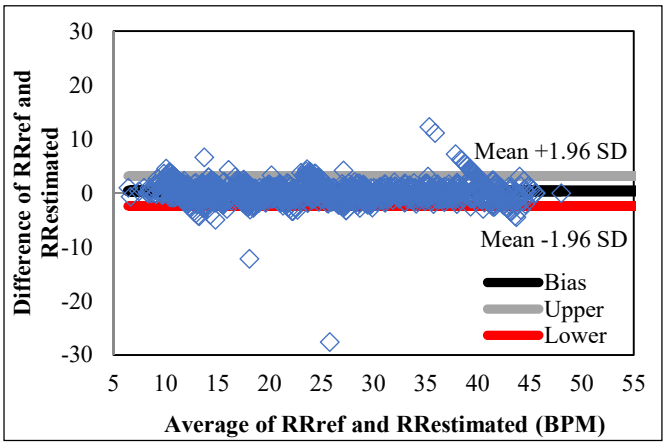


Fig. 5 Bland-Altman plot representation of the proposed method

The LOA value for the signals from the Capnobase dataset is between – 2.4 to 3.10 BPM with a bias of 0.348 BPM and σE 1.40 BPM.

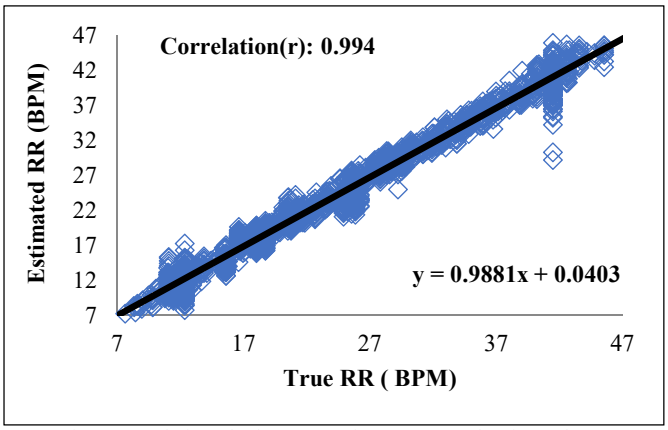


Fig. 6 Correlation plot between reference RR and estimated RR

As observed from Fig. 5, maximum points lie within the agreement limit of 95%, indicating that the error between RR_{ref} and $RR_{estimated}$ is low and within an acceptable limit.

The Pearson correlation plot is shown in fig. 6. A correlation coefficient value (r) of 0.994 indicates a high correlation between RR_{ref} and $RR_{estimated}$.

Fig. 7 shows the tracking between $RR_{estimated}$ and RR_{ref} at each window for the Subject-009 of the Capnobase dataset. Fig. 7 shows the potential of the proposed method in estimating the RR in real time.

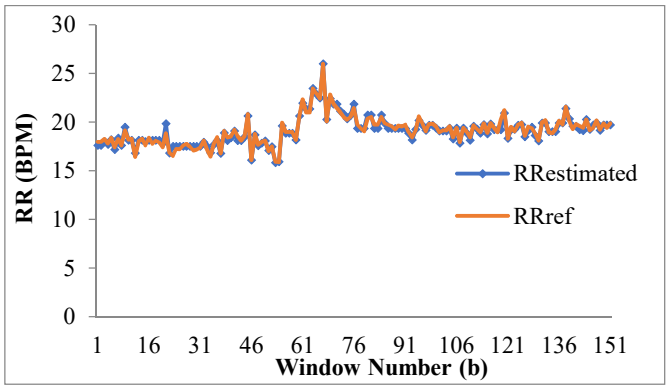


Fig. 7 Tracking of estimated RR with reference RR of Subject-009

The work is further extended to address the constraints of a real-time monitoring wearable device. Energy consumption and computational complexity are the constraints in a wearable device as the system has limited resources. The sampling rate is a parameter used to improve the system's performance to reduce energy consumption. Thus, the proposed work also tested at a lower sampling rate to reduce energy consumption for the proposed methodology. A lower sampling rate helps preserve battery life but may affect the system's accuracy. The proposed methodology is also tested on a sampling rate of 125 Hz instead of 300 Hz for randomly selected two subjects (Subject 0016 and Subjects 0029). The results obtained at 125 Hz are similar to those obtained at 300 Hz. Table 3 shows the MAE at two different sampling rates. Thus, lower energy consumption at a lower sampling rate and computation time makes the proposed FDM-based framework be used in real-time monitoring wearable systems.

Table 3: MAE values of the signals for subjects 0016 and 0029 at different sampling frequencies

Subject Id	MAE @ sampling frequency	
	300 Hz	125 Hz
Subject 0016	0.825	0.829
Subject 0029	0.746	0.747

CONCLUSIONS

Ensuring quality of the respiratory signal obtained from a PPG sensor signal in real-time is one of the big challenges in estimating RR. This work introduces Fourier analysis-based FDM method to denoise and estimates vital physiological parameter RR from the PPG signal. Using time-frequency domain combined methods and applying FIBFs enabled the estimation of RR accurately and reliably. The FDM-based RR estimation method is validated using signals from the publicly available benchmark dataset Capnobase. The RMSE and MAE values estimated using the proposed method are 1.22 and 1.02, respectively. In the future, we aim

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on February 06,2023 at 05:14:55 UTC from IEEE Xplore. Restrictions apply.

to extract other respiratory-induced variability and classify the instantaneous RR using a deep learning algorithm.

REFERENCES

[1] S. C. Gandevia and D. K. McKenzie, "Respiratory rate: The neglected vital sign," Medical Journal of Australia, vol. 189, no. 9, p. 532, 2008.

[2] Pankaj, A. Kumar, R. Komaragiri, and M. Kumar, "A Review on Computation Methods Used in Photoplethysmography Signal Analysis for Heart Rate Estimation," Archives of Computational Methods in Engineering, 2021.

[3] V. Hartmann, H. Liu, F. Chen, W. Hong, S. Hughes, and D. Zheng, "Toward accurate extraction of respiratory frequency from the photoplethysmogram: Effect of measurement site," Front Physiol, vol. 10, no. JUN, 2019.

[4] X. Zhang and Q. Ding, "Respiratory rate monitoring from the photoplethysmogram via sparse signal reconstruction," Physiol Meas, vol. 1105, p. 1105.

[5] Pankaj, A. Kumar, M. Kumar, and R. Komaragiri, "STSR: Spectro-Temporal Super-Resolution Analysis of a Reference Signal Less Photoplethysmogram for Heart Rate Estimation During Physical Activity," IEEE Trans Instrum Meas, vol. 71, pp. 1–10, 2022.

[6] Y. Zhou, Y. Zheng, C. Wang, and J. Yuan, "Extraction of Respiratory Activity from Photoplethysmographic Signals Based on an Independent Component Analysis Technique: Preliminary Report," Instrum Sci Technol, vol. 34, no. 5, pp. 537–545, Oct. 2006.

[7] K. H. Shelley, A. A. Awad, R. G. Stout, and D. G. Silverman, "The use of joint time frequency analysis to quantify the effect of ventilation on the pulse oximeter waveform," J Clin Monit Comput, vol. 20, no. 2, pp. 81–87, 2006.

[8] P. S. Addison, J. N. Watson, M. L. Mestek, and R. S. Mecca, "Developing an algorithm for pulse oximetry derived respiratory rate (RR oxi): a healthy volunteer study," pp. 45–51, 2012.

[9] A. Garde, W. Karlen, J. M. Ansermino, and G. A. Dumont, "Estimating respiratory and heart rates from the correntropy spectral density of the photoplethysmogram," PLoS One, vol. 9, no. 1, 2014.

[10] X. Zhang and Q. Ding, "Biomedical Signal Processing and Control Respiratory rate estimation from the photoplethysmogram via joint sparse signal reconstruction and spectra fusion," Biomed Signal Process Control, vol. 35, pp. 1–7, 2017.

[11] J. Lee and K. H. Chon, "An autoregressive model-based particle filtering algorithms for extraction of respiratory rates as high as 90 breaths per minute from pulse oximeter," IEEE Trans Biomed Eng, vol. 57, no. 9, pp. 2158–2167, Sep. 2010.

[12] A. Garde, W. Karlen, P. Dehkordi, J. M. Ansermino, and G. A. Dumont, "Empirical mode decomposition for respiratory and heart rate estimation from the photoplethysmogram," Comput Cardiol (2010), vol. 40, pp. 799–802, 2013.

[13] M. A. Motin, C. K. Karmakar, and M. Palaniswami, "Ensemble empirical mode decomposition with principal component analysis: a novel approach for extracting respiratory rate and heart rate from photoplethysmographic signal," IEEE J Biomed Health Inform, vol. 22, no. 3, pp. 766–774, 2018.

[14] R. Katiyar, V. Gupta, and R. B. Pachori, "Fbse-ewt-based approach for the determination of respiratory rate from ppg signals," IEEE Sens Lett, vol. 3, no. 7, Jul. 2019.

[15] IEEE Engineering in Medicine and Biology Society. Annual International Conference (42nd : 2020 : Online), IEEE Engineering in Medicine and Biology Society, Institute of Electrical and Electronics Engineers, and Canadian Medical and Biological Engineering Society. : 20-24 July 2020, Montreal, Canada.

[16] M. N. I. Shuzan et al., "A Novel Non-Invasive Estimation of Respiration Rate from Motion Corrupted Photoplethysmograph Signal Using Machine Learning Model," IEEE Access, vol. 9, pp. 96775–96790, 2021.

[17] K. H. Chon, S. Dash, and K. Ju, "Estimation of Respiratory Rate From Photoplethysmogram Data Using Time-Frequency Spectral Estimation," IEEE Trans Biomed Eng, vol. 56, no. 8, pp. 2054–2063, 2009.

[18] P. Singh, S. D. Joshi, R. K. Patney, and K. Saha, "The Fourier decomposition method for non-linear and non-stationary time series analysis," Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, vol. 473, no. 2199, 2017.

[19] Pankaj, A. Kumar, R. Komaragiri, and M. Kumar, "Reference signal less Fourier analysis based motion artifact removal algorithm for wearable photoplethysmography devices to estimate heart rate during physical exercises," Comput Biol Med, vol. 141, no. May 2021, p. 105081, 2022.

[20] W. Karlen, "CapnoBase IEEE TBME Respiratory Rate Benchmark." Scholars Portal Dataverse, 2021. doi: doi/10.5683/SP2/NLB8IT.

[21] J. Zhang, G. Scebba, W. Karlen, and S. Member, "Photoplethysmogram Derived Respiratory Rate," pp. 5939–5942, 2020.

[22] K. W., R. S., A. JM, and D. GA, "Multiparameter respiratory rate estimation from the photoplethysmogram," IEEE Trans Biomed Eng, vol. 60, no. 7, pp. 1946–1953, 2013.

[23] P. Dehkordi, A. Garde, B. Molavi, J. M. Ansermino, and G. A. Dumont, "Extracting instantaneous respiratory rate from multiple photoplethysmogram respiratory-induced variations," Front Physiol, vol. 9, no. JUL, pp. 1–10, 2018.

[24] T. Zhu, A. E. W. Johnson, Y. Yang, G. D. Clifford, and D. A. Clifton, "Bayesian fusion of physiological measurements using a signal quality extension Bayesian fusion of physiological measurements using a signal quality extension," 2018.

[25] M. A. Motin, C. K. Karmakar, and M. Palaniswami, "Selection of empirical mode decomposition techniques for extracting breathing rate from PPG," IEEE Signal Process Lett, vol. 26, no. 4, pp. 592–596, 2019.

[26] S. Khreis, D. Ge, H. A. Rahman, and G. Carrault, "Breathing Rate Estimation Using Kalman Smoother with Electrocardiogram and Photoplethysmogram," vol. 9294, no. c, pp. 1–12, 2019.

[27] M. A. F. Pimentel et al., "Toward a robust estimation of respiratory rate from pulse oximeters," IEEE Trans Biomed Eng, vol. 64, no. 8, pp. 1914–1923, 2017.

[28] T. Zhu et al., "Patient-Specific Physiological Monitoring and Prediction Using Structured Gaussian Processes," IEEE Access, vol. 7, pp. 58094–58103, 2019.

[29] T. Zhu and D. A. Clifton, "Comparison of parametric and non-parametric Bayesian inference for fusing sensory estimates in physiological time-series analysis," no. August 2019.

[30] D. A. Birrenkott, M. A. F. Pimentel, P. J. Watkinson, and D. A. Clifton, "A Robust Fusion Model for Estimating Respiratory Rate from Photoplethysmography and Electrocardiography," no. c, pp. 1–9, 2017.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on February 06,2023 at 05:14:55 UTC from IEEE Xplore. Restrictions apply.

Check for updates

Original Research

SAGE Open
January-March 2023: 1–17
© The Author(s) 2023
DOI: 10.1177/21582440221135107
journals.sagepub.com/home/sgo
SAGE

Impact of Competition in Credit Rating Industry: Evidence From India

Chandan Sharma[1], Archana Singh[1], and Rajan Yadav[1]

Abstract

Credit rating agencies (CRAs) play an essential role in efficiently allocating capital in the financial system. However, several researchers have highlighted shortcomings of CRAs, leading to drastic consequences such as the 2008 financial crisis. The paper focuses on the competition among CRAs as one of the critical drivers of issues plaguing the credit rating industry. The paper uses quantitative and regression techniques to check the impact of competition on a firm's credit rating. The paper utilizes firm-level financial and credit rating data from India. The paper uses dual ratings to check whether competition leads to rating inflation and shopping practices in the credit rating industry. The paper finds that CRA inflates a firm's credit rating due to competition from other CRAs. Rating shopping is also evident in the credit rating industry, driven by competition between CRAs to gain new clients. The paper's findings also indicate that increased competition for large-size firms business leads to CRAs showing leniency when rating such firms. The results raise the need for regulators to actively monitor and control the competition among CRAs to ensure the accuracy of credit ratings.

Keywords

credit rating, competition, rating inflation, rating shopping

Introduction

Competition among firms operating in an industry is natural in the modern-day economy, where government intervention or control is minimal in businesses. Competition among firms has several benefits, such as higher quality products or services at lower prices for end customers and improved business efficiency (EU Commission, 2021; FTC, 2021). However, in the absence of relevant checks in critical industries such as financial systems, unwanted effects of competition can have a contagion effect on the economy. Adverse consequences of the financial industry competition, including the credit rating industry, were seen in the 2008 financial crisis in the US, which led to a worldwide economic crisis (Chu & Rysman, 2019; Griffin et al., 2013; Kim & Park, 2016). The paper focuses on the competition among Credit Rating Agencies (CRAs), which play a vital role in estimating risk in today's financial system.

Credit rating by CRAs is mandatory for a corporate to borrow money through capital markets and even from a bank in several countries (Duff & Einig, 2015). Credit rating determines the borrowing rate of corporate (Tang, 2009) and lenders' capital requirements in certain countries (Reserve Bank of India, 2015). Thus, credit rating plays an essential role in the fair pricing of risk in the financial system. However, large-scale corporate default and financial crises have raised questions about CRAs' credit rating accuracy and a possible upward bias in credit rating (Bonsall et al., 2015; Sangiorgi & Spatt, 2017; Zhuo et al., 2017).

CRAs' actions have come under the scrutiny of investors and regulators several times in the past. The first significant event was the Enron bankruptcy in 2001, followed by the Worldcom bankruptcy in 2002, in which CRAs' role was criticized (Bedendo et al., 2018; Coskun, 2008). Following the Enron and Worldcom debacle, the Sarbanes–Oxley Act, 2002, and the Credit Rating Agency Reform Act, 2006 was passed in the US to increase credit rating quality by CRAs by increasing competition, transparency and accountability in the

[1]Delhi School of Management, Delhi Technological University, India

Corresponding Author:
Chandan Sharma, Delhi School of Management, Delhi Technological University, Shahbad Daulatpur, Delhi 110042, India.
Email: chandansharma@dtu.ac.in

Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License (https://creativecommons.org/licenses/by/4.0/) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).

credit rating industry (Coskun, 2008). However, despite the steps taken, CRAs' role in the 2008 financial crisis again came under scrutiny. Authorities concluded that CRAs inflated ratings were one of the driving forces of the financial crisis.

Consequently, the U.S. government filed civil lawsuits against CRAs for defrauding investors, which CRAs settled by paying heavy fines (The U.S. DoJ, 2015, 2017). Similarly, the regulator in India—Securities Exchange Board of India—had initiated actions against domestic CRAs for erroneous ratings, which the CRAs settled by paying fines (Business Today, 2018; CNBC TV18, 2020; Livemint, 2020). In each instance, the commonality was a sudden default or severe downgrade of a high-rated entity or instrument.

The main reason for such issues in the rating industry is the issuer-paid revenue model followed by CRAs and their competition to gain market share and increase revenues. The Issuer-pays model is an approach where the issuer of debt pays for evaluating its securities. Several researchers have highlighted that CRAs face a conflict of interest between providing informative ratings to investors and satisfying issuers' rating preferences (Camanho et al., 2022; Cavallaro & Trotta, 2019; Mathis et al., 2009; Partnoy, 2006).

Researchers have highlighted how competition between CRAs leads to such issues in the credit rating industry. Competition could result in CRAs assigning inflated ratings to a firm (Beatty et al., 2019; Bolton et al., 2012) as well as the tendency of firms to do rating shopping, that is, move from one CRA to another to get a higher rating (Goldstein & Yang, 2019; Huang & Shen, 2019). The paper analyzes the prevalence of such rating issues due to competition in the Indian credit rating industry.

The paper employs empirical techniques to test the impact of competition on credit rating. The paper focuses on testing the presence of rating inflation and rating shopping due to competition in the credit rating industry. Several researchers have analyzed the impact of competition on credit rating, but mainly through an indirect approach using theoretical models with assumptions (Bolton et al., 2012; Camanho et al., 2022; Chu & Rysman, 2019; Lee & Schantl, 2019) or using market share data of CRAs as a proxy of competition (Bae et al., 2019; Becker et al., 2011; Flynn & Ghent, 2018; Vu et al., 2022). Existing literature on competition in the credit rating industry has primarily focused on the US (Becker et al., 2011; Flynn & Ghent, 2018; Griffin et al., 2013) and other markets such as Canada, Korea, and China (Bae et al., 2019; Chu & Rysman, 2019; Vu et al., 2022).

As far as the author knows, this paper is the first to investigate the impact of competition among CRAs in India's domestic credit rating market. Moreover, it employs a distinct approach compared to existing works of literature. The paper uses dual rating, that is, multiple CRAs rating a firm, as a direct measure for competition to analyze its incremental impact on the firm's credit rating.

The paper makes several contributions to the literature. The paper finds additional evidence for rating inflation in the credit rating industry due to competition. Unlike existing studies that use theoretical models or market share as a proxy of competition, the paper uses a direct measure of competition in the form of dual ratings and finds that CRA inflates a firm's rating when facing direct competition from other CRAs. In addition, the paper provides evidence that points toward competition leading to CRAs being lenient while rating larger firms. The paper also provides evidence showing that competition leads to the initial rating assigned to a firm by a CRA being higher than the firm's existing rating from another CRA, resulting in rating shopping in the credit rating industry.

The paper findings raise concern regarding the impact of competition among CRAs on their role as mitigants of risk in the financial system. It augments the existing literature supporting the unintended consequences of competition in the credit rating industry. The paper's findings are helpful for regulators across countries to understand how competition between CRAs could manifest itself in the credit rating industry. Regulators need to regulate the competition in the industry from posing a systemic risk to the financial system over time. The paper's findings suggest that investors should be more cautious in utilizing credit ratings, especially when multiple CRAs are involved or when issuers change CRA.

Literature Review

CRAs' role in risk mitigation in the financial system has become increasingly important, given investors' higher reliance on credit rating for decision-making (Boot et al., 2006). The different rating levels assigned by CRAs to instruments or issuers correspond to different risk levels (Moody's Investors Service, 2021; S&P Global, 2021). Therefore, as per the risk-return tradeoff (Modigliani & Pogue, 1974), investors have a higher expectation of return from instruments rated at a higher risk level, with other conditions remaining the same. Thus, the capital cost of firms rated at a higher rating level (or lower risk level) will be lower and vice versa (Cantor & Packer, 1995). Consequently, credit rating changes affect a firm's future cost of capital (Tang, 2009).

Several researchers have explored the impact of credit ratings on firms' capital structure and found that credit rating changes can materially impact a firm's investment

and financing decisions. Kisgen (2006) concluded that firms issue less debt relative to equity when nearing a rating change compared to other firms. Kisgen (2009) found that firms reduce debt and are less likely to issue debt following a downgrade in credit rating. Similarly, Aktan et al. (2019) reported that firms issue less net debt than equity post a broad credit rating level change. The above highlights firms' concerns regarding their credit ratings and firms motivation to improve them.

The primary source of revenue for CRAs is the fee receipt for the rating service. The prevalent model in rating markets is the "Issuer-pays" model, wherein the issuer/firm, whose debt/bond is rated by the CRA, pays for the rating service. The issuer-pays model leads to a conflict of interest for the CRAs. CRAs are looking to increase revenue by providing rating services to new debt issuers and gaining more business from existing issuers. The impact of this conflict of interest on CRAs rating becomes more severe due to other CRAs competing for the same rating business. The conflict of interest and competition could lead to rating inflation, that is, CRAs assigning higher ratings to firms, which is not commensurate with issuers' creditworthiness and rating shopping, that is, firms moving from one CRA to another in order to get better ratings.

Competition among companies is natural in any industry, but unchecked competition in the credit rating industry leads to rating shopping by issuers and inflation of rating by CRAs (Gannon, 2012; He et al., 2012; Jiang & Packer, 2019; Xia & Strobl, 2012). The adverse impact of competition among CRAs is visible even in the domestic rating market. Park and Lee (2018) analyzed the Korean domestic rating market and found that CRAs and firms' actions can lower credit rating quality in a competitive market. Singh and Chavan (2020) highlighted that credit ratings sometimes lag the asset quality of borrowers in India, raising concern about the rating quality of Indian CRAs. This paper looks into India's domestic credit rating market and analyzes the impact of competition among CRAs on firms' domestic ratings.

Globally, the three major credit rating agencies are—S&P Global Ratings, Moody's Investor Service, and Fitch Ratings. These three CRAs account for around 95% of global rating markets (European CEO, 2020). The global credit rating market can be divided into two major parts—international and national or domestic ratings. Although the three CRAs dominate the global rating market, several other CRAs operate in different countries' national debt markets. A nation-specific regulator generally governs the operations of these CRAs in a national debt market.

There are currently seven CRAs in India accredited by the regulator—Securities and Exchange Board of India (SEBI)—for assigning credit ratings to domestic capital market issuance by firms (SEBI, 2021). These seven rating agencies have also been accredited to assign credit ratings to bank loans, which banks utilize for risk-weighting their disbursed loan for capital adequacy purposes (Reserve Bank of India, 2015). However, the top four agencies—CRISIL, ICRA, CARE, and India Ratings—accounted for over 88% of the domestic credit rating market in FY20 and are the oldest CRAs in India. Therefore, the paper focuses on these four agencies to understand the impact of competition on a firm's credit rating. Regulators in India has been actively trying to improve the credit rating quality given by CRAs. The regulator has imposed penalties on CRAs for violation of regulations or erroneous credit ratings in the past. CRAs had reached a settlement with SEBI for alleged breach of regulations (Business Standard, 2017; Business Today, 2018). SEBI has also fined several CRAs for lapses in assigning appropriate credit ratings (CNBC TV18, 2020; Livemint, 2020).

Regulators have brought several changes to address rating inflation by CRAs and rating shopping by issuers. However, the regulations have primarily been brought in response to severe lapses by CRAs. In 2002, the Sarbanes–Oxley Act was passed by regulators in the US, which ultimately led to the "Credit Agency Reform Act of 2006" augmenting the disclosure requirement and accountability of CRAs. Dodd-Frank Act, enacted in 2010 in the aftermath of the 2008 financial crisis, dealt with the conflict of interest in credit rating business, the liability of CRAs, and improving rating quality (Picciau, 2018). SEBI introduced CRA regulations in India through SEBI (Credit Rating Agencies) Regulations, 1999 (SEBI, 2022). SEBI has brought several amendments from time to time in response to failures of CRAs in India. In 2010, SEBI increased the disclosures, audit, and compliance requirements for CRAs (Prakash et al., 2017). Similarly, amendments have been brought in by SEBI in 2012, SEBI in 2016, SEBI in 2018 to increase transparency in the working of CRAs and ensure timely action by CRAs.

However, despite the regulator's continuous focus on improving credit ratings' reliability, the issues discussed above continue in the credit rating industry. The paper focuses on analyzing these issues of rating inflation and rating shopping among CRAs due to competition in the credit rating industry. Most researchers have analyzed the impact of competition on rating agencies indirectly. Several researchers have studied the competition among CRAs using a theoretical model with assumptions regarding the credit industry. Bolton et al. (2012) used a model to analyze the impact of competition on credit rating and found that competition reduces efficiency, increases rating shopping, and could result in rating inflation. Camanho et al. (2022) used a theoretical model

to analyze competition among CRAs and found that competition exacerbates the problem of rating inflation in the credit rating industry. Lee and Schantl (2019) used a model to analyze how the dynamics between competition among CRAs and their gatekeeper's role impact rating inflation in the industry.

Some researchers have used market share data as a proxy for competition to understand its impact on credit ratings. Vu et al. (2022) used market share data to investigate the impact on sovereign ratings due to competition between CRAs and concluded that competition lowers the quality of the ratings. Flynn and Ghent (2018) reported that incumbent CRAs inflate ratings as the market share of new entrants increases. Becker et al. (2011) used an increase in Fitch's market share to measure increased competition and found that the rating quality of CRAs declined as the market share of Fitch increased. However, Bae et al. (2015) contradicted the above finding and concluded controlling for unobservable industry effects; there is no linkage between competition measured by Fitch market share and rating inflation. Beatty et al. (2019) found that Moody's and Fitch, following the recalibration of their municipal debt rating scale in 2010, increased the credit rating of municipal bonds without a change in credit quality, resulting in increased market share. However, this paper tries to directly analyze whether the rating actions of CRAs are impacted due to competition by using dual ratings, that is, multiple CRAs rating a firm as a measure of competition among CRAs.

Hypothesis Development and Methodology

The paper's main objective is to test whether competition leads to rating inflation and rating shopping in the credit rating industry. The paper adopts multiple approaches to achieve this objective. The paper analyzes the rating inflation tendency of one of the top four domestic CRA (designated as CRA "A") due to competition from the other three of the top four domestic CRAs (designated as CRA "1," "2," and "3"). The tendency of issuers to indulge in rating shopping due to competition is investigated by analyzing the initial ratings given by other CRAs (CRA "1," "2," and "3") to firms already rated by CRA "A" and vice versa. The paper uses *t*-tests and regression techniques to empirically analyze the impact of competition among CRAs on firms' credit ratings.

The first approach compares key financial metrics of firms rated by only CRA "A" with those rated by more than one CRA through hypothesis 1. In cases where dual ratings are present, if firms' key metrics are significantly worse than those rated only by CRA "A" for the same rating category, the presence of rating inflation could be inferred.

> Hypothesis 1: There will be no difference in key credit metrics of firms rated by multiple CRAs compared to those rated by a single CRA for the same rating category.

As per hypothesis 1, in case of no rating inflation by CRA "A," the mean of key credit metrics of firms where more than one CRA (CRA "A" and one or more CRAs) are rating the firm should be equal to those firms in which only CRA "A" is rating the firms for the same rating category. Hypothesis 1 is tested using a two-sample *t*-test to check whether the mean credit metrics of the two groups are equal.

The second approach involves using regression to determine whether a CRA provides an inflated rating in the presence of competition. Here, it is directly tested, through hypothesis 2, that a CRA will provide a higher rating, controlling for firm-specific factors, to a firm if other CRAs also rate the same firm.

> Hypothesis 2: A CRA rating for a firm is not influenced by another CRA rating of the firm.

Therefore, in case of hypothesis 2 is rejected, it will mean that CRA "A" will tend to give a higher rating to a firm if another CRA is also rating the firm. Hypothesis 2 is tested using Ordinary Least Squares (OLS) regression and Ordered Probit regression, employed in several previous credit ratings related papers (Cubas-Díaz & Martínez Sedano, 2018; Gupta et al., 2017; Park & Lee, 2018). The firm's credit rating is converted into numerical values to run the regression, explained later in the Results and Discussion section. In case of rejection of hypothesis 2, the dummy variable coefficient related to dual ratings in the regression should be of appropriate sign and significant.

The third approach involves testing whether rating shopping is prevalent in the industry using hypothesis 3.

> Hypothesis 3: There are equal chances of a firm rated by a CRA getting an initial rating that is higher or lower, or equal to its existing rating from a new CRA.

In the case of rating shopping due to competition in the credit rating industry, the rating assigned by a new CRA to a firm will invariably be higher than the firm's current rating. Therefore, hypothesis 3 will be rejected. For hypothesis 3, statistics of the difference between a firm rating from CRA "A" and the initial rating assigned by other CRAs is analyzed to check the presence of rating shopping. In addition, to test hypothesis 3, the paper also uses Ordinary Least Squares (OLS) regression and Ordered Probit regression. In case CRA "A" assigns a higher initial rating to a firm rated by other CRAs, the dummy variable coefficient related to dual rating should

Table 1. Rating Category Wise Sample Distribution.

Rating category	AAA	AA +	AA	AA-	A +	A
No. of observations	208	111	491	502	730	640
Percentage of total observations	1.4	0.8	3.4	3.5	5.0	4.4
Dual ratings percentage	70.7	43.2	36.0	29.3	21.1	29.1
Rating category	A−	BBB +	BBB	BBB−	BB +	BB
No. of observations	872	1,044	1,379	1,765	1,428	1,292
Percentage of total observations	6.0	7.2	9.5	12.1	9.8	8.9
Dual ratings percentage	25.2	21.5	21.1	21.4	19.6	18.7
Rating category	BB−	B +	B	B−	C	D
No. of observations	861	979	604	243	102	1,310
Percentage of total observations	5.9	6.7	4.2	1.7	0.7	9.0
Dual ratings percentage	18.2	18.0	23.7	24.7	26.5	18.4

be positively correlated to the dependent variable (dependent variable is calculated by the difference between current rating by other CRAs and initial credit rating given by CRA 'A'). The dependent variable should also be positive.

Data and Summary Statistics

The paper utilizes data from the Centre for Monitoring Indian Economy (CMIE)—ProwessIQ Database for credit rating and firms' financial information. Several Indian studies have used the above database (Jaiswall & Bhattacharyya, 2016; Jaiswall & Raman, 2021; Jawed et al., 2019). The paper analyzes the competition among CRAs by focusing on India's domestic credit rating industry. The credit rating of firms used in the paper represents the Long term rating assigned to the firm's long-term bank loans or capital market debt. The paper utilizes the rating action of CRAs from FY08-FY20 for analysis as RBI regulation on bank loans rating came into existence from FY08. The sample firms include all firms for which credit rating is available in the database except financial firms (National Industrial Classification Code 64000-66999). The reason for excluding financial firms is that such firms' credit ratings are driven by significantly different factors than non-financial firms. If a firm does not witness a credit rating action in a period from a CRA, the previous period credit rating given by CRA is taken as the firm's outstanding rating from the CRA for analysis.

The rating-wise and year-wise sample distribution taken in the study is shown in Tables 1 and 2. The sample contains 14,561 firm years of observations. These observations correspond to the rating given to firms by CRA "A", as the paper primarily focuses on the impact of competition on ratings assigned by CRA "A". Sample observations are equally divided between investment grade (above BB + rating) and non-investment grade (below BBB- rating) observations. The observations are distributed normally with a peak at the boundary of the investment-grade rating of BBB-. Firm-year observations in which dual ratings are present, that is, one or more agencies are rating along with CRA "A," are around 22% of total observations.

Table 2. Year Wise Sample Distribution.

Year	Frequency	Percentage	Dual ratings percentage
2008	2	0.0	50.0
2009	11	0.1	63.6
2010	11	0.1	72.7
2011	647	4.4	15.3
2012	1,499	10.3	17.3
2013	1,755	12.1	19.8
2014	1,864	12.8	21.9
2015	1,857	12.8	24.5
2016	1,757	12.1	26.5
2017	1,547	10.6	24.1
2018	1,574	10.8	25.9
2019	1,578	10.8	25.3
2020	459	3.2	14.6
Total	14,561	100.0	22.6

Table 2 shows rating distribution across different years. The number of observations is lower in years immediately after implementing the regulation of credit rating for bank loans in 2008. However, from 2012 the number of observations is equally distributed at around 10% each year. The percentage of firm ratings in which dual agencies are present is more than 20% in almost all years.

Results and Discussion

Comparison of Key Financial Metrics of Firm Rated by Multiple CRAs

To test hypothesis 1, we examine key credit metrics of leverage (Debt/EBITDA), interest coverage (EBITDA/

Interest Expenses), and financial metrics of profitability (EBITDA/Sales) and sales to analyze the metrics of firms with or without dual ratings. We use a two-sample *t*-test to compare the mean credit metrics of the two groups. Several studies have highlighted that these financial metrics have a significant relationship with credit rating (Blume et al., 1998; Gray et al., 2006; Gupta et al., 2017). These financial metrics are winsorized at their 1% and 99% percentiles to eliminate outliers. Several other papers have done the winsorization of data (Baker & Wurgler, 2006; Bonsall et al., 2018; Chen et al., 2019; Chiang et al., 2019).

Table 3 shows the key financial metrics of firms. The firms have been divided into two groups: (1) Group 1—where only CRA "A" is the rating firm and (2) Group 2—dual ratings, that is, one or more CRA is the rating firm along with CRA "A". The mean and median leverage of firms, measured as total debt/EBITDA, in Group 2 is higher than Group 1 firms. This higher leverage of group 2 firms is visible in both investment grade and non-investment grade rating categories. Using a two-sample *t*-test, we find that the mean debt/EBITDA of Group 1 is significantly lower than the mean of Group 2 across rating categories at a 1% level. The mean interest coverage ratio of Group 1 is significantly (at a 1% level) higher than the mean interest coverage ratio of Group 2 in the investment-grade rating category but not in the non-investment grade rating category. The mean profitability of Group 2 is significantly higher than that in Group 1 (at a 1% level) in both investment and non-investment grade ratings. The mean sales of firms in Group 2 are significantly higher than firms in Group 1 across rating categories. Rating Category "C" and "D" have not been included in the analysis. This is because the firms with these ratings have a very high risk of default or are expected to be in default (SEBI, 2011)and timely servicing of debt rather than financial metrics is the primary driver of firm rating at this level. In addition, CRAs define the investment-grade category ratings as having moderate to the highest level of safety. In contrast, non-investment-grade category ratings are defined as moderate to very high risk of default for timely servicing of obligations. Therefore, the accuracy of rating in the investment-grade category assumes more significance from a risk perspective.

The above analysis indicates that hypothesis 1 is rejected, that is, the Financial metrics of firms in Group 1 and Group 2 should be equal. We see that two key credit metrics of firms—leverage and interest coverage—in Group 2 are significantly worse off than Group 1—especially across the investment-grade categories. Thus, CRA "A" assigns a higher rating to a firm with a weaker credit profile when faced with competition in the form of dual ratings. The above analysis confirms rating inflation by CRAs in India due to competition in the credit rating industry. In addition, the size of firms in Group 2, as measured by sales, is significantly larger than those in Group 1, indicating that larger firms are more likely to get inflated ratings due to competition among CRAs. This could be partly explained based on large firms being better positioned to pay fees for credit rating services to multiple CRAs. In addition, large firms have a higher level of debt and thus are likely to pay higher rating fees, as fees are linked to the debt (ICRA Limited, 2021; India Ratings, 2021), and thus are more sought after by multiple CRAs. He et al. (2012) drew a similar conclusion that large issuers received inflated ratings than small issuers in the US mortgage-backed securities market.

Effect of Competition on the Credit Rating of a Firm by a CRA

As per hypothesis 2, it is tested whether CRA "A" will tend to give a higher rating if one or more CRA are also rating the firm. The following equation is used to test the hypothesis 2:

$$CR_{it} = \beta_{\circ} + \beta_1 CRA^j D_{it} + \beta_2 Size_{it} + \beta_3 Leverage_{it} + \beta_4 Profitability_{it} + \beta_5 Interest\ Coverage_{it} + \xi_{it} \quad (1)$$

Where, CR_{it} is the credit rating of firm i at time t at CRA "A"
$CRA^j D_{it}$ is a dummy variable for dual rating which takes a value "1" if CRA j gives a rating to the firm i at time t, which has also been rated by CRA "A" and "0" otherwise
$Size_{it}$, the natural logarithm of total assets of firm i at period t;
$Leverage_{it}$, total debt divided by total debt and total equity of firm i at period t
$Profitability_{it}$, EBITDA divided by total assets of firm i at period t
$Interest\ Coverage_{it}$, EBITDA divided by total finance expenses of firm i at period t

Control variables of Size, Leverage, Profitability, and Interest Coverage account for firm-specific characteristics driving the firm's rating, as used in several other studies (Gupta et al., 2017; Jiang et al., 2012; Kisgen, 2006). As explained earlier, rating levels "C" and "D" have been excluded from our analysis. The paper employs Ordinary Least Squares (OLS) regression and Ordered Probit regression to test hypothesis 2. The dependent variable CR_{it} is generated by converting the credit rating of firm into numerical values as per Table 4. In case, CRA "A" inflates the rating due to competition from other CRAs, then the numerical value of rating of a firm by CRA "A," which is being rated by

Table 3. Key Financial Metrics of Firm With or Without Dual CRAs Across Rating Categories.

Key Metrics of Firms	Group-1: Firms rated by only CRA "A"			Group-2 : Firms rated by CRA "A" where at least one more CRA rating is outstanding		
	N	Mean	Median	*N*	Mean	Median
All ratings						
Leverage	9,174	2.77	2.94	2,946	4.34***	3.25***
Interest coverage	8,918	7.19	2.45	2,899	5.44***	2.53
Profitability	8,719	0.16	0.1	2,862	0.18***	0.11***
Sales	8,719	6,576	1,749	2,862	27,658.15***	4,067.5***
Investment grade ratings						
Leverage	5,064	2.39	2.06	1,838	3.79***	2.79***
Interest Coverage	4,957	10.96	3.82	1,823	7.03***	3.16***
Profitability	4,870	0.2	0.12	1,807	0.22**	0.13**
Sales	4,870	9,927.84	3,091.5	1,807	40,078.2***	8,600.5***
Non-investment grade ratings						
Leverage	4,110	3.25	4.07	1,108	5.25***	3.94***
Interest coverage	3,961	2.48	1.72	1,076	2.74	1.93***
Profitability	3,849	0.11	0.08	1,055	0.12	0.09**
Sales	3,849	2,335.13	893.6	1,055	6,384.58***	1,279***
AAA Category						
Leverage	57	1.39	0.21	145	1.01	1.52***
Interest coverage	57	58.1	28.91	144	24.77***	8.57**
Profitability	57	0.29	0.28	143	0.33	0.2*
Sales	57	120,000	97,600.2	143	200,000***	210,000**
AA Category						
Leverage	706	1.73	0.97	356	3.73***	2.48***
Interest coverage	703	21.93	9.82	356	8.34***	3.93***
Profitability	691	0.25	0.16	355	0.28	0.18*
Sales	691	21,962.36	11,717.9	355	71,207.84***	28,982.4***
A Category						
Leverage	1,537	2.14	1.67	540	2.98*	2.49***
Interest coverage	1,521	13.49	5.02	540	6.05***	3.5***
Profitability	1,507	0.22	0.14	537	0.23	0.14
Sales	1,507	9,300.54	4,183.6	537	20,924.05***	9,397.5***
BBB Category						
Leverage	2,764	2.71	2.72	797	4.86***	3.13***
Interest coverage	2,676	5.64	2.85	783	3.85***	2.46***
Profitability	2,615	0.17	0.11	772	0.16	0.11
Sales	2,615	4,745.22	2,122.3	772	10,336.69***	3,863.05***
BB Category						
Leverage	2,270	3.91	3.73	575	4.43	3.56
Interest coverage	2,195	3.07	1.99	560	3.48	2.26***
Profitability	2,110	0.11	0.08	546	0.12	0.1***
Sales	2,110	2,407.05	1,019.65	546	6,586.07***	1,426.1***
B Category						
Leverage	1,007	6.04	5.02	311	5.99	4.36***
Interest coverage	965	2.14	1.58	295	2.36	1.77***
Profitability	952	0.12	0.08	289	0.1	0.07
Sales	952	1,229.41	589.3	289	1,778.65***	918.4***

*, **, and *** indicate significance at the .1, .05, and .01 levels, respectively.

another CRA, will be lower compared to the rating of a firm where CRA "A" is the only CRA rating the firm, controlling for firm-specific factors. This means that for CRA "A" to give an inflated rating due to competition, the coefficient of the dummy variable $CRA^{j}D_{it}$, should be negative and significant.

The OLS and Ordered probit regression results using equation (1) are presented in Tables 5 to 7. Table 5 s1hows whether the ratings given by CRA "A" are inflated when CRA "1" is also rating the same firm. The results show separately β1, the dummy variable coefficient, for investment and non-investment grade ratings.

Table 4. Table for Rating Scale conversion to Numerical Scale.

Rating Level	Numerical value
AAA	1
AA +	2
AA	3
AA−	4
A +	5
A	6
A−	7
BBB +	8
BBB	9
BBB−	10
BB +	11
BB	12
BB−	13
B +	14
B	15
B−	16
C	17
D	18

For non-investment grade category ratings, β_1, the dummy variable coefficient, is positive and significant in OLS and ordered probit regression. However, the coefficient is negative and significant at a 1% level for investment grade. The negative coefficient of dummy variables in investment-grade category rating indicates that CRA "1" presence leads to CRA "A" inflating the issuer rating, that is, CRA "A" assigns better/higher ratings when faced with direct competition from CRA "1." As discussed earlier, the investment category has more significance than the non-investment category from a risk perspective.

Table 6 shows whether the ratings given by CRA "A" are inflated when CRA "2" is also rating the same firm. The results show that for both investment and non-investment grade category ratings, the dummy variable coefficient related to the dual ratings is negative and significant at 1% in both OLS and oprobit regression. The negative coefficient of dummy variable related to dual ratings indicates that in the presence of CRA "2" rating a firm, CRA "A" rating for the firm is inflated.

Table 7 shows rating inflation by CRA "A" in the presence of CRA "3". The results show that the dummy variable's coefficient related to the dual ratings is negative and significant for all rating categories at a 1% level. Tables 5 to 7 show that CRA 'A' inflates the rating level if another CRA, CRA "1" or CRA "2" or CRA "3", is rating the firm. As per OLS regression, the rating of a firm by CRA 'A' in the presence of other CRAs is 0.21-0.44 notches more than the rating of a firm in which only CRA "A" is assigning the rating, controlling for firm-specific characteristics.

Table 5. Impact on Credit Rating of a firm by CRA "A" due to the Presence of CRA "1".

Dependent variable: Rating code, CR_{it}

	OLS regression			Probit regression		
	All ratings	Investment grade ratings	Non-investment grade ratings	All ratings	Investment grade ratings	Non-investment grade ratings
main						
CRA^1D_{it}	0.284*** (0.078)	−0.317*** (0.076)	0.368*** (0.066)	0.080*** (0.028)	−0.155*** (0.037)	0.261*** (0.048)
Coefficient economic size				0.21	−0.377	0.438
Sales	−1.140*** (0.016)	−0.703*** (0.016)	−0.272*** (0.018)	−0.412*** (0.007)	−0.331*** (0.008)	−0.195*** (0.013)
Leverage	0.409*** (0.033)	0.238*** (0.029)	0.163*** (0.039)	0.143*** (0.012)	0.104*** (0.013)	0.112*** (0.028)
Profitability	−2.336*** (0.275)	−1.137*** (0.256)	−0.868*** (0.269)	−0.800*** (0.099)	−0.511*** (0.121)	−0.670*** (0.193)
Interest coverage	−0.002*** (0.000)	−0.001*** (0.000)	−0.002** (0.001)	−0.001*** (0.000)	−0.001*** (0.000)	−0.002** (0.001)
N	10,414	6,584	3,830	10,414	6,584	3,830
adj. R^2	0.360	0.263	0.073			
F	1,170.222	471.237	61.357			

*, **, and *** indicate significance at the .1, .05, and .01 levels, respectively.

Table 6. Impact on Credit Rating of a Firm by CRA "A" due to the Presence of CRA "2."

Dependent variable: Rating code, CR_{it}	OLS Regression			Probit regression		
	All ratings	Investment grade ratings	Non-investment grade ratings	All ratings	Investment grade ratings	Non-investment grade ratings
CRA^2D_{it}	−0.229** (0.097)	−0.216** (0.087)	−0.037 (0.101)	−0.093*** (0.035)	−0.094** (0.041)	−0.059 (0.075)
Coefficient economic size				−0.243	−0.230	−0.099
Sales	−1.123*** (0.016)	−0.707*** (0.016)	−0.263*** (0.018)	−0.406*** (0.007)	−0.333*** (0.008)	−0.188*** (0.013)
Leverage	0.415*** (0.033)	0.232*** (0.029)	0.163*** (0.039)	0.145*** (0.012)	0.100*** (0.013)	0.111*** (0.028)
Profitability	−2.335*** (0.275)	−1.184*** (0.257)	−0.780*** (0.270)	−0.800*** (0.099)	−0.530*** (0.121)	−0.605*** (0.193)
Interest coverage	−0.002*** (0.000)	−0.001*** (0.000)	−0.002** (0.001)	−0.001*** (0.000)	−0.001*** (0.000)	−0.002** (0.001)
N	10,414	6,584	3,830	10,414	6,584	3,830
adj. R^2	0.359	0.262	0.066			
F	1,167.783	468.215	54.718			

*, **, and *** indicate significance at the .1, .05, and .01 levels, respectively.

However, to determine the economic impact of dual CRA on a firm's credit rating, the paper employs the method used by Alp (2013). As per this method, the dual rating dummy coefficient estimates, β1 in the ordered probit regression, are in units of the latent variable CRit. The economic impact of the dummy variable on the credit ratings can be determined by converting the coefficient β1 to units of rating notches. This is done by dividing the estimated coefficient by the average distance (measured in terms of the latent variable) between the rating categories. The economic impact of the dummy variable for dual ratings, CRA^jD_{it} measures average rating difference between firms rated by CRA "A" where other CRAs are present (i.e., $CRA^jD_{it} = 1$) and where CRA "A" is the only CRA (i.e., $CRA^jD_{it} = 0$). It can be seen that in the case of CRA "1," "2," and "3" presence, CRA "A" rated firms on an average, are likely to have 0.38, 0.23, and 0.54 notches higher rating than those firms where CRA "A" is the only CRA rating the firms. This corresponds to a one-notch higher rating by CRA "A" for approximately one out of every four firms rated along with CRA "2," one out of every three firms rated with CRA "1," and one out of every two firms rated with CRA "3." The range of dummy variable, CRA^jD_{it}, coefficient's economic value for different CRAs could be attributed to the reputational concerns of CRA "A" viz.a.viz other CRAs in the credit rating industry.

Thus, the results from Tables 5 to 7 indicate that hypothesis 2 is rejected, that is, a CRA rating for a firm is not influenced by another CRA rating the firm, especially in the more critical investment-grade category. The results show that the firm rating given by CRA "A" is inflated in the presence of other CRAs, as the dummy variable coefficient is negative and significant in all the cases. CRA "A" actions could be attributed to it providing a higher or equal rating than other CRA to prolong its existing relationship with the firm. In the absence of a higher or equal rating, the firm may drop CRA "A." This corresponds to the multiple CRAs rating the firm, improving the firm's rating to a higher but permissible level due to competition as compared to the rating of a single CRA-rated firm. Ultimately, the issuer-pays model makes CRA's revenue dependent upon a firm's relationship, and the competition forces CRA to inflate ratings. Our findings align with the existing literature that competition among CRAs leads to rating inflation (Bolton et al., 2012; Camanho et al., 2022; Lee & Schantl, 2019; Zhuo et al., 2017).

Impact on Initial Rating by a CRA Due to Competition

The tendency of rating shopping by firms can be analyzed using the initial rating assigned by a CRA to a firm when another CRA is already rating the firm. In the case

Table 7. Impact on Credit Rating of a Firm by CRA "A" Due to the Presence of CRA "3".

Dependent variable: Rating code, CR_{it}	OLS regression			Probit regression		
	All ratings	Investment grade ratings	Non-investment grade ratings	All ratings	Investment grade ratings	Non-investment grade ratings
CRA^3D_{it}	−0.449*** (0.145)	−0.438*** (0.125)	0.067 (0.173)	−0.209*** (0.053)	−0.221*** (0.060)	0.033 (0.126)
Coefficient Economic Size				−0.548	−0.539	−0.056
Sales	−1.121*** (0.016)	−0.703*** (0.016)	−0.264*** (0.018)	−0.405*** (0.007)	−0.331*** (0.008)	−0.189*** (0.013)
Leverage	0.419*** (0.033)	0.236*** (0.029)	0.163*** (0.039)	0.147*** (0.012)	0.103*** (0.013)	0.112*** (0.028)
Profitability	−2.349*** (0.275)	−1.194*** (0.256)	−0.785*** (0.270)	−0.808*** (0.099)	−0.538*** (0.121)	−0.612*** (0.193)
Interest Coverage	−0.002*** (0.000)	−0.001*** (0.000)	−0.002** (0.001)	−0.001*** (0.000)	−0.001*** (0.000)	−0.002** (0.001)
N	10,414	6,584	3,830	10,414	6,584	3,830
adj. R^2	0.359	0.263	0.066			
F	1,169.051	469.866	54.722			

*, **, and *** indicate significance at the .1, .05, and .01 levels, respectively.

of rating shopping due to competition, the rating assigned by a new CRA will invariably be higher than the current rating. For testing hypothesis 3, we compare the initial rating assigned by other CRAs to a firm rated by CRA "A" to understand whether other CRAs assign a higher initial rating than the current rating given by CRA "A". In addition, we test hypothesis 3 by using regression to check whether CRA "A" assigns a higher initial rating to firms with existing ratings by other CRAs.

Initial Rating assigned by other CRAs to CRA "A" rated firms. We first convert the rating scale into a numerical scale for this analysis using Table 3. Then we compare the initial rating given by other CRAs and CRA "A". We calculate the difference between the current rating given by CRA "A" and the initial rating assigned by other CRAs. If rating shopping occurs, the difference would invariably be positive. The summary statistics of this rating difference are present in Table 8. The mean and median difference between CRA "A" rating and initial rating by other CRAs are positive for all the CRAs. The difference is also positive for investment and non-investment grade categories with all the CRAs. Table 8 shows CRA 1, 2, and 3 assign a higher rating to a firm with an existing rating from CRA "A" in more than 50% of cases, while a lower rating is assigned in only less than 10% of cases. This lends credence to the argument that competition leads to rating shopping in the credit rating industry, that is, invariably a firm gets higher than the current rating when it approaches a new CRA for rating. Competition among CRAs to acquire new business from the firm and the Issuer-pays business model is the primary reason for the rating shopping occurring in the credit rating industry.

Table 9 shows CRA "A" assigns a higher initial rating in more than 50% of cases in which CRA "1" has an incumbent rating, while in around 10% of cases, it assigns a lower rating. If a firm already has an incumbent rating from CRA "2" and CRA "3", CRA "A" assigns a higher initial rating in around 30% of cases. In comparison, it assigns a lower initial rating in only around 13% of cases. Thus, there is a skew toward CRA "A" assigning a higher or equal initial rating to firms where other CRA are already present.

Initial rating assigned by CRA "A" to other CRAs rated firms. We also test hypothesis 3 by testing, using regression, whether the initial rating assigned by CRA "A" to a firm rated by other CRAs is higher, controlling for firm-specific characteristics. The hypothesis is tested using the following equation:

Table 8. Summary of the Difference Between Credit Rating of a Firm by CRA "A" and Initial Rating by Other CRAs.

				Percentage of observations where new CRA rating is		
	N	Mean	Median	Higher	Equal	Lower
Comparison of initial rating of CRA "1" with rating of CRA "A"						
All rating	367	0.89	1.00	61.04	30.52	8.45
Investment grade rating	175	0.61	1.00	50.86	41.14	8.00
Non-Investment grade rating	192	1.16	1.00	70.31	20.83	8.85
Comparison of initial rating of CRA "2" with rating of CRA "A"						
All Rating	304	0.72	1.00	53.62	40.13	6.25
Investment grade rating	179	0.30	0.00	37.43	54.75	7.82
Non-Investment grade rating	125	1.31	1.00	76.80	19.20	4.00
Comparison of initial rating of CRA "3" with rating of CRA "A"						
All rating	167	1.02	1.00	52.10	44.91	2.99
Investment grade rating	123	0.58	0.00	39.84	56.91	3.25
Non-investment grade rating	44	2.25	2.00	86.36	11.36	2.27

Table 9. Summary of the Difference Between the Initial Rating of a Firm by CRA "A" and Rating by Other CRAs.

				Percentage of observations where new CRA rating is		
	N	Mean	Median	Higher	Equal	Lower
Comparison of initial rating of CRA "A" with rating of CRA "1"						
All Rating	503	0.75	1.00	52.3	37.0	10.7
Investment grade rating	288	0.73	1.00	53.5	40.6	5.9
Non-investment grade rating	215	0.77	1.00	50.7	32.1	17.2
Comparison of initial rating of CRA "A" with rating of CRA "2"						
All Rating	251	0.42	0.00	28.3	58.2	13.5
Investment grade rating	199	0.43	0.00	26.6	64.8	8.5
Non-Investment grade rating	52	0.38	0.00	34.6	32.7	32.7
Comparison of initial rating of CRA "A" with rating of CRA "3"						
All Rating	104	0.65	0.00	33.7	53.8	12.5
Investment grade rating	82	0.62	0.00	28.0	63.4	8.5
Non-Investment grade rating	22	0.77	1.00	54.5	18.2	27.3

$$D.CR^{j}_{it} = \beta_{\circ} + \beta_1 ICR^{j}_{it} + \beta_2 Size_{it} + \beta_3 Leverage_{it} + \beta_4 Profitability_{it} + \beta_5 Interest\ Coverage_{it} + \xi_{it} \quad (2)$$

Where, $D.CR^{i}_{A}$ is the difference in credit rating (converted to numerical scale) of a firm i at time t by CRA 'j' and CRA "A"

ICR^{j}_{it} is a dummy variable which takes a value "1" if a firm i gets a rating assigned by CRA "A," when it was already being rated by CRA "j"

$Size_{it}$, the natural logarithm of total assets of firm i at period t;

$Leverage_{it}$, total debt divided by total debt and total equity of firm i at period t

$Profitability_{it}$, EBITDA divided by total assets of firm i at period t

$Interest\ Coverage_{it}$, EBITDA divided by total finance expenses of firm i at period t

Control variables are the same as in equation (1). Rating level "C" and "D" has been excluded from our analysis. The test employs Ordinary Least Squares (OLS) and Ordered Probit regression. In Case CRA "A" assigns a higher initial rating to a firm rated by other issuers, then the difference between credit rating by other CRAs and CRA "A," that is, the dependent variable should be positive and the independent dummy variable, ICR^{j}_{it}, should be positively correlated to it. The results of equation (2) are shown in Tables 10 to 12.

Table 10 shows the coefficient estimates of regression of dependent variable $D.CR^{j}_{it}$, the difference between an existing rating of a firm "1" and initial rating by CRA "A," on the dummy variable ICR^{j}_{it}, while controlling for firm-specific factors, as per equation (2). Similarly, Tables 11 and 12 present coefficient estimates of regression 2 of CRA "1" with CRA "A" and CRA "2" with CRA "A", respectively.

Table 10. Rating Shopping: Initial Rating by CRA "A" to a Firm With CRA "1" as Incumbent CRA.

Dependent variable: Rating difference, $D.CR^j_{it}$	OLS regression			Probit regression		
	All ratings	Investment grade ratings	Non-investment grade ratings	All ratings	Investment grade ratings	Non-investment grade ratings
ICR^j_{it}	0.612*** (0.098)	0.345*** (0.093)	0.887*** (0.207)	0.435*** (0.060)	0.380*** (0.076)	0.474*** (0.100)
Sales	−0.038* (0.020)	−0.085*** (0.019)	−0.343*** (0.067)	−0.016 (0.012)	−0.067*** (0.016)	−0.142*** (0.033)
Leverage	−0.020 (0.024)	−0.026 (0.018)	0.647 (0.484)	−0.018 (0.014)	−0.026* (0.015)	0.314 (0.230)
Profitability	1.108** (0.473)	−0.012 (0.395)	6.964*** (1.526)	0.708** (0.284)	0.268 (0.316)	2.836*** (0.732)
Interest coverage	−0.000 (0.001)	−0.000 (0.001)	−0.048** (0.021)	−0.000 (0.001)	−0.000 (0.001)	−0.020** (0.010)
N	1,659	1,067	592	1,659	1,067	592
adj. R^2	.027	.031	.105			
F	10.095	7.752	14.796			

*, **, and *** indicate significance at the .1, .05, and .01 levels, respectively.

The coefficient of ICR^j_{it} in the OLS regression varies from 0.603 to 1.126 at a significance level of 1% in Tables 10 to 12. The dummy variable coefficient is also positive (0.435–0.616) and significant in probit regression, which along with the positive mean and median rating difference in Table 9, implies that CRA "A" tends to assign a higher initial rating to a firm than its existing rating. The economic magnitude of rating shopping can be seen from Table 9, which indicates that the initial rating assigned by CRA "A" to a firm, on average, is 0.42 to 0.75 notches higher than its current rating by another CRA. Thus, based on these findings, hypothesis 3 is rejected. This could be explained because CRA "A" ability to increase its revenue depends on acquiring new firms for ratings due to the issuer-pay model. Therefore, CRA "A" assigns a rating higher than the existing rating of the firm from another CRA to ensure the firm extends the relationship beyond the initial rating. Consequently, the initial rating assigned by a CRA to a firm is generally higher than the firm's existing rating from other CRAs. The results confirm rating shopping in the credit rating industry stemming from the competition among rating agencies.

The paper's findings confirm how competition among CRAs worsens the quality of the ratings in the credit rating industry. The findings suggest that due to competition, the issues of rating inflation and rating shopping were prevalent in the Indian domestic credit rating industry during FY08–FY20. The study finds that all four major CRAs in India tend to assign a higher than the existing rating to a firm already rated by another CRA. However, the study primarily focuses on the impact of competition on credit ratings assigned by CRA "A"; therefore, credit ratings by other CRAs may not be affected similarly. The paper does not look into the benefits accruing to a CRA from such practices in the absence of relevant data on CRAs fees. The motivation behind CRAs' actions, when faced with competition, is primarily attributed to the issuer-pays model followed in the credit rating industry.

The paper's findings are consistent with existing literature that rating inflation and shopping are evident in the credit rating industry driven by the industry's issuer-pays business model and competition among CRAs. Vu et al. (2022) found that S&P and Moody inflate sovereign credit ratings due to competition from Fitch. Becker et al. (2011) established that increased competition in the corporate bond market from Fitch resulted in inflated ratings from other incumbents. However, Bae et al. (2015) found no evidence of rating inflation due to competition from Fitch. Bae et al. (2019) presented evidence that the quality of bond ratings by DBRS deteriorated when faced with increased competition from S&P in Canada. Park and Lee (2018) showed the presence of

Table 11. Rating Shopping: Initial Rating by CRA "A" to a Firm With CRA "2" as Incumbent CRA.

Dependent variable: Rating difference, $D.CR^j_{it}$						
	OLS regression			Probit regression		
	All ratings	Investment grade ratings	Non-investment grade ratings	All ratings	Investment grade ratings	Non-investment grade ratings
ICR^j_{it}	0.603*** (0.110)	0.329*** (0.111)	1.125*** (0.357)	0.531*** (0.081)	0.395*** (0.093)	0.793*** (0.185)
Sales	−0.009 (0.021)	−0.080*** (0.022)	−0.277*** (0.099)	−0.001 (0.015)	−0.070*** (0.018)	−0.057 (0.052)
Leverage	0.238*** (0.088)	0.202** (0.084)	−0.124 (0.937)	0.079 (0.064)	0.040 (0.069)	0.130 (0.442)
Profitability	−2.913*** (0.547)	−2.471*** (0.543)	−3.862 (2.901)	−1.318*** (0.397)	−0.933** (0.442)	−1.008 (1.207)
Interest Coverage	−0.001 (0.003)	0.005 (0.003)	−0.114 (0.102)	−0.001 (0.002)	0.002 (0.002)	−0.051* (0.026)
N	1,048	815	143	1,048	815	233
adj. R^2	.052	.047	.133			
F	12.494	9.060	5.372			

*, **, and *** indicate significance at the .1, .05, and .01 levels, respectively.

Table 12. Rating Shopping: Initial Rating by CRA "A" to a Firm With CRA "3" as Incumbent CRA.

Dependent variable: Rating difference, $D.CR^j_{it}$						
	OLS regression			Probit regression		
	All ratings	Investment grade ratings	Non-investment grade ratings	All ratings	Investment grade ratings	Non-investment grade ratings
ICR^j_{it}	1.126*** (0.267)	0.881*** (0.252)	1.552 (1.114)	0.668*** (0.130)	0.616*** (0.145)	0.808** (0.322)
Sales	0.209*** (0.045)	0.157*** (0.046)	0.160 (0.344)	0.108*** (0.022)	0.085*** (0.027)	0.032 (0.077)
Leverage	−0.021 (0.042)	−0.055 (0.037)	1.050*** (0.381)	−0.031 (0.020)	−0.047** (0.020)	0.353* (0.192)
Profitability	−0.089 (0.844)	−0.095 (0.738)	−12.211 (9.097)	−0.162 (0.394)	−0.162 (0.402)	−1.076 (2.392)
Interest Coverage	−0.000 (0.000)	−0.000 (0.000)	0.207* (0.104)	−0.000 (0.000)	−0.000 (0.000)	0.027 (0.021)
N	440	364	49	440	364	76
adj. R^2	.069	.052	.203			
F	7.540	4.957	3.441			

*, **, and *** indicate significance at the .1, .05, and .01 levels, respectively.

rating shopping in the Korean bond market. Flynn and Ghent (2018) also reported rating shopping in the structured finance market in the US. Credit rating influences investors' risk pricing and the regulator's capital requirements prescribed for banks. Therefore, such issues in the credit rating industry are likely to impact efficient capital allocation.

Conclusion

The paper investigated how competition among CRAs in the credit rating industry impacts firms' credit rating. The results of this paper indicate that the competition among CRAs influences firms' credit rating. The paper focuses on empirical analysis of firms' credit ratings by the top four CRAs in India, focusing on how competition by other CRAs impacts firm rating of CRA "A". The paper found empirical evidence of rating inflation by a CRA due to competition from other CRAs. The paper's findings also indicate that CRAs tend to be lenient when rating large-size firms in the face of competition. In addition, the paper results also indicate that competition leads to a firm getting a higher initial rating from a CRA than the firm's existing rating, resulting in rating shopping in the credit rating industry. The findings of the study are consistent with prior literature, which supports that competition among CRAs worsens the quality of credit ratings (Becker et al., 2011; Bolton et al., 2012; Camanho et al., 2022; Flynn & Ghent, 2018; Park & Lee, 2018; Vu et al., 2022).

The paper findings have important implications for regulators and investors. In the past, regulators have sought to address the credit rating industry's issues by increasing competition by allowing more players in the credit rating market. In the US, Congress passed the Credit Rating Agency Reform Act in 2006, while in Europe, CRA III regulation, enacted in 2013–14, had provisions for increasing competition in the credit rating industry. In India, three additional CRAs have been accredited for credit rating since 2008. The study's findings indicate that the regulators need to be cautious while allowing more competition in the credit rating industry due to its adverse effects on rating quality. The study also informs investors about the hazards of relying solely on credit ratings for risk estimation and the upward bias in large issuers' credit ratings.

Penalizing CRAs, through monetary fines or temporary suspension, for misrating could be a way to make CRAs more cautious and curb the practice of rating inflation and rating shopping. However, this is more of a reactive approach and needs to be complemented with closer regulatory supervision and increased disclosure requirements to improve transparency and consistency in ratings. Controlling the allocation of cases to different CRAs could help curb rating shopping, as this will restrict firms' movement from one CRA to another for favorable ratings. Periodic rotation of a firm's rating between CRAs could address the rating inflation issue in the industry, with CRAs being aware that their relationship continuance with a firm is independent of the firm's rating. However, regulators also need to strike a balance between supervision and allowing CRA to rate firms freely, as a too-conservative approach by CRAs could raise the cost of capital for firms and result in inefficient allocation of resources.

There are certain limitations to the findings of the study. The existing competitive dynamics between CRAs may change in the future based on incentives for CRAs to indulge in such practices. The study considers only credit ratings for non-financial corporates, but domestic CRAs also assign ratings in the area of public finance, structured finance, and financial corporates. Future research can focus on how credit ratings in other areas are affected due to competition between CRAs. Future studies can also analyze how the relative market position in the industry impacts a CRA response to competition faced by it.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Chandan Sharma https://orcid.org/0000-0003-4725-1715

References

Aktan, B., Çelik, Ş., Abdulla, Y., & Alshakhoori, N. (2019). The impact of credit ratings on capital structure. *ISRA International Journal of Islamic Finance*, *11*(2), 226–245. https://doi.org/10.1108/ijif-03-2018-0028

Alp, A. (2013). Structural shifts in credit rating standards. *The Journal of Finance*, *68*(6), 2435–2470. https://doi.org/10.1111/jofi.12070

Bae, K. H., Driss, H., & Roberts, G. S. (2019). Does competition affect ratings quality? Evidence from Canadian corporate bonds. *Journal of Corporate Finance*, *58*, 605–623. https://doi.org/10.1016/j.jcorpfin.2019.07.009

Bae, K.-H., Kang, J.-K., & Wang, J. (2015). Does increased competition affect credit ratings? A reexamination of the effect of Fitch's market share on credit ratings in the

corporate bond market. *Journal of Financial and Quantitative Analysis*, *50*(5), 1011–1035. https://doi.org/10.1017/s0022109015000472

Baker, M., & Wurgler, J. (2006). Investor sentiment and the cross-section of stock returns. *The Journal of Finance*, *61*(4), 1645–1680. https://doi.org/10.1111/j.1540-6261.2006.00885.x

Beatty, A., Gillette, J., Petacchi, R., & Weber, J. (2019). Do rating agencies benefit from providing higher ratings? Evidence from the consequences of municipal bond ratings recalibration. *Accounting Research Journal*, *57*(2), 323–354. https://doi.org/10.1111/1475-679x.12263

Becker, B., Milbourn, T., Dinc, S., Gopalan, R., Greenwood, R., Hollifield, B., Mian, A., Meltz, M., Pearson, N., Petersen, M., & Pollet, J. (2011). How did increased competition affect credit ratings?. *Journal of Financial Economics*, *101*(3), 493–514. https://doi.org/10.1016/j.jfineco.2011.03.012

Bedendo, M., Cathcart, L., & El-Jahel, L. (2018). Reputational shocks and the information content of credit ratings. *Journal of Financial Stability*, *34*, 44–60. https://doi.org/10.1016/j.jfs.2017.12.003

Blume, M. E., Lim, F., & Mackinlay, A. C. (1998). The declining credit quality of U.S. corporate debt: Myth or reality? *The Journal of Finance*, *53*(4), 1389–1413. https://doi.org/10.1111/0022-1082.00057

Bolton, P., Freixas, X., & Shapiro, J. (2012). The credit ratings game. *The Journal of Finance*, *67*(1), 85–111. https://doi.org/10.1111/j.1540-6261.2011.01708.x

Bonsall, S., Koharki, K., & Neamtiu, M. (2015). The effectiveness of credit rating agency monitoring: Evidence from asset securitizations. *The Accounting Review*, *90*(5), 1779–1810. https://doi.org/10.2308/accr-51028

Bonsall, S. B., Green, J. R., & Muller, K. A. (2018). Are credit ratings more rigorous for widely covered firms? *The Accounting Review*, *93*(6), 61–94. https://doi.org/10.2308/accr-52044

Boot, A. W. A., Milbourn, T. T., & Schmeits, A. (2006). Credit ratings as coordination mechanisms. *Review of Financial Studies*, *19*(1), 81–118. https://doi.org/10.1093/rfs/hhj009

Business Standard. (2017). *Crisil settles case with SEBI, pays Rs 28 lakh*. https://www.business-standard.com/article/pti-stories/crisil-settles-case-with-sebi-pays-rs-28-lakh-118013101525_1.html

Business Today. (2018). *CARE settles case with SEBI, pays Rs 43 lakh*. https://www.businesstoday.in/pti-feed/care-settles-case-with-sebi-pays-rs-43-lakh/story/277063.html

Camanho, N., Deb, P., & Liu, Z. (2022). Credit rating and competition. *International Journal of Economics and Finance*, *27*, 2873–2897. https://doi.org/10.1002/ijfe.2303

Cantor, R., & Packer, F. (1995). The credit rating industry. *The Journal of Fixed Income*, *5*(3), 10–34. https://doi.org/10.3905/jfi.1995.408153

Cavallaro, G., & Trotta, A. (2019). Role and regulation of Credit Rating Agencies: The European Way. *International Journal of Accounting and Financial Reporting*, *9*(1), 209. https://doi.org/10.5296/ijafr.v9i1.14021

Chen, Z., Harford, J., & Kamara, A. (2019). Operating leverage, profitability, and capital structure. *Journal of Financial and Quantitative Analysis*, *54*(1), 369–392. https://doi.org/10.1017/s0022109018000595

Chiang, C., Dai, W., Fan, J., Hong, H., & Tu, J. (2019). Robust measures of earnings surprises. *The Journal of Finance*, *74*(2), 943–983. https://doi.org/10.1111/jofi.12746

Chu, C. S., & Rysman, M. (2019). Competition and strategic incentives in the market for credit ratings: Empirics of the financial crisis of 2007. *American Economic Review*, *109*(10), 3514–3555. https://doi.org/10.1257/aer.20140773

CNBC TV18. (2020). *RCom case: SEBI slaps Rs 1 crore fine on CARE ratings for lapses in assigning credit rating*. https://www.cnbctv18.com/finance/rcom-case-sebi-slaps- rs-1-crore-fine-on-care-ratings-for-lapses-in-assigning-credit-rating-6460201.htm

Coskun, D. (2008). Credit rating agencies in a post-enron world: Congress revisits the NRSRO concept. *Journal of Banking Regulation*, *9*(4), 264–283. https://doi.org/10.1057/jbr.2008.17

Cubas-Díaz, M., & Martínez Sedano, M. (2018). Do credit ratings take into account the sustainability performance of companies? *Sustainability*, *10*(11), 4272. https://doi.org/10.3390/su10114272

Duff, A., & Einig, S. (2015). Debt issuer: Credit rating agency relations and the trinity of solicitude: An empirical study of the role of commitment. *Journal of Business Ethics*, *129*(3), 553–569. https://doi.org/10.1007/s10551-014-2175-y

EU Commission. (2021). *Competition policy*. Author. https://ec.europa.eu/competition-policy/index_en

European CEO. (2020). *New players are finally disrupting the credit rating sector*. https://www.europeanceo.com/finance/new-players-are-finally-disrupting-the-credit-rating-sector/

Flynn, S., & Ghent, A. (2018). Competition and credit ratings after the fall. *Management Science*, *64*(4), 1672–1692. https://doi.org/10.1287/mnsc.2016.2604

FTC. (2021). *FTC fact sheet: how competition works*. https://doi.org/10.1017/cbo9780511510854.006

Goldstein, I., & Yang, L. (2019). Good disclosure, bad disclosure. *Journal of Financial Economics*, *131*(1), 118–138. https://doi.org/10.1016/j.jfineco.2018.08.004

Gray, S., Mirkovic, A., & Ragunathan, V. (2006). The determinants of credit ratings: Australian evidence. *Australian Journal of Management*, *31*(2), 333–354. https://doi.org/10.1177/031289620603100208

Griffin, J. M., Nickerson, J., & Tang, D. Y. (2013). Rating shopping or Catering? An examination of the response to competitive pressure for CDO credit ratings. *Review of Financial Studies*, *26*(9), 2270–2310. https://doi.org/10.1093/rfs/hht036

Gupta, R., Gupta, D. B., & Hardeep, C. (2017). Financial determinants of credit ratings of Indian companies. *International Journal of Business Insights & Transformation*, *10*(2), 30–37.

He, J. J., Qian, J. Q., & Strahan, P. E. (2012). Are all ratings created equal? The impact of issuer size on the pricing of mortgage-backed securities. *The Journal of Finance*, *67*(6), 2097–2137. https://doi.org/10.1111/j.1540-6261.2012.01782.x

Huang, Y.-L., & Shen, C.-H. (2019). What role does the investor-paid rating agency play in China? Competitor or information provider. *International Review of Economics & Finance*, *63*, 253–272. https://doi.org/10.1016/j.iref.2018.11.007

ICRA Limited. (2021). *General nature of compensation arrangements with rated entities*. https://www.icra.in/Rating/ShowRatingPolicyReport/?id = 56

India Ratings. (2021). *Compensation arrangements with rated entities for bank loan ratings*. https://www.indiaratings.co.in/regulatory-disclosurs/bank-loan-ratings

Jaiswall, S. S., & Raman, K. K. (2021). Sales growth, CEO pay, and corporate governance in India. *Journal of Accounting Auditing & Finance*, *36*(2), 249–277. https://doi.org/10.1177/0148558x19825672

Jaiswall, S. S. K., & Bhattacharyya, A. K. (2016). Corporate governance and CEO compensation in Indian firms. *Journal of Contemporary Accounting & Economics*, *12*(2), 159–175. https://doi.org/10.1016/j.jcae.2016.06.001

Jawed, M. S., Dhaigude, A. S., & Tapar, A. V. (2019). The sectoral effect of demonetization on the economy: Evidence from early reaction of the Indian stock markets. *Cogent Economics & Finance*, *7*(1), 1–18. https://doi.org/10.1080/23322039.2019.1595992

Jiang, J., Harris Stanford, M., & Xie, Y. (2012). Does it matter who pays for bond ratings? Historical evidence. *Journal of Financial Economics*, *105*(3), 607–621. https://doi.org/10.1016/j.jfineco.2012.04.001

Jiang, X., & Packer, F. (2019). Credit ratings of Chinese firms by domestic and global agencies: Assessing the determinants and impact. *Journal of Banking & Finance*, *105*, 178–193. https://doi.org/10.1016/j.jbankfin.2019.05.011

Kim, S., & Park, S. (2016). Credit rating inflation during the 2000s: Lessons from the U.S. state governments. *International Journal of Economics and Financial Issues*, *6*(1), 13–19.

Kisgen, D. J. (2006). Credit ratings and capital structure. *The Journal of Finance*, *61*(3), 1035–1072. https://doi.org/10.1111/j.1540-6261.2006.00866.x

Kisgen, D. J. (2009). Do firms target credit ratings or leverage levels? *Journal of Financial and Quantitative Analysis*, *44*(6), 1323–1344. https://doi.org/10.1017/s002210900999041x

Lee, K., & Schantl, S. F. (2019). Financial reporting and credit ratings: On the effects of competition in the rating industry and rating agencies' gatekeeper role. *Accounting Research Journal*, *57*(2), 545–600. https://doi.org/10.1111/1475-679x.12264

Livemint. (2020). *IL&FS: Sebi enhances penalty to ₹1 crore on 3 credit rating agencies*. https://www.livemint.com/market/stock-market-news/il-fs-sebi-enhances-penalty-to-rs-1-crore-on-3-credit-rating-agencies-11600824065515.html

Mathis, J., McAndrews, J., & Rochet, J. C. (2009). Rating the raters: Are reputation concerns powerful enough to discipline rating agencies? *Journal of Monetary Economics*, *56*(5), 657–674. https://doi.org/10.1016/j.jmoneco.2009.04 .004

Modigliani, F., & Pogue, G. A. (1974). An introduction to risk and return: Concepts and evidence. *Financial Analysts Journal*, *30*(2), 68–80.

Moody's Investors Service. (2021). *What is a credit rating?* Author. https://ratings.moodys.io/ratings

Park, G., & Lee, H. Y. (2018). Opportunistic behaviors of credit rating agencies and bond issuers. *Pacific-Basin Finance Journal*, *47*, 39–59. https://doi.org/10.1016/j.pacfin.2017.11.003

Partnoy, F. (2006). How and why credit rating agencies are not like other gatekeepers.*Financial gatekeepers: can they protect investors?* (Working Paper No. 07–46). https://papers.ssrn.com/sol3/papers.cfm?abstract_id = 900257

Picciau, C. (2018). The evolution of the liability of credit rating agencies in the United States and in the European Union: Regulation after the crisis. *European Company and Financial Law Review*, *15*(2), 339–402. https://doi.org/10.1515/ecfr-2018-0012

Prakash, S., Ayachit, A., & Garg, S. (2017). *Regulation of credit rating agencies in India*. https://vidhilegalpolicy.in/wp-content/uploads/2020/06/170731_CRAReport.pdf

Reserve Bank of India. (2015). *Master circular – Basel III capital regulations*. https://www.rbi.org.in/scripts/BS_ViewMasCirculardetails.aspx?id = 9859

S&P Global. (2021). *Intro to credit ratings - A credit rating is an informed opinion*. Author. https://www.spglobal.com/ratings/en/about/intro-to-credit-ratings

Sangiorgi, F., & Spatt, C. (2017). Opacity, credit rating shopping and bias. *Management Science*, *63*(12), 4016–4036. https://doi.org/10.1287/mnsc.2016.2558

SEBI. (2011). *Standardisation of rating symbols and definitions*. Securities and Exchange Board of India, Circular No: CIR/MIRSD/4/2011, pp. 1–8. https://www.sebi.gov.in/legal/circulars/jun-2011/standardisation-of-rating-symbols-and-definitions_19860.html

SEBI. (2022). *Securities and exchange board of India (Credit Rating Agencies) regulations, 1999* [Last amended on January 24, 2022]. https://www.sebi.gov.in/legal/regulations/jan-2022/securities-and-exchange-board-of-india-credit-rating-agencies-regulations-1999-last-amended-on-january-24-2022_55541.html

Singh, S., & Chavan, P. (2020). *Efficacy of credit ratings in assessing asset quality : an analysis of large borrowers**. Reserve Bank of India. https://rbidocs.rbi.org.in/rdocs/Bulletin/PDFs/2EFFICACYAD3A2D69BC3D486EA52B172436F041DA.PDF

Tang, T. T. (2009). Information asymmetry and firms' credit market access: Evidence from Moody's credit rating format refinement. *Journal of Financial Economics*, *93*(2), 325–351. https://doi.org/10.1016/j.jfineco.2008.07.007

The U.S. DoJ. (2015). *Justice department and state partners secure $1.375 billion settlement with S&P for defrauding investors in the lead up to the financial crisis*. https://www.justice.gov/opa/pr/justice-department-and-state-partners-secure-1375-billion-settlement-sp-defrauding- investors

The U.S. DoJ. (2017). *Justice department and state partners secure nearly $ 864 million settlement with moody's arising from conduct in the lead up to the financial crisis settlement involves significant commitment by moody's to improve business practices*. https://www.justice.gov/opa/pr/justice-department-and-state-partners-secure-nearly-864-million-settlement-moody-s-arising

Vu, H., Alsakka, R., & Ap Gwilym, O. (2022). Does competition improve sovereign credit rating quality? *Journal of International Financial Markets Institutions and Money*, *76*, 101478. https://doi.org/10.1016/j.intfin.2021.101478

Xia, H., & Strobl, G. (2012). The issuer-pays rating model and ratings inflation: Evidence from corporate credit ratings. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2002186

Zhuo, X., Xu, G., & Wang, Y. (2017). The issuer-pays business model and competitive rating market: Rating network structure. *The Journal of Real Estate Finance and Economics*, *55*(2), 216–241. https://doi.org/10.1007/s11146-016-9563-2

2023 13th International Conference on Cloud Computing, Data Science & Engineering (Confluence) | 978-1-6654-6263-1/23/$31.00 ©2023 IEEE | DOI: 10.1109/Confluence56041.2023.10048868

Impact of Word Embedding Methods on Software Vulnerability Severity Prediction Models

Ruchika Malhotra
Department of Software Engineering
Delhi Technological University
Delhi, India
ruchikamalhotra@dtu.ac.in

Vidushi
Department of Software Engineering
Delhi Technological University
Delhi, India
vidushirf@gmail.com

***Abstract*— Software vulnerability severity prediction has recently gained popularity in software engineering. To perform this task, a prediction model has to be developed. The model's performance greatly depends on the feature vectors used to train it. These feature vectors are formed using textual data present on software vulnerability and converting it to numerical form. There are various methods for feature vector representation using word embeddings. Unlike traditional methods such as TF-IDF, word embedding methods using deep learning map the words into vectors, preserving the semantic relationships between words as well as doing automatic feature selection, which holds great importance when dealing with textual data. Therefore, word embedding methods using deep learning showed promising results over traditional TF-IDF. Hence, in this paper, we conducted a controlled experiment to examine the effect of various word embedding methods for feature vector representation on the performance of the prediction model. The models developed in this study used TF-IDF, word2vec, and GloVe, coupled with SVM, LSTM, and Bi-LSTM classification algorithms. The classification performance of different word embedding methods is analyzed on five different vulnerability datasets of Mozilla products. The experiments showed that the best performance is achieved by using the GloVe method with the Bi-LSTM algorithm giving an average AUC value of 0.81, which is an improvement of 12.5% from the traditional method of feature vector representation.**

***Keywords*— Feature Vectors, Word Embeddings, Software Vulnerability Severity, Prediction Model**

I. INTRODUCTION

With human dependency on software systems for almost all their activities, the requirement arises for different types of software systems. Therefore, it is important to take care of the security aspect of these software systems [1-3]. Nowadays, software systems are huge and complex. Hence, it is hard to build software that has zero errors. The count of software vulnerabilities being reported is also increasing exponentially [4], [5]. If a vulnerability is reported, its scope of getting exploited increases massively [6]. Therefore, it is required to develop a model, which automates the task of predicting the severity of the vulnerability. This is required for assigning higher priority to more severe vulnerabilities so that they can be fixed first. The model that can perform this task is named Software Vulnerability Severity Prediction Model (SVSPM). Severity to vulnerabilities is assigned based on various types of vulnerability scoring frameworks [7-9]. Among these frameworks, the Common Vulnerability Scoring System (CVSS) [10] by the Forum of Incident Response and Security Teams (FIRST) is considered to be a standard framework. Once a vulnerability is discovered, it is documented in the Common Vulnerabilities and Exposures (CVE) database [11]. This database contains vulnerability reports which have vulnerability descriptions for all vulnerabilities. This vulnerability description is textual data.

When such textual data act as input to SVSPM, the challenge arises to convert the textual data to the numeric form, so that it becomes meaningful to various classification algorithms. This leads to the requirement of applying various Feature Vector Representation (FVR) methods. Methods such as Term Frequency- Inverse Document Frequency (TF-IDF), Word2Vec, and Global Vector (GloVe) enable words to be expressed in numbers [12]. Traditional methods such as TF-IDF are simple, but it doesn't capture the relationships between tokens ending up with a large dimension feature vector. Whereas methods such as Word2Vec preserve the semantics and context of words, also allowing customize the length of feature vectors [13], thus helping in automatic feature selection.

Therefore, in this study, SVSPMs are developed experimenting with various word embedding methods of FVR using Machine Learning (ML) and Deep Learning (DL) algorithms for classification, and their results are compared.

The contribution of the paper can be highlighted in the points mentioned below:

- Comparison of the DL approach is made against the traditional ML approach with different FVR methods.
- Comprehensive discussion is done over the performance of the models developed.

The following RQs are addressed in this study:

RQ1: What is the effect of different word embedding methods for FVR on the performance of SVSPM?

The effectiveness of various word embedding methods used for FVR is measured on the dataset of five different products of Mozilla using Area Under Curve (AUC) as a performance measure.

RQ2: Which word embedding method for FVR gave the best results?

Addressing this RQ, the most effective word embedding method of FVR is found, and final judgment is given, stating the best model.

RQ3: Whether the performance of SVSPM giving the best results, boosted significantly?

Based on the results of the Friedman test and Wilcoxon Signed Rank test, a significant difference in the performance of SVSPM is found.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 01,2023 at 07:28:33 UTC from IEEE Xplore. Restrictions apply.

II. Related work

This section briefly reviews the previous work done in recent years related to our research, i.e., severity prediction of software vulnerabilities using textual data available on vulnerabilities, both in terms of the topic as well as in terms of techniques used.

A. Software Vulnerability Severity Prediction

Kaya et al. [14] in their experiments used four data balancing methods, seven classification algorithms, and three feature types. According to the result analysis, it can be stated that, among feature types, when text-based features are used better results are obtained. Also, ensemble-based classifiers provided better results. Liu et al. [30] used three deep learning and one ensemble method for developing the model. TextRCNN came out to be the best-performing one. Nakagawa et al. [15] proposed a character-level Convolution Neural Network (CNN) to conduct the experiment and found that it outperforms the conventional one in terms of accuracy. Bulut et al. [16] used Latent Dirichlet Allocation (LDA) for document representation and Naïve Bayes (NB), Fisher's Linear Discriminant Algorithm (FLDA), and Random Tree (RT) for classification. Huang et al. [17] used TF-IDF, Information Gain (IG), and Deep Neural Network (DNN) for building the model. It outperformed when compared with Support Vector Machine (SVM), NB, and K-Nearest Neighbours (K-NN). Chen et al. [18] introduced a new approach that is based on Term Frequency and Inverse Gravity Moment (TF-IGM) and showed that it outperformed TF-IDF. Kudjo et al. [19] used Random Forest (RF), DNN, K-NN, Logistic Regression (LR), and SVM. Wang et al. [20] used Principal Component Analysis (PCA) for document representation and Extreme Gradient Boosting (XGBoost) for classification. XGBoost outperformed as compared to SVM, LR, and RF. Aota et al. [21] used SVM, LR, Decision Tree (DT), and RF. Result Analysis gave the conclusion that RF performed the best. Kudjo et al. [22], after performing preprocessing of data applied DNN, LR, K-NN, and RF for classification.

B. DL-Based Word Embedding Methods

Word embeddings providing low-dimension feature vector representation, thus helping in automatic feature selection are gaining popularity in the NLP community [23-24]. In the field of software vulnerability, Sahin et al. [25], used word2vec among word embedding models for FVR along with CNN, Long Short Term Memory (LSTM), and XGBoost. It is concluded that LSTM outperformed others. Fang et al. [26] used FastText and LightGBM (Light Gradient Boosting Machine) which outperformed when compared with SVM, RF, and LR using TF-IDF for document representation. Sharma et al. [27] used GloVe and CNN for model development.

In these recent works, the main focus of the author lies in finding the best-performing model based on the classification algorithm. Moreover, the use of DL-based word embedding for FVR is very limited. Bag-of-Words (BOW) along with TF-IDF is mostly used. Hence, in this study, we decided to focus on the performance comparison of SVSPMs based on various word embedding methods for FVR.

III. Experimental Settings

In this section of the paper, we listed the methods which are used to carry out the experiment of this study.

A. Datasets

Vulnerability reports of five different products of Mozilla are used as datasets in this study. These vulnerability reports are offered by the CVE [11]. CVSS framework [10] is followed for scoring the severity of the vulnerability. It assigns a score to vulnerability in the range of 0 to 10. According to version v3.0 of CVSS, a 'None' rating is given to vulnerabilities with a 0.0 score, a 'Low' rating is given to vulnerabilities with a score of 0.1-3.9, a 'Medium' rating is given to vulnerabilities with a score 4-6.9, 'High' rating is given to vulnerabilities with a score 7-8.9 and 'Critical' rating is given to vulnerabilities with a score 9-10. Fig. 1 presents an example of a vulnerability and all the attributes present in the vulnerability report.

#	CVE ID	CWE ID	# of Exploits	Vulnerability Type(s)	Publish Date	Update Date	Score	Gained Access Level	Access	Complexity	Authentication	Conf.	Integ.	Avail.
1	CVE-2021-38502			Exec Code	2021-11-03	2022-07-12	4.3	None	Remote	Medium	Not required	Partial	None	None

Thunderbird ignored the configuration to require STARTTLS security for an SMTP connection. A MITM could perform a downgrade attack to intercept transmitted messages, or could take control of the authenticated session to execute SMTP commands chosen by the MITM. If an unprotected authentication method was configured, the MITM could obtain the authentication credentials, too. This vulnerability affects Thunderbird < 91.2.

Fig.1. Example of Vulnerability Data

B. Word Embedding Methods for FVR

In this study, we have used TF-IDF, word2vec, and GloVE for FVR. These methods are described briefly below in the text.

- TF-IDF: It [32] comes under traditional word embedding as it is based on the frequency of the word in the document. A weight is calculated for every word based on its frequency and it is assigned to that word.
- Word2vec: In this work, word2vec continuous skip-gram model [33] is used as the word embedding technique.
- GloVe: It [34] is a word-level pre-trained word embedding that uses a word co-occurrence matrix for generating word embeddings.

Although TF-IDF is simple and easy to implement, but it does not capture the semantic and syntactic similarity between words, like word2vec. Above that, GloVe defines the global context, unlike word2vec which defines a local context. Therefore, looking at the advantages and disadvantages of these methods, in this study we decided to compare their performances.

C. Classification Algorithms

In this study, we choose to work with SVM, LSTM, and Bi-LSTM. The idea of choosing these algorithms is to compare traditional ML algorithm with DL algorithms. These three algorithms are described briefly below in the text:

- SVM: It is a supervised ML algorithm. It works well for both classifications and regression problems [28]. The model uses a set of points located as wide as possible, which follow different categories. These sets of points are nothing but data instances. An optimal hyperplane is found which basically separates these instances, maximizing the margin among them.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 01,2023 at 07:28:33 UTC from IEEE Xplore. Restrictions apply.

- LSTM and Bi-LSTM: CNN can also be used to work on textual data. But it is incapable of linking the present information with the past information. Therefore, Bi-LSTM can be used. It [29] is a kind of LSTM model which belongs to the category of Recurrent Neural Network (RNN). It can process the data in both forward and backward directions making Bi-LSTM different from the LSTM network. Good results are recorded using Bi-LSTM in natural language processing [30-31].

IV. Proposed Framework

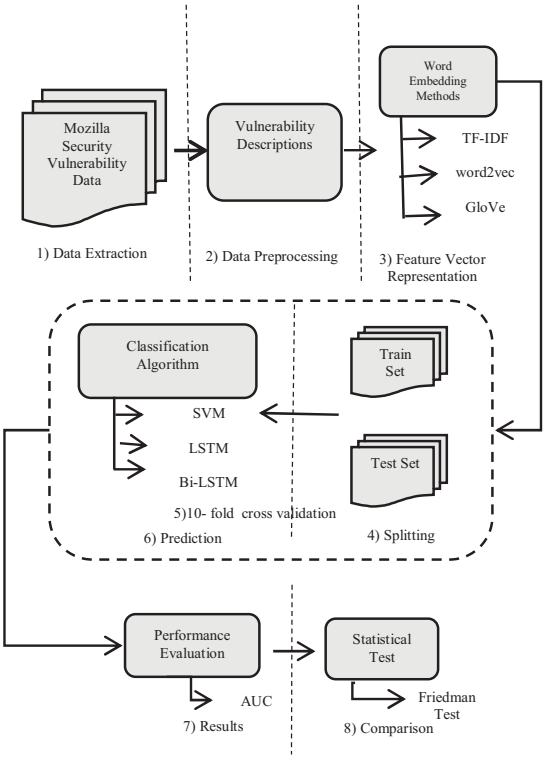


Fig. 2. Proposed Framework

In this section of the paper, the methodology followed to develop the SVSPM is explained. As depicted in Fig. 2, the methodology can be summarized in eight steps. In Steps 1, 2, and 3, input data to the classification model is generated from raw unstructured data. This data is preprocessed using traditional steps of preprocessing i.e., tokenization, stop word removal, and stemming. In the case of TF-IDF, all three steps of data preprocessing are applied. Although, when word2vec and GloVe are used, stop words are not removed as they provide context to their neighboring words. After data preprocessing is done, in Step 3, feature vectors are formed. The clear benefit of word2vec and Glove over TF-IDF is it preserves the semantics of the words as well as produces low-dimension feature vectors, helping in automatic feature selection as well. What impact it will have on the performance of the SVSPM is analyzed in the next section of the paper. The data once ready in numeric form is split into training and testing sets according to the 70/30 ratio in Step 4. In Step 6, the model is trained over the training set using the classification algorithms, and in Step 7, performance is analyzed of these trained models over the test set using AUC. AUC as a performance measure is a popular choice among researchers as it finds a balance between both specificity and sensitivity values and is considered to be a good choice when dealing with imbalanced data problems [35]. In Step 8, it is found if the results are statistically significant or not and finally conclusion is derived.

V. Result Analysis

In this section, the RQs listed in the paper's introduction section are addressed.

RQ1: What is the effect of different word embedding methods for FVR on the performance of SVSPM?

To capture the effect of the different word embedding methods for FVR on the performance of SVSPM, the AUC values are recorded for all the models developed and for all levels of severity of the vulnerability, as seen in Table 1. When these values are analyzed, it can be concluded that the word embedding methods affects the performance of SVSPM as different methods lead to different value of AUC.

TABLE I. Result Values of all SVSPM

Database	Model	Low	Medium	High	Critical	Mean
MF	M1	0.68	0.69	0.61	0.75	0.68
	M2	**0.71**	0.77	0.7	0.82	0.75
	M3	**0.71**	0.73	0.62	0.82	0.72
	M4	**0.71**	**0.78**	**0.79**	**0.91**	**0.8**
MT	M1	0.63	0.66	0.7	0.79	0.7
	M2	0.67	0.68	0.73	0.82	0.73
	M3	**0.73**	0.79	0.73	0.83	0.77
	M4	0.71	**0.81**	**0.77**	**0.85**	**0.79**
MS	M1	0.62	0.72	0.6	0.81	0.69
	M2	**0.68**	0.71	0.64	0.7	0.68
	M3	0.65	0.78	0.64	0.78	0.71
	M4	0.61	**0.83**	**0.79**	**0.84**	**0.77**
MF ESR	M1	0.69	0.79	0.72	0.84	0.76
	M2	0.69	0.69	0.71	0.84	0.73
	M3	0.72	0.7	0.72	0.85	0.75
	M4	**0.78**	**0.89**	**0.81**	**0.88**	**0.84**
MT ESR	M1	0.69	0.77	0.71	0.88	0.76
	M2	0.71	0.72	0.71	0.85	0.75
	M3	**0.72**	0.71	0.72	0.98	0.78
	M4	0.69	**0.86**	**0.81**	**0.98**	**0.84**

MF: Mozilla Firefox, MT: Mozilla Thunderbird MS: Mozilla Seamonkey, M1: TF-IDF+SVM, M2: word2vec+LSTM, M3: word2vec+Bi-LSTM, M4: GloVe+Bi-LSTM

RQ2: Which word embedding method for FVR gave the best results?

To address this RQ, the values recorded in Table 1 are analyzed, and the highest values are made bold, for all levels of severity of the vulnerability as well as for the mean case. We have got a total of twenty-seven highest values, which are made bold. Out of these twenty-seven values, twenty-two are from the GloVe+Bi-LSTM combination. Hence, we can conclude that out of TF-IDF, word2vec, and GloVe, the best performance is given by GloVe in combination with Bi-LSTM.

RQ3: Whether the performance of SVSPM giving the best results, boosted significantly?

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 01,2023 at 07:28:33 UTC from IEEE Xplore. Restrictions apply.

To check whether the performance of SVSPMs using different word embedding methods for FVR boosted significantly, Friedman Test is used. The choice of the Friedman Test is made as the data we are having is not following normal distribution and we are having more than two samples to investigate [36]. The null and alternate hypotheses are framed as follows:

- Ho: The performance of SVSPMs developed using different FVR methods is the same and does not differ significantly.
- H1: The performance of SVSPMs developed using different FVR methods differs significantly.

Above stated hypothesis is tested considering the significance level (α) to be 0.05, i.e., 95% confidence level. The p-value obtained for low, medium, high, critical, and mean cases are 0.071, 0.026, 0.008, 0.011, and 0.014 respectively. This indicates that for a low level as the p-value is greater than 0.05, the result is not significant (H_o accepted), but for all other four cases including the mean case, the results are significant (H_o rejected).

Doing further analysis, the results obtained after the application of the Friedman test are recorded in Table 2. It gives the mean rank for all four levels as well as for the average case, for all the models. A large value of rank obtained by a particular model indicates that that particular model is better. Except at a low level, for all other four cases, the highest rank is obtained by the model developed using GloVe+Bi-LSTM, proving that this model performed best showing significant improvement in the performance. Hence, the finding of RQ2 is also validated mathematically using a statistical test for hypothesis testing.

TABLE II. RESULT OF THE FRIEDMAN TEST

Model	Mean Rank (Low)	Mean Rank (Medium)	Mean Rank (High)	Mean Rank (Critical)	Mean Rank (Average)
M1	1.40	2	1.4	1.7	1.8
M2	2.70	1.8	2.1	1.6	1.6
M3	3.40	2.2	2.5	2.8	2.6
M4	2.50	4	4	3.9	4

M1: TF-IDF+SVM, M2: word2vec+LSTM, M3: word2vec+Bi-LSTM, M4: GloVe+Bi-LSTM

TABLE III. RESULT OF THE WILCOXON SIGNED RANK TEST

Pair	p-value (Medium)	p-value (High)	p-value (Critical)	p-value (Mean)
M1 vs M2	0.69	0.144	0.854	0.686
M1 vs M3	0.79	0.066	0.138	0.08
M1 vs M4	**0.042**	**0.043**	**0.043**	**0.042**
M2 vs M3	0.416	1	0.066	0.176
M2 vs M4	**0.043**	**0.042**	**0.043**	**0.043**
M3 vs M4	**0.042**	**0.042**	0.068	**0.043**

M1: TF-IDF+SVM, M2: word2vec+LSTM, M3: word2vec+Bi-LSTM, M4: GloVe+Bi-LSTM

The posthoc test is also applied to check the statistical difference pairwise, hence validating our results even more strongly. Wilcoxon Signed rank test is applied for those cases in which the results of the Friedman test came out to be significant. The results obtained are documented in Table 3. Significant values are made bold. The results of eleven pairs out of twelve containing GloVe+Bi-LSTM are significant, proving that the GloVe+Bi-LSTM model outperformed others.

VI. CONCLUSION AND FINDINGS

The feature vector representation is a very important aspect of developing a prediction model which uses textual data as input. Therefore, in this paper, we compared the performance of various word embedding methods used for FVR. The methods investigated in this study are TF-IDF, word2vec, and GloVe. SVM, LSTM, and Bi-LSTM algorithms are used for the classification task. Thus, we can say that a comparison of the DL approach is made against the traditional ML approach with different word embedding methods for FVR. After doing the result analysis, it is concluded that the SVSPM developed using GloVE+Bi-LSTM gave the best performance giving an average AUC value across all databases of 0.81. The main problem associated with textual data is high dimensionality and the other is the imbalanced nature of data. In this study, the word embedding methods explored except TF-IDF, also do automatic feature selection thus helping in dimensionality reduction too. Therefore, in the future, more such methods which are existing in the literature can be explored. Experiments can be carried out and the effect on the performance of the model can be analyzed using a hybrid approach, derived from a combination of two methods of FVR. In order to tackle the imbalanced nature of data, data balancing techniques can be explored to develop an effective data balancing module for enhancing the performance of SVSPM.

REFERENCES

[1] G. McGraw, Software security: building security in, 1st ed., vol. 1, Addison-Wesley Professional, 2006.

[2] S. Furnell, "Cybercrime in society," Connected Minds, Emerging Cultures: Cybercultures in Online Learning, pp. 209, 2009.

[3] G. E. Violettas, T. L. Theodorou, and G. C. Stephanides, "E-learning software security: Tested for security vulnerabilities and issues", e-Learning, Best Practices in Management, Design and Development of e-Courses: Standards of Excellence and Creativity, pp. 233–240, 2013.

[4] Y. Younan, "25 years of vulnerabilities: 1988–2012," Sourcefire Vulnerability Research Team, 2013.

[5] R. Coulter, Q.L. Han, L. Pan, J. Zhang, and Y. Xiang, "Data-driven cyber security in perspective–intelligent traffic analysis," IEEE transactions on cybernetics, vol. 50, no. 7, pp. 3081-3093, July 2020.

[6] L. Bilge, and T. Dumitra¸s. "Before we knew it: an empirical study of zero-day attacks in the real world", In Proceedings of the 2012 ACM Conference on Computer and Communications Security, pp. 833–844, 2012.

[7] C. Microsoft, "Microsoft security response center security bulletin severity rating system," https://technet.microsoft.com/zhcn/security/gg309177.aspx, 2002, [Online; accessed 30-March-2017].

[8] US-CERT, "Uscert vulnerability note field descriptions," http://www.kb. cert.org/vuls/html/fieldhelp, 2006, [Online; accessed 30-March-2017].

[9] I. SANS, "Sans critical vulnerability analysis archive," http://www.sans. org/newsletters/cva/, [Online; accessed 30-March-2017].

[10] FIRST, "Common vulnerability scoring system (cvss) version 2.0," https: //www.first.org/cvss/v2/guide#i1.2, 2007, [Online; accessed 30-March2017].

[11] C. MITRE, "Common vulnerabilities and exposures (cve)," https://cve. mitre.org/, [Online; accessed 30-March-2017].

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 01,2023 at 07:28:33 UTC from IEEE Xplore. Restrictions apply.

[12] Q. Jiao and S. Zhang, "A Brief Survey of Word Embedding and Its Recent Development," Advanced Information Technology, Electronic and Automation Control Conference, pp. 1697-1701, 2021

[13] D.R. Kubal, and A.V. Nimkar. "A survey on word embedding techniques and semantic similarity for paraphrase identification", International Journal of Computational Systems Engineering, vol. 5, no.1, pp. 36-52, 2019.

[14] S. Liu, G. Lin, Q.L. Han, S. Wen, J. Zhang, and Y. Xiang, "DeepBalance: Deep-Learning and Fuzzy Oversampling for Vulnerability Detection," IEEE Transactions on Fuzzy Systems, vol. 28, no. 7, pp. 1329-1343, July 2020.

[15] S. Nakagawa, et al. "Character-level convolutional neural network for predicting severity of software vulnerability from vulnerability description." IEICE Transactions on Information and Systems, vol. 102, no. 9, pp. 1679-1682, 2019.

[16] F. G. Bulut, H. Altunel, and A. Tosun, "Predicting Software Vulnerabilities Using Topic Modeling with Issues," International Conference on Computer Science and Engineering, pp. 739-744, 2019.

[17] G. Huang, Y. Li, Q. Wang, J. Ren, Y. Cheng and X. Zhao, "Automatic Classification Method for Software Vulnerability Based on Deep Neural Network," IEEE Access, vol. 7, pp. 28291-28298, 2019.

[18] J. Chen, P.K. Kudjo, S. Mensah, S.A. Brown, G. Akorfu, "An automatic software vulnerability classification framework using term frequency-inverse gravity moment and feature selection", Journal of Systems and Software, vol. 167, 2020.

[19] P.K. Kudjo, and J. Chen, "A cost-effective strategy for software vulnerability prediction based on bellwether analysis", Proceedings of the 28th ACM SIGSOFT International Symposium on Software Testing and Analysis, pp. 424-427, 2019.

[20] P. Wang, Y. Zhou, B. Sun and W. Zhang, "Intelligent Prediction of Vulnerability Severity Level Based on Text Mining and XGBboost," Eleventh International Conference on Advanced Computational Intelligence, pp. 72-77, 2019

[21] M. Aota, H. Kanehara, M. Kubo, N. Murata, B. Sun and, T. Takahashi, "Automation of Vulnerability Classification from its Description using Machine Learning," IEEE Symposium on Computers and Communications (ISCC), pp. 1-7, 2020.

[22] P.K. Kudjo, J. Chen, S. Mensah, "The effect of Bellwether analysis on software vulnerability severity prediction models", Software Quality Journal, vol. 28, pp. 1413–1446, 2020.

[23] L.C. Yu, J. Wang, K.R. Lai, and X. Zhang, "Refining word embeddings for sentiment analysis", Proceedings of the Conference on Empirical Methods in Natural Language Processing, pp. 534–539, 2017.

[24] L. Li, B. Qin, and T. Liu, "Contradiction detection with contradiction-specific word embedding, Algorithms", Algorithms, vol. 10, 2017.

[25] S.E. Sahin, and A. Tosun, "A conceptual replication on predicting the severity of software vulnerabilities", Proceedings of the Evaluation and Assessment on Software Engineering, pp. 244-250, 2019.

[26] Y. Fang, Y. Liu, C. Huang, L. Liu, "FastEmbed: Predicting vulnerability exploitation possibility based on ensemble machine learning algorithm", Plos one, vol.15, 2020.

[27] R. Sharma, R. Sibal, and S. Sabharwal, "Software vulnerability prioritization using vulnerability description", International Journal of System Assurance Engineering and Management, pp. 58–64, 2021.

[28] I. Arora, and A. Saha, "Software Defect Prediction: A Comparison Between Artificial Neural Network and Support Vector Machine", Advanced Computing and Communication Technologies. Advances in Intelligent Systems and Computing, vol 562, 2018.

[29] M. Schuster and, K. K. Paliwal, "Bidirectional recurrent neural networks," IEEE Transactions on Signal Processing, vol. 45, no. 11, pp. 2673-2681, Nov. 1997.

[30] W. Yin, K. Kann, M. Yu, and H. Schütze, "Comparative study of CNN and RNN for natural language processing", arXiv:1702.01923, February 2017.

[31] K.S. Tai, R. Socher, and C.D. Manning, "Improved semantic representations from tree-structured long short-term memory networks", arXiv:1503.00075, May 2015.

[32] S. Qaiser, and R Ali, "Text mining: use of TF-IDF to examine the relevance of words to documents" International Journal of Computer Applications, vol. pp. 25-29, 2018.

[33] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient Estimation of Word Representations in Vector Space", arXiv:1301.3781 [cs], January 2013.

[34] Y. Chen, "Convolutional neural networks for sentence classification", MS thesis. University of Waterloo, 2015.

[35] H. He and, E. A. Garcia, "Learning from Imbalanced Data", IEEE Transactions on Knowledge and Data Engineering, vol. 21, no. 9, pp. 1263-1284, September 2009.

[36] D.L. Olson, and D. Delen, "Performance evaluation for predictive modelling", Advanced data mining techniques, pp. 137-147, 2008.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 01,2023 at 07:28:33 UTC from IEEE Xplore. Restrictions apply.

Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Influence of SiC microparticles and multi-pass FSW on weld quality of the AA6082 and AA5083 dissimilar joints

SUMIT JAIN (sumitjain4@gmail.com)
Delhi Technological University (DTU)
R. S. MISHRA
Delhi Technological University (DTU)
HUSAIN MEHDI
Meerut Institute of engineering and technology

Research Article

Keywords: Reinforced joints, tensile strength, %strain, microhardness, SiC microparticles

Posted Date: February 20th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2568962/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Abstract

In the current research, the influence of multi-pass (one, two and three passes) of friction stir welding and SiC microparticles on the tensile strength, microhardness and %strain of dissimilar reinforced joints of AA6082 and AA5083 was examined. Tool rotating speed of 900 rpm, transverse speed of 45 mm/min, tool tilt angle of 2º and 8% volume percentage of SiC microparticles were considered as constant parameters for multi-pass FSW. The finding demonstrated that increasing the number of FSW passes from one to three augmented the dispersion pattern of Sic microparticles. The grain refinement of multi-pass reinforced joints was achieved by the pinning effect of SiC microparticles and dynamic recrystallization. The three-passes FSWed reinforced joint exhibited the highest tensile strength (247.17 MPa), %stain (13.1%) and microhardness (126.6 HV) due to the higher grain refinement.

1. Introduction

The fabrication of metal matrix composites employing reinforcing particles (RPs) is a creative technique for augmenting the mechanical characteristics (stiffness, hardness and strength, etc.), corrosion, and wear resistance [1–3]. The distributed reinforcing particles, via the "pinning effect" inhibit the grain boundaries from expanding during dynamic recrystallization, resulting in equiaxed tiny grains that ameliorate the tribological and mechanical characteristics of welded or processed materials [4–6]. To enhance tribological and mechanical characteristics, micro/nanoparticles-based aluminum metal matrix composites through FSP were fabricated to produce surface and bulk composites [7–9]. The transformation from traditional materials to lightweight materials is recently developing in almost all industries for fabricating high-strength and lightweight structural components. Aluminum alloys offer exceptional mechanical characteristics like higher strength, lightweight, and good corrosion resistance. Therefore, aluminum alloys have attracted great interest for structural applications in almost all industrial domains, including aerospace, marine and automobile. Fusion welding methods find inappropriate for welding low melting point materials due to their limitations of high residual stresses, surface cracks, distortion and porosity [10–12]. Thus, friction stir welding (FSW) is solid-state in nature and seems to be the most effective method for joining dissimilar aluminum alloys.

Several researchers also employed FSW to produce aluminum metal matrix composites (AMMCs) utilizing various micro/nano-sized RPs to ameliorate the weld properties of related aluminum alloys. Recent studies have examined the viability of fabricating AMMCs with different reinforcing particles, including AA2124 with SiC particles, AA7005 with Al_2O_3 particles and AA2009 with SiC particles [13–15]. The incorporation of micro/nano particles into the stir zone helped the researchers improve the weld properties of numerous metal alloys. Compared to the FSWed samples without SiC particles, improved weld qualities of dissimilar AA7075/AA2024 were produced with different volume percentages (5, 8, and 13%) of nano-sized SiC particles. It was found that 5% SiC nanoparticle inclusion provided the best mechanical properties [16]. The reinforcement particles of Al_2O_3 outperformed Sic reinforcement particles and a hybrid of the two in FSW of the dissimilar AA6101 and AA1350 in terms of mechanical and wear characteristics [17].

The strain-hardenable AA5083 has exceptional corrosion resistance additionally, it has superb weldability [18]. Consequently, AA5083 is often used not only in the aerospace and automotive industries but also in shipbuilding and armored vehicles [19, 20]. AA6082 is also frequently utilized in a wide range of automotive and aerospace applications. AA6082 provides an impressive combination of corrosion resistance, strength, and surface characteristics [21]. Dissimilar welding of AA6082 and AA5083 has drawn the considerable interest of almost all the industrial sectors due to their broad applications and superior weldability. The motivation behind this study to improve weld quality of these alloys which will be tremendously beneficial in a number of industrial domains. Therefore, in the present research the consequences of multi-pass FSW and SiC microparticles (SiCp) on weld quality of AA5083 and AA6082 dissimilar joint was investigated. The consequences of multi-pass FSW and SiC particles on weld quality of AA5083 and AA6082 dissimilar joint has not yet been studied, to the author's knowledge.

2. Material And Experimentation

The materials utilized in this study to develop dissimilar FSWed joints were: AA5083 and AA6082 with dimensions 150x40x6 mm as the base and the SiC particles (SiCp) ranging from 50 to 80 µm as reinforcing material. A friction stir welding machine equipped with position controls was employed for multi-pass FSW. The Adjacent surfaces of the base plates were milled to incorporate the SiC powder. To include 8% SiCp, the depth and width of the milled slot were taken as 3 mm and 0.48 mm, respectively as depicted in Fig. 1a.. The SiCp was mixed with ethanol, and a thick slurry was made. The milled slots were cleaned with acetone before filling the SiCp slurry. The slurry of SiCp was filled manually in the milled slots. A detailed and systematic explanation of composite development utilizing the groove approach was previously published [22, 23]. The multi-pass FSW was performed at constant rotating speed, transverse speed, tool tilt angle, and volume fraction of SiCp of 900 rpm, 40 mm/min, 2°, and 8%, respectively. The position of materials The FSW parameters condition and volume fraction of SiCp were chosen based on preliminary trials to obtain defect-free FSWed reinforced joints. Prior to multi-pass FSW, a pin-less tool of shoulder diameter 19.5 mm was employed (Fig. 1b) to close the groove from top at a constant rotating speed and transverse speed of 900 rpm and 40 mm/min respectively .The parametric setting of FSW and volume percentage of SiCp were chosen based on machine capability and trail runs. The profiles of tool pin significantly impact the flow of material during FSW. The cylindrical threaded-pin tool revealed defect-free FSWed joints with better material flow than the cylindrical pin tool and cylindrical tapered-pin tool [24]. Therefore, in the current research, multi-pass FSW was performed utilizing a cylindrical threaded-pin tool made of H13 tool steel. Figure 1b illustrates the dimensional characteristics and pictorial view of the FSW tool. In the case of welding dissimilar materials, positioning of materials is an essential consideration for defect-free dissimilar joints. Therefore, in the current research, plates of AA6082 were situated on the advancing side, as suggested by previous studies [25, 26]. Table 1 illustrates the parametric setting of multi-pass FSW and the reinforcing condition of SiCp. The defect free FSWed joints after multi-pass FSW are depicted in Fig. 2.

The test sample extraction plan for various tests is cited in Fig. 3. A universal testing machine (Model: UT-04-0100, BISS Pvt. Ltd., India) of a capacity of 100 kN was employed to perform tensile tests.

Microhardness indentations were performed under 100gm load and separated by a 1 mm distance utilizing a Micro-Vickers hardness tester (Model: XHVT1000Z/V, Jinan victory inst. Co. Ltd.). To analyze the consequences of SiC particles and multi-pass FSW, samples were cross-sectioned and prepared for microstructural evolution. The etching of prepared samples was performed with Keller reagent. Optical-microscopy (OM) and Field-emission-scanning-electron-microscopy (FESEM) were utilized to evaluate the microstructural behavior of the stir zone. OM was performed utilizing a metallurgical inverted microscope (Model: Suxma-met, Conation Technologies), while FESEM analysis was performed using FESEM (Model: JSM-IT100, Jeol) equipped with an energy-dispersive X-ray spectroscopy (EDS) detector (Model: 51-ADD0076, Oxford Instrument).

3. Results And Discussion

The grain structure, FESEM photo-image and elemental composition obtained using EDS of AA5083 and AA6082 are cited in Fig. 4 (a-d). The mean grain size of AA5083 and AA6082 was found to be ~ 35 and ~ 39 µm, respectively. The FESEM photo-image and elemental composition obtained using EDS of SiCp are cited in Fig. 4 (e, f). The Due to heat input and dynamic recrystallization (DRX) resulting from severe plastic deformation, three distinct zones are identified in the weld zone of the FSWed joints: the stir zone (SZ), the thermo-mechanical affected zone (TMAZ), and heat-affected zone (HAZ) [27, 28]. Compared to the HAZ and TMAZ, the SZ exhibited higher grain refinement [29].

Table 1
Experimental condition of various FSWed joints

FSWed joints	Tool rotating speed (rpm)	Transverse speed (mm/min)	Tool angle (degree)	Number of FSW passes	Volume fraction of SiCmp (%)
One-pass FSWed unreinforced joint	900	40	2	1	0
One -pass FSWed reinforced joint	900	40	2	1	8
Two-passes FSWed reinforced joint	900	40	2	2	8
Three-passes FSWed reinforced joint	900	40	2	3	8

3.1 Microstructural characterization

To examine the influences of multi-pass FSW and SiC particles on the weld quality of AA6082 and AA5083 dissimilar joints, an investigation associated with the distribution of SiCp and grain structure of SZ was carried out. Figure 5 depicts the FESEM photo-image of the SZ of one-pass FSWed unreinforced joint and multi-pass (one, two and three) FSWed reinforced joints. The defect-free joining of AA6082 and

AA5083 with the proper intermixing was observed in the one-pass FSWed unreinforced joint, as depicted in Fig. 5a. Figure 5 (b-d) depicts the FESEM photo-image of the SZ of one, two and three-passes FSWed reinforced joints and confirms the consequences of the increasing FSW passes on the dispersion pattern of particles.

The clustering of SiCp along with poor bonding with AMM was noticed in the SZ of the one-pass FSWed reinforced joint as delineated in Fig. 5b. This demonstrates that one FSW pass was not producing sufficient strain to prevent clustering of SiCp [30]. The clustering of SiCp was found absent as the FSW passes enhanced from one to two, but the dispersion pattern of SiCp was observed as non-uniform in the two-passes FSWed reinforced joint, as cited in Fig. 5c [31]. The augmented dispersion pattern of SiCp may be attributed to the repeated strain developed by the tool's stirring action resulting in disjoining of clustering of SiCp as the FSW passes enhanced from one to two. Further improved dispersion of SiCp was found after implementing three passes of FSW. Therefore the uniform dispersion of SiCmp was observed after implementing three FSW passes, as cited in Fig. 5d. The absence of SiCp clustering revealed the excellent bonding between SiCp and surrounded AMM [32, 33]. The uniformity of SiCmp in the SZ of three-passes FSWed reinforced joints was confirmed by the EDS mapping analysis. It is evident from the Si elemental mapping (Fig. 5e) that the dispersion pattern of SiCmp is uniform after three passes of FSW. Referring to Fig. 6 (a-c), The EDS analysis of FSWed unreinforced joint and reinforced joints after one and three FSW passes also confirms the incorporation of SiCmp in FSWed reinforced joints. Considering the EDS analysis finding, the silicon percentage in the one-pass FSWed unreinforced joint was 0.56%, respectively. In contrast, the silicon percentage in the one-pass and three-passes FSWed reinforced joint was 6.98% and 7.66%, respectively. The carbon content in the one-pass and three-passes reinforced FSWed joint was 2.02% and 2.83%, respectively.

Figure 7 depicts the grain structure of the SZ of FSWed unreinforced joint and various reinforced joints. The stretched grain structure of AA6082 and AA5083 transformed into equiaxed and fined grains due to DRX [34]. The grain size of the SZ is mainly influenced by the temperature effect, DRX and RPs in the AMM. The dominating factors among these three phenomena decide the grain size. The temperature effect arises due to increased heat input at lower transverse speeds and higher rotating speeds, which induces grain growth and coarsening of grains [35]. DRX occurs at high temperature owing to severe plastic deformation resulting in the conversion of the low-angle boundaries to the high-angle boundaries and nucleates the new grains at developmental regions, which diminishes the grain size [36]. The RPs in the SZ serve as grain boundary barriers, inhibiting grain development via Zener-pinning effect [37]. Consequently, at constant rotating speed (900 rpm) and transverse speed (45 mm/min), the phenomena of DRX and the influence of SiCp dominated in diminishing the grain size. Whereas, Due to the absence of SiCp in FSWed unreinforced joint, only the DRX phenomenon dominated in diminishing the grain size. Thus, higher grain refinement was achieved in the three-passes FSW reinforced joints due to the DRX and the effect of uniformly distributed SiCp [38].

The coarse gain structure of AA 6082 and AA5083 exhibits the mean grain size of 39 and 35 μm, respectively. An abrupt reduction in the grain size was observed after implementing one-pass of FSW

owing to DRX [39]. Therefore, a mean grain size of 19.4 μm (Fig. 7a) was found in the one-pass FSWed unreinforced joint, as delineated in Fig. 7a. The grain size of the one-pass FSWed reinforced joint was found to be smaller (12.6 μm) than one-pass FSWed unreinforced joint, as cited in Figs. 7a and b. This may be ascribed to the presence of SiCp in the SZ, as the same rotating speed and transverse speed were used for unreinforced and reinforced joints. As the FSW passes proceeded from one to three, it was observed that the grain size of SZ was further diminished. The further reduction in the grain size may be ascribed to the dispersion pattern of SiCp [40]. The more uniformly distributed RPs decreased the grain size by providing more obstacles to grain boundaries and restricting grain growth by the pinning effect [41]. When dislocations begin to accumulate during plastic deformation owing to RPs, the particles-induced nucleation-based DRX is feasible [42, 43]. The presence of more uniformly dispersed SiCp enhances the pinning effect, which results in higher grain refinement. Consequently, the grain size of the two-passes and three-passes FSWed reinforced joints was found to be 7.2 (Fig. 7c) and 3.4 μm (Fig. 7d), respectively. Thus, the grain structure of multi-pass FSWed reinforced joints evident that equiaxed and fine grains were developed owing to the domination of DRX and the pinning effect of uniformly dispersed SiCp [43].

3.2 Influence on tensile properties

In Fig. 8, the average tensile characteristics of the base materials and the various FSWed joints was presented. In the case of unreinforced joints, grain size is the key variable in influencing the mechanical characteristics of the FSWed joint. However, the size, quantity, dispersion pattern of RPs, and the bonding strength between the RPs and AMM also influence the mechanical characteristics of the FSWed reinforced joints [44]. It can be revealed from Fig. 8 that the tensile strength of the FSWed unreinforced and reinforced joints are lesser than that of the base materials. In contrast, the one-pass FSWed reinforced joint exhibits higher tensile strength than that of the unreinforced joint. This is ascribed to the existence of SiCp, which inhibits dislocation boundary movement and prevents grain development, resulting in reduced gain size [45]. The bonding between AMM and RPs, grain size, and dislocation density all significantly influenced the tensile strength of the FSWed reinforced joint [46]. The improved dispersion of SiCp was achieved by the increase in FSW passes. More uniform dispersion of RPs developed more barriers to grain development and further diminished grain size [47]. The tensile strength is inversely correlated with grain size, as stated by the Hall-Patch equation [48]. The grain size of FSWed reinforced joints was observed to be smaller than that of the base materials due to the pinning effect of SiCp and DRX [49]. Consequently, the three-passes FSWed reinforced joint exhibited a higher tensile strength of 247.17 MPa. Whereas the one-pass FSWed unreinforced joint exhibited the minimum tensile strength of 206.83 MPa. These observations are consistent with those of Jamalian et al. [50]. The %strain of one-pass FSWed unreinforced joint and reinforced joint was found to be higher than AA6082 but lower than AA5083. The %strain of the one-pass FSWed reinforced joint was observed lower than that of unreinforced joints and both the base alloys. But the %strain of two and three-passes FSWed reinforced joints was 12.2, and 13.1%, respectively, which is higher than the base materials.

Figure 9(a-d) depicts the fractography of the fractured tensile samples of various FSWed joints. The fracture occurred towards the AA6082 side at HAZ, where the hardness value was observed as minimal [51]. The dimple fracture, symptomatic of the ductile approach of failure, was noticed in all tensile samples. Under the tensile stress, the test samples' edges formed a shear plane with a cup-cone shape [52]. In the FSWed reinforced joints, the ductile fracture with honeycomb dimples was identified (Fig. 9). The smaller dimple size was observed in the one-pass FSWed reinforced joint (Fig. 9b) compared to the one-pass FSWed unreinforced joint. From Fig. 9 (c, d), it can be noticed that the dimple size was further decreased as the FSW passes were enhanced from one to three. The diminished grain size was found in the three-passes FSWed reinforced joint via pinning effect of uniformly dispersed SiCp. The fractography of raptured surface (Fig. 9d) also revealed fine and equiaxed dimples developed by the micro voids' coalescence [53, 54]. The ruptured zone was found at the HAZ towards AA6082, confirmed by hardness variation, which led to augmented %strain. Thus, it can be concluded that the embedded SiCmp ameliorated the tensile chrematistics of the FSWed joints. This may be correlated to the enhanced grain refinement by DRX and pinning effect of uniformly dispersed SiCp and excellent bonding between the SiCp and the AMM, resulting in higher resistance to fracture [55].

3.3 Influence on microhardness

The microhardness distribution in the weld zone of several FSWed joints of AA5083 and AA6082 was evaluated using a Vickers microhardness tester. The asymmetrical microhardness variation was noted in the weld zone of FSWed joints due to the irregular plastic flow in the retreating and advancing side [56]. The mean microhardness of base materials (AA6082 and AA5083) and at the center of the weld zone of FSWed joints is cited in Fig. 10a.

The one-pass FSWed unreinforced joint exhibited a microhardness of 93.2 HV, which is lower than that of AA5083 but higher than AA6082. According to the microhardness findings, one-pass FSWed unreinforced joint exhibits the lowest hardness compared to the multi-pass FSWed reinforced joints. The microhardness was enhanced by incorporating SiCp and implementing one, two and three passes of FSW. Therefore, the microhardness of one, two and three passes FSWed joints was observed as 105.8, 113.1 and 126.6.4 HV, respectively. The three-passes FSWed reinforced joint has the maximum microhardness of 126.6 HV among all the FSWed joints. The distribution of microhardness in the weld zone of the various FSWed joints is cited in Fig. 10b. The SZ indicates the higher microhardness owing to smaller grain size caused by the pinning effect of SiCp and DRX compared to the TMAZ and HAZ. In contrast, the microhardness in the HAZ was noticed lesser owing to coarsening of grains at high temperature and over-aging [57]. The microhardness of the FSWed reinforced joints was also augmented owing to the hard nature of SiCp. The three-passes FSWed reinforced joint revealed the highest microhardness due to better grain refinement than other FSWed joints.

4. Conclusions

In the current work, dissimilar FSWed reinforced joints of AA5083 and AA6082 were developed using 8% volume fraction of SiCp. The impact of SiC microparticles along with multi-pass FSW on the weld quality of AA5083 and AA6082 dissimilar FSWed joints was investigated. The following are the main finding from the obtained results:

1. In one-pass FSWed unreinforced joint, only DRX was involved in the grain refining. In contrast, the improved grain refinement in the multi-pass FSWed reinforced joints was accomplished by the pinning effect of SiCp and DRX.
2. The improved dispersion of SiCmp was found as the FSW passes enhanced from one to three. The uniformly dispersed SiCp was observed in the three-passes FSWed reinforced joint.
3. The multi-pass FSWed reinforced joints revealed a higher tensile strength than the unreinforced joint. Furthermore, the three-passes FSWed reinforced joint exhibited a higher tensile strength of 247.17 MPa.
4. The %strain of one-pass FSWed unreinforced joint and reinforced joint was found to be higher than AA6082 but lower than AA5083. The %strain of one-pass FSWed reinforced joint was to be found lower than unreinforced joints and both the base alloys. But the %strain of two and three-passes FSWed reinforced joints was 19.2, and 21.4%, respectively, which is higher than that of both the base materials.
5. The one-pass FSWed unreinforced joint exhibited a microhardness of 93.2 HV, which is lower than that of AA5083 and higher than AA5083. The microhardness was enhanced by incorporating SiCp and implementing one, two and three passes of FSW. Therefore, the microhardness of one, two and three passes FSWed joints was observed as 105.8, 113.1 and 129.6.4 HV, respectively.

Declarations

Compliance with Ethical Standards : No

Competing Interests: No competing interests.

Consent to Participate: No objection.

Consent for Publication: No objection.

Informed consent: Not applicable.

Conflict-of-Interest: No

Availability of Data and Material: All the data and information mentioned in this article are related to this research.

Author's Contribution: Sumit Jain and R.S. Mishra: Writing and reviewing of article.

Husain Mehdi: Drafting and reviewing.

Funding statement: Not applicable

Acknowledgements: The authors would like to express gratitude to Department at the Mechanical Engineering in DTU, Delhi.

References

1. Hartley P, Hartley DE (2002) “FSW Implementation on the Space Shuttle’s External Tank,” Friction Stir Welding Technology for Defense Applications, New Orleans. LA, pp 14–15
2. Chawla KK (2012) Composite Materials: Science and Engineering. Springer Science & Business Media, Heidelberg, NY
3. Kaczmar J, Pietrzak K, Wlosinski W (2000) The production and application of metal matrix composite materials. J Mater Process Technol 106(1):58–67
4. Alidokht SA, Abdollah-Zadeh A, Soleymani S et al (2011) Microstructure and tribological performance of an aluminium alloy based hybrid compositeproduced by friction stir processing. Mater Des 32(5):2727–2733
5. Faraji G, Asadi P (2011) Characterization of AZ91/alumina nanocomposite produced by FSP. Mater Sci Eng A 527:2320–2330 [6] Barmouz M, Besharati Givi MK. Fabricat
6. Barmouz M, Asadi P, Besharati Givi MK et al (2011) Investigation ofmechanical properties of Cu/SiC composite fabricated by FSP: effect of SiCparticles’ size and volume fraction. Mater Sci Eng A 528(3):1740–1749
7. Tonelli L, Morri A, Toschi S et al (2019) Effect of FSP parameters and tool geometry on microstructure, hardness, and wear properties of AA7075 with and without reinforcing B C ceramic particles. Int J Adv Manuf Technol 102(9–12):3945–3917
8. Deore HA, Mishra J, Rao AG et al (2019) Effect of filler material and post-process ageing treatment on microstructure, mechanical properties and wear behaviour of friction stir processed AA 7075 surface composites. Surf Coat Technol 374:52–64
9. Khodabakhshi F, Nosko M, Gerlich AP (2018) Effect of graphene nano-platelets (GNPs)on the microstructural characteristics and textural development of an Al-Mg alloy during friction-stir processing. Surf Coat Technol 335:288–305
10. Ahmad Fauzi MN, Uday MB, Zuhailawati, Ismail H (2010) Microstructure and mechanical properties of alumina-6061aluminum alloy joined by friction welding. Mater Des 31:670–676
11. Kuryntsev SA, Review A, Ti A (2022) /Cu)—Methods and Techniques, Microstructure and Properties. Materials. ; 15(1):122. https://doi.org/10.3390/ma15010122
12. Bohm M, Kowalski M (2020) Fatigue life estimation of explosive cladded transition joints with the use of the spectral method for the case of a random sea state. Mar Struct 71:102739. https://doi.org/10.1016/j.marstruc.2020.102739

13. Uzun H (2007) Friction stir welding of SiC particulate reinforced AA2124 aluminium alloy matrix composite. Mater Des 28(5):1440–1446
14. Ceschini L, Boromei I, Minak G et al (2007) Effect of friction stir welding on microstructure, tensile and fatigue properties of the AA7005/10 vol.%Al2O3 composite. Compos Sci Technol 67(3–4):605–615
15. Feng AH, Xiao BL, Ma ZY (2008) Effect of microstructural evolution on mechanical properties of friction stir welded AA2009/SiC composite. Compos Sci Technol 68(9):2141–2148
16. Kumar KS, Murigendrappa SM, kumar H (2019) Experimental investigation on effects of varying volume fractions of SiC nanoparticle reinforcement on microstructure and mechanical properties in friction stir welded dissimilar joints of AA2024-T351 and AA7075-T651.Journal of Materials Research;455
17. Ashok Kumar R, Ragav GR, Nagarajan KJ et al (2019) Effect of hybrid reinforcement at stirred zone of dissimilar aluminum alloys during friction stir welding. Metall Res Technol 116:631
18. Mirihanage W, Munasinghe N (2004) : Modification of AA 5083 weld joint characteristics, International Symposium of Research Students on Materials Science and Engineering, Chennai, India, ISRS pp. 1–6
19. Peasura P, Watanapa A (2012) Influence of shielding gas on aluminum alloy 5083 in gas tungsten arc welding. Procedia Eng 29:2465–2469
20. Cetinel H, Manisa MA, Microstructure and Mechanical Properties of AA 5083 and AA 6061 Welds Joined with AlSi5 and AlSi12 Wires*, 3rd International Conference on Welding Technologies and Exhibition (ICWET'14) in Manisa, Turkey
21. Erdoğan AA, Feyzullahoğlu E, Fidan S, Sinmazçelik T (2020) Investigation of erosive wear behaviors of AA6082-T6 aluminum alloy. Proceedings of the Institution of Mechanical Engineers, Part L: Journal of Materials: Design and Applications. ;234(3):520–530. doi:10.1177/1464420719899686
22. Rathee S, Maheshwari S, Siddiquee AN et al (2017) Effect of tool plunge depth on reinforcement particles distribution in surface composite fabrication via friction stir processing. Def Technol 13(2):86–91
23. Rathee S, Maheshwari S, Siddiquee AN et al (2017) Analysis of microstructural changes in enhancement of surface properties in sheet forming of Al alloys via friction stir processing. Mater Today: Proc. ;4(2):452–458
24. Ilangovan M, Rajendra Boopathy S, Balasubramanian V (2015) Effect of tool pin profile on microstructure and tensile properties of friction stir welded dissimilar AA 6061–AA 5086 aluminium alloy joints. Def Technol 11(2):174–184. 10.1016/j.dt.2015.01.004
25. Donatus U, Thompson GE, Zhou X, Wang J, Beamish K (2015) Flow patterns in friction stir welds of AA5083 and AA6082 alloys. Mater Design 83:203–213. 10.1016/j.matdes.2015.06.006
26. Velaphi Msomi;Vuyani Moni (2021) The influence of materials positioning on microstructure and mechanical properties of friction stir welded AA5083/AA6082 dissimilar joint. Adv Mater Process Technol. (), –. doi:10.1080/2374068x.2021.1878732

27. Husain Mehdi, Mishra RS (2021) Effect of Friction Stir Processing on Mechanical Properties and Wear Resistance of Tungsten Inert Gas Welded Joint of Dissimilar Aluminum Alloys. J of Materi Eng and Perform 30:1926–1937
28. Mishra HM, R.S., Technologies (2020)1–23. https://doi.org/10.1080/2374068X.2020.1829952.
29. Srinivasu R, Rao AS, Reddy GM, Rao KS (2015) Friction stir surfacing of cast A356 aluminum silicon alloy with boron carbide and molybdenum disulphide powders. Def Technol 11:140–146
30. Sumit Jain RS, Mishra (2021) Effect of Al2O3 nanoparticles on microstructure and mechanical properties of friction stir-welded dissimilar aluminum alloys AA7075-T6 and AA6061-T6, https://doi.org/10.1177/09544089211065534
31. Mohammadzadeh Jamalian H, Ramezani H, Ghobadi H, Ansari M, Yari S, Givi B, Mohammad Kazem (2016) Processing–structure–property correlation in nano-SiC-reinforced friction stir welded aluminum joints. J Manuf Process 21:180–189. 10.1016/j.jmapro.2015.12.008
32. Asadi P, Faraji G, Besharati Givi MK (2010) Producing of AZ91/SiC composite by Friction stir processing (FSP).IntJAdvManufTechnol;51:247–60
33. Husain Mehdi RS, Mishra (2020) Influence of Friction Stir Processing on Weld Temperature Distribution and Mechanical Properties of TIG-Welded Joint of AA6061 and AA7075. Trans Indian Inst Met 73:1773–1788. https://doi.org/10.1007/s12666-020-01994-w
34. Srivastava M, Rathee S, Siddiquee AN et al (2019) Investigation on the Effects of Silicon Carbide and Cooling Medium during Multi-Pass FSP of Al-Mg/ SiC Surface Composites. Silicon 11:2149–2157
35. Humphreys FJ, Hatherly M (2004) Chapter 11 - grain growth following recrystallization, in: Recrystallization and Related Annealing Phenomena.Elsevier Ltd. ;333–378
36. Husain Mehdi RS, Mishra (2021) Effect of friction stir processing on mechanical properties and heat transfer of TIG welded joint of AA6061 and AA7075. Def Technol 17(3):715–727
37. Husain Mehdi RS, Mishra (2019) Study of the influence of friction stir processing on tungsten inert gas welding of different aluminum alloy. SN Appl Sci 1:712. https://doi.org/10.1007/s42452-019-0712-0
38. Sun YF, Fujii H (2011) The effect of SiC particles on the microstructure and mechanical properties of friction stir welded pure copper joints. Mater Sci Eng A 528:5470–5475
39. Srivatsan T, Narendra N, Troxell J (2000) Tensile deformation and fracture behavior of an oxide dispersion strengthened copper alloy. Mater Des 21:191–198
40. Salah AN, Mabuwa S, Mehdi H, Msomi V, Kaddami M, Mohapatra P (2022) Effect of Multipass FSP on Si-rich TIG Welded Joint of Dissimilar Aluminum Alloys AA8011-H14 and AA5083-H321: EBSD and Microstructural Evolutions. Silicon https://doi.org/10.1007/s12633-022-01717-4
41. Sharma DK, Patel V, Badheka V et al (2019) Fabrication of hybrid surface composites AA6061/(B 4 C + MoS) via friction stir processing. J Tribology 141:52201–52210
42. Paidar M, Sarab ML (2016) Friction stir spot welding of 2024-T3 aluminum alloy with SiC nanoparticles. J Mech Sci Technol 30:365–370. https://doi.org/10.1007/s12206-015-1241-4

43. Paidar M, Asgari A, Ojo OO et al (2018) Mechanical Properties and Wear Behavior of AA5182/WC Nanocomposite Fabricated by Friction Stir Welding at Different Tool Traverse Speeds. J of Materi Eng and Perform 27:1714–1724. https://doi.org/10.1007/s11665-018-3297-7
44. Lloyd DJ (1994) Particle reinforced aluminum and magnesium matrix composites. Int Mater Rev 39:1
45. Yadav D, Bauri R (2011) Processing, microstructure and mechanical properties of nickel particles embedded aluminium matrix composite. Mater Sci Eng A 528:1326–1333
46. Rani P, Mishra R (2022) Influence of Reinforcement with Multi-Pass FSW on the Mechanical and Microstructural Behavior of Dissimilar Weld Joint of AA5083 and AA6061. Silicon 14:11219–11233. https://doi.org/10.1007/s12633-022-01863-9
47. Kumar KS, Murigendrappa SM, kumar H (2019) Experimental investigation on effects of varying volume fractions of SiC nanoparticle reinforcement on microstructure and mechanical properties in friction stir welded dissimilar joints of AA2024-T351 and AA7075-T651.Journal of Materials Research;455
48. Paidar M, Bokov M, Mehrez S et al (2021) Improvement of mechanical and wear behavior by the development of a new tool for the friction stir processing of Mg/B4C composite. Surf Coat Technol 426:127797
49. Chang CI, Lee CJ, Huang JC (2004) Relationship between grain size and Zener–Holloman parameter during friction stir processing in AZ31 Mg alloys. Scripta Mater 51:509–514
50. Jamalian HM, Ramezani H, Ghobadi H, Ansari M, Yari S, Givi MKB (2016) Processing–structure–property correlation in nano-SiC-reinforced friction stir welded aluminum joints. J Manuf Process 21:180–189
51. Azizieh M, Kokabi AH, Abachi P (2011) Effect of rotational speed and probe profile on microstructure and hardness of AZ31/Al2O3 nanocomposites fabricated by friction stir processing. Mater Design 32(4):2034–2041
52. Husain Mehdi, Mishra RS (2020) Influence of Friction Stir Processing on Weld Temperature Distribution and Mechanical Properties of TIG-Welded Joint of AA6061 and AA7075. Trans Indian Inst Met 73:1773–1788
53. Salehi M, Saadatmand M, Aghazadeh mohandesi J (2012) Optimization of process parameters for producing AA6061/SiC nanocomposites by friction stir processing. Trans Nonferrous Met Soc China 22(5):1055–1063
54. Pandey C, Saini N, Mahapatra MM, Kumar P (2017) Study of the fracture surface morphology of impact and tensile tested cast and forged (C&F) Grade 91 steel at room temperature for different heat treatment regimes. Eng Fail Analy 71:p131–147
55. Magdy M, El-Rayes; Ehab A, El-Danaf (2012) The influence of multi-pass friction stir processing on the microstructural and mechanical properties of Aluminum Alloy 6082., 212 (5),1157–1168
56. McNelley TR, Swaminathan S, Su JQ (2008) Recrystallization mechanisms during friction stir welding/processing of aluminum alloys. Scr Mater 58:349–354

57. Husain Mehdi RS, Mishra (2020) Investigation of mechanical properties and heat transfer of welded joint of AA6061 and AA7075 using TIG + FSP welding approach. J Adv Join Processes 1:100003

Figures

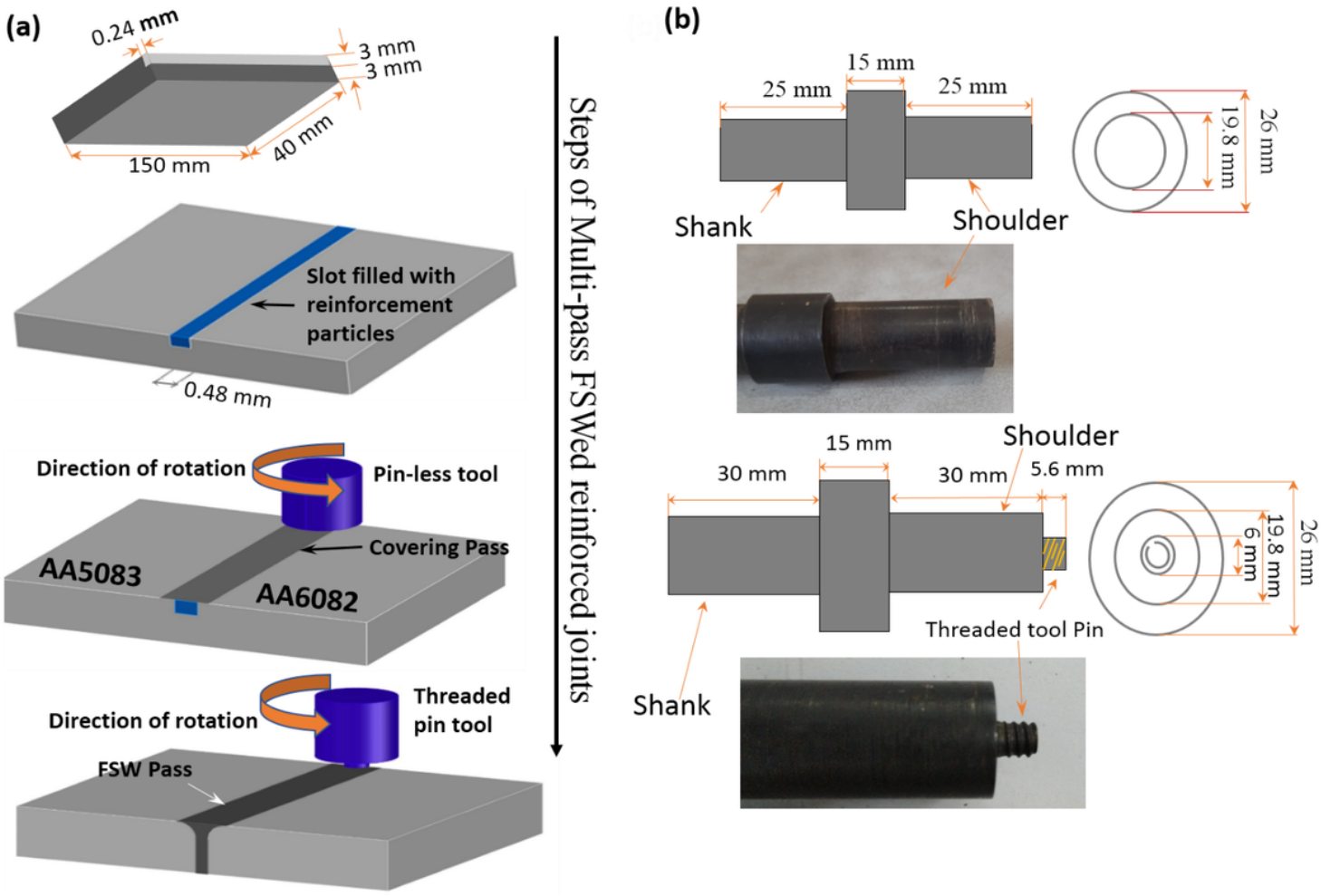


Figure 1

a) Dimensions Slot milled on the base plate; (b) Experimental set-up; (c), FSW tool's Dimensions; (d) Dimensional features of test samples.

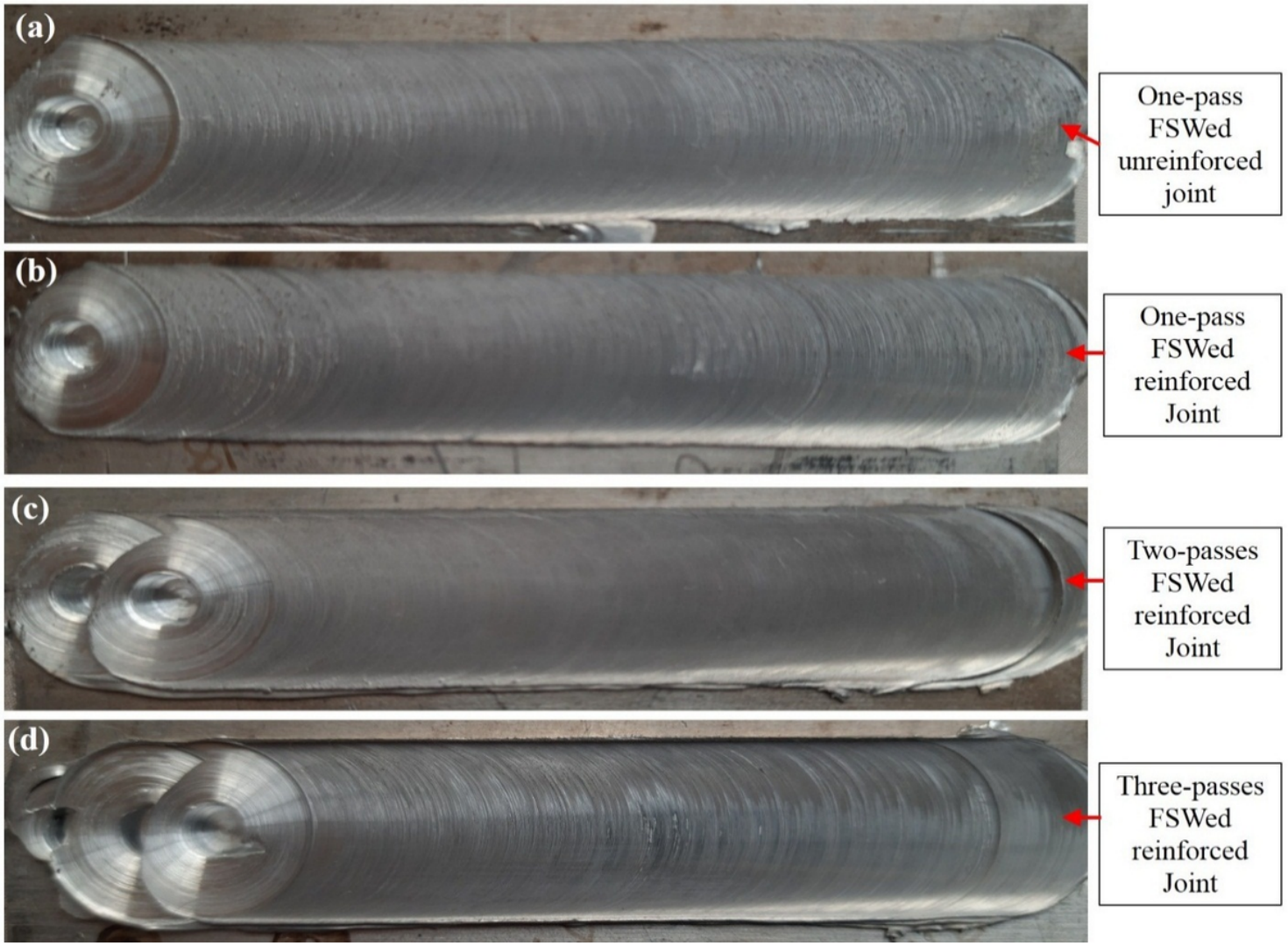


Figure 2

Various FSWed plates of AA5083 and AA6082

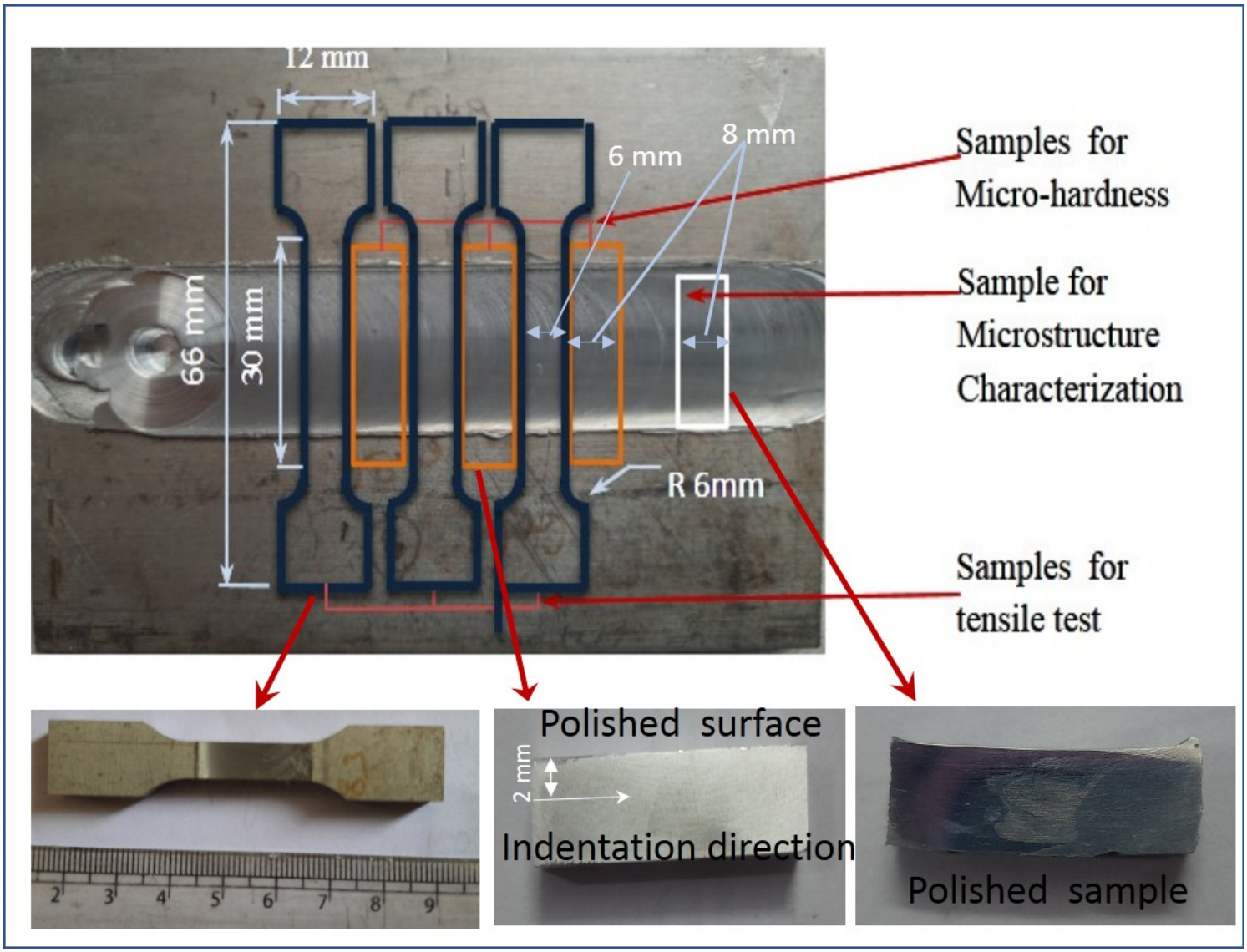


Figure 3

Dimensional features of test samples.

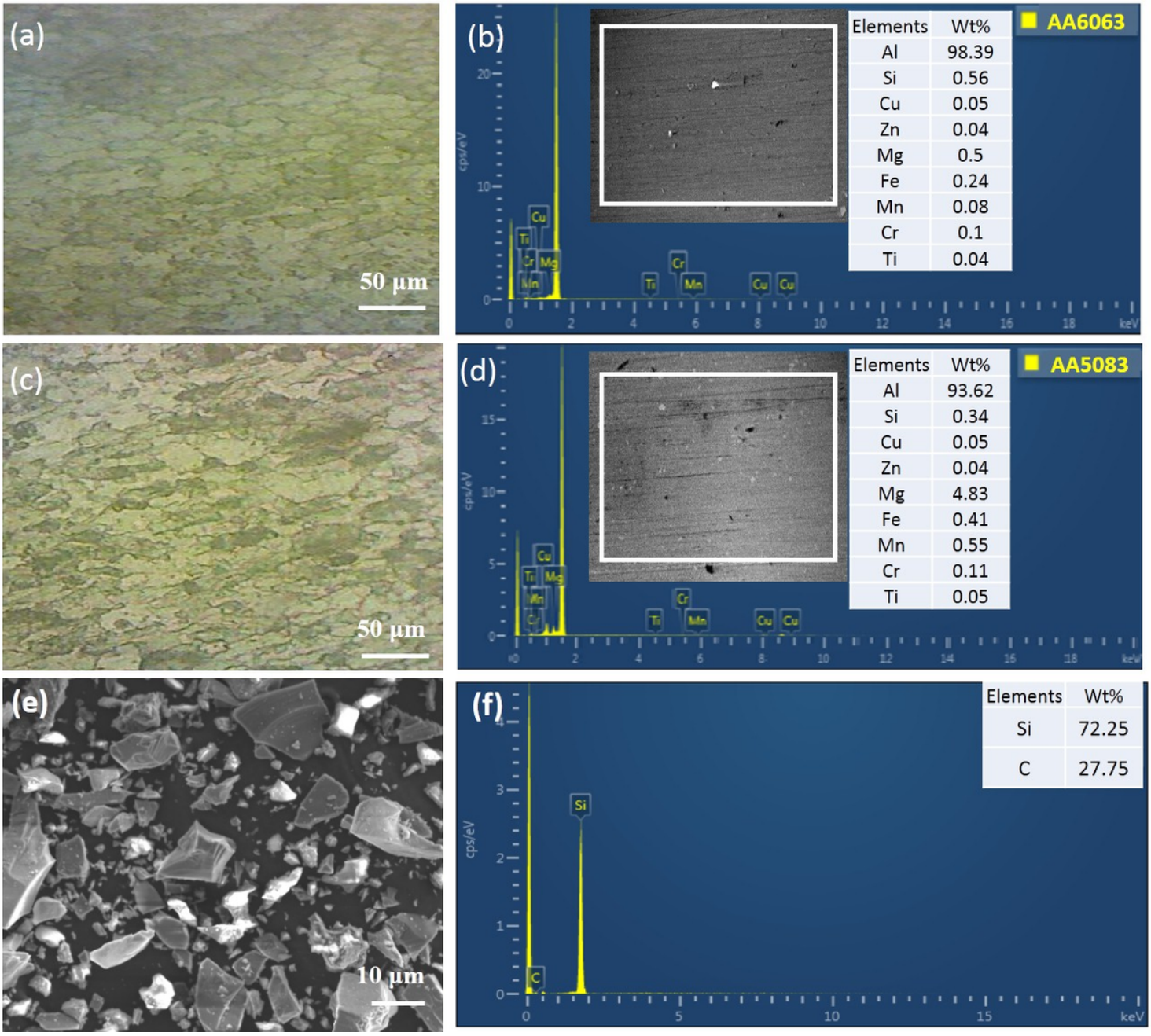


Figure 4

(a-d) Grain structure, FESEM photo-image and EDS elemental configuration of base materials; (a, b) AA6082, (c, d) AA5083; (e) FESEM photo-image of SiCp, (f) EDS elemental configuration of SiCp

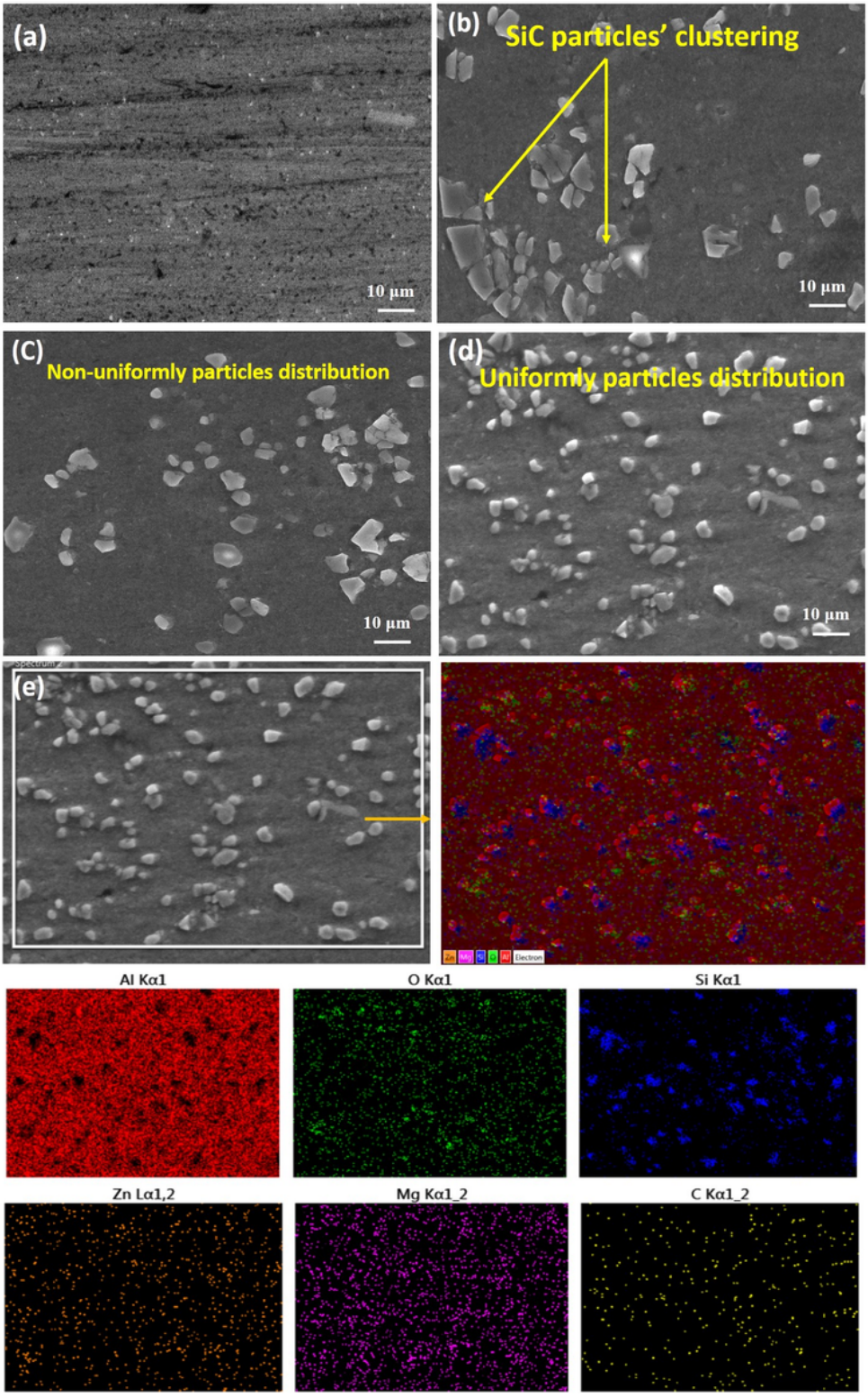


Figure 5

(a) FESEM photo-images of one-pass FSWed unreinforced joint; (b-d) FESEM photo-images of FSWed reinforced joint with; (b) one-pass, (c) Two-passes, (d) Three-passes; (e) Elemental (EDS) mapping of three-passes FSWed reinforced Joint

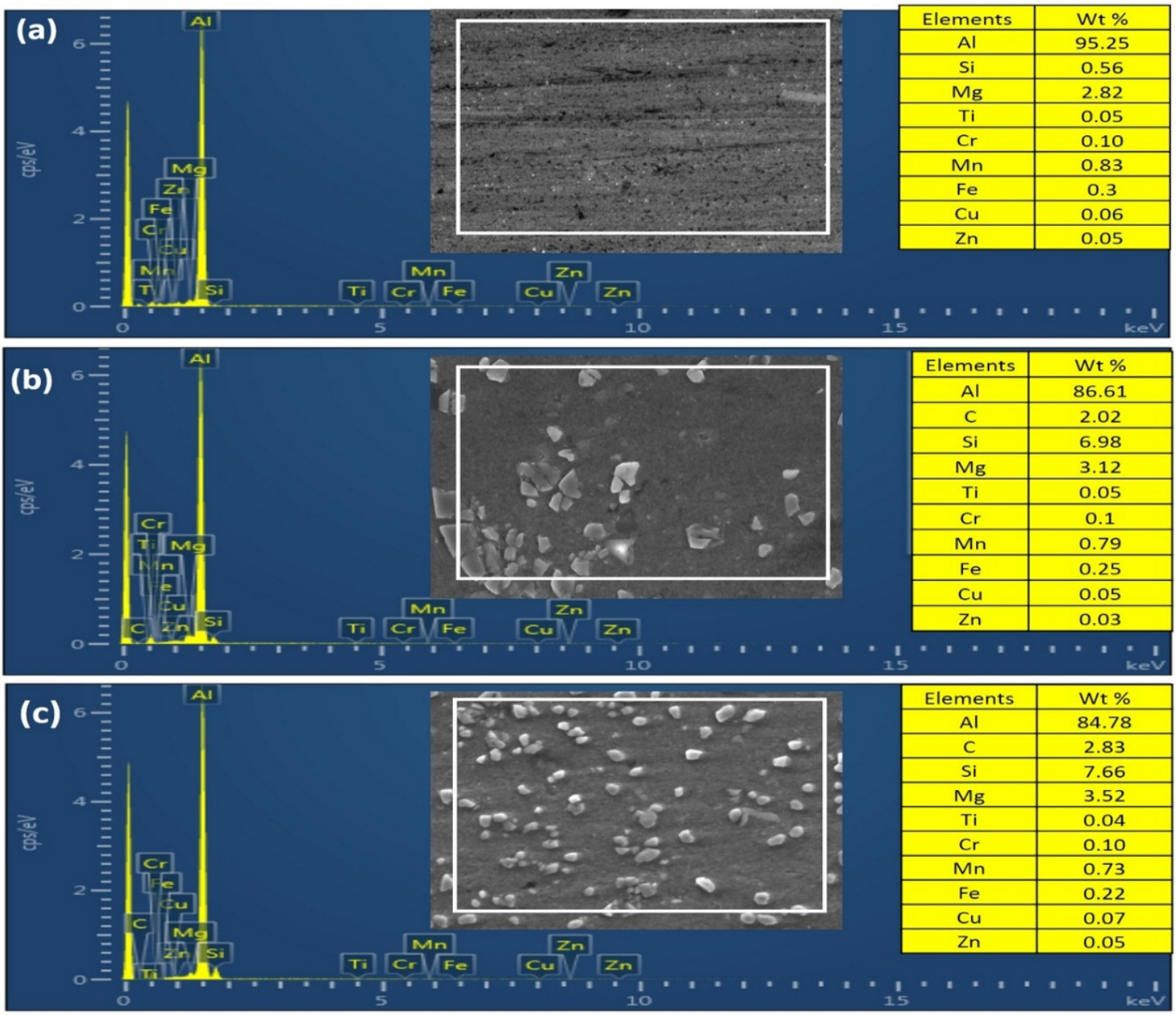


Figure 6

(a) EDS analysis of one-pass FSWed unreinforced joint, (b, c) EDS analysis of FSWed reinforced joints with; (b) One-pass (c) Three-passes.

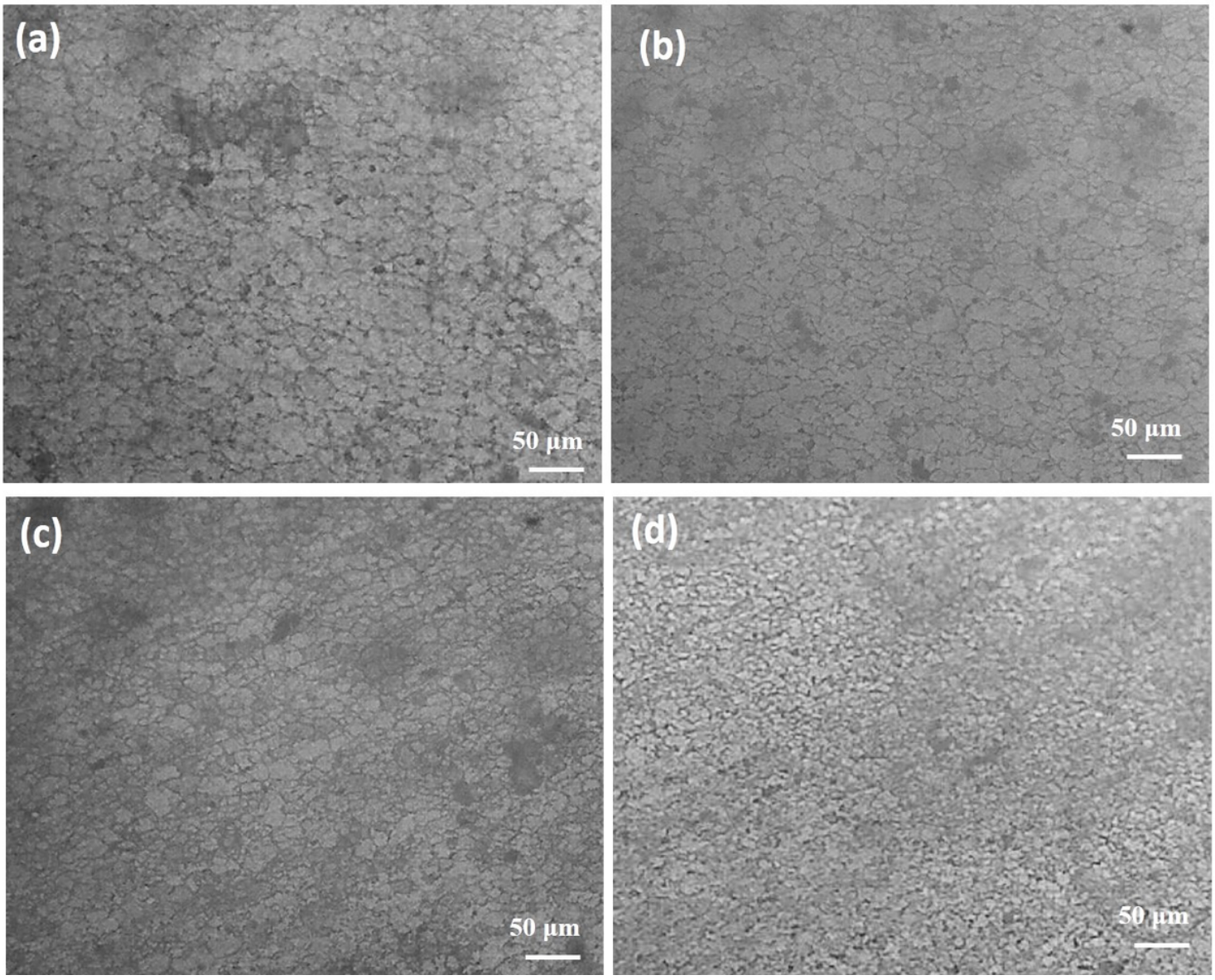


Figure 7

(a-d) Optical micrograph of the SZ; (a) Grain structure of one-pass FSWed unreinforced joint, (b-d) Grain structure of FSWed reinforced joints with, (b) one-pass, (c) Two-passes, (d) Three-passes.

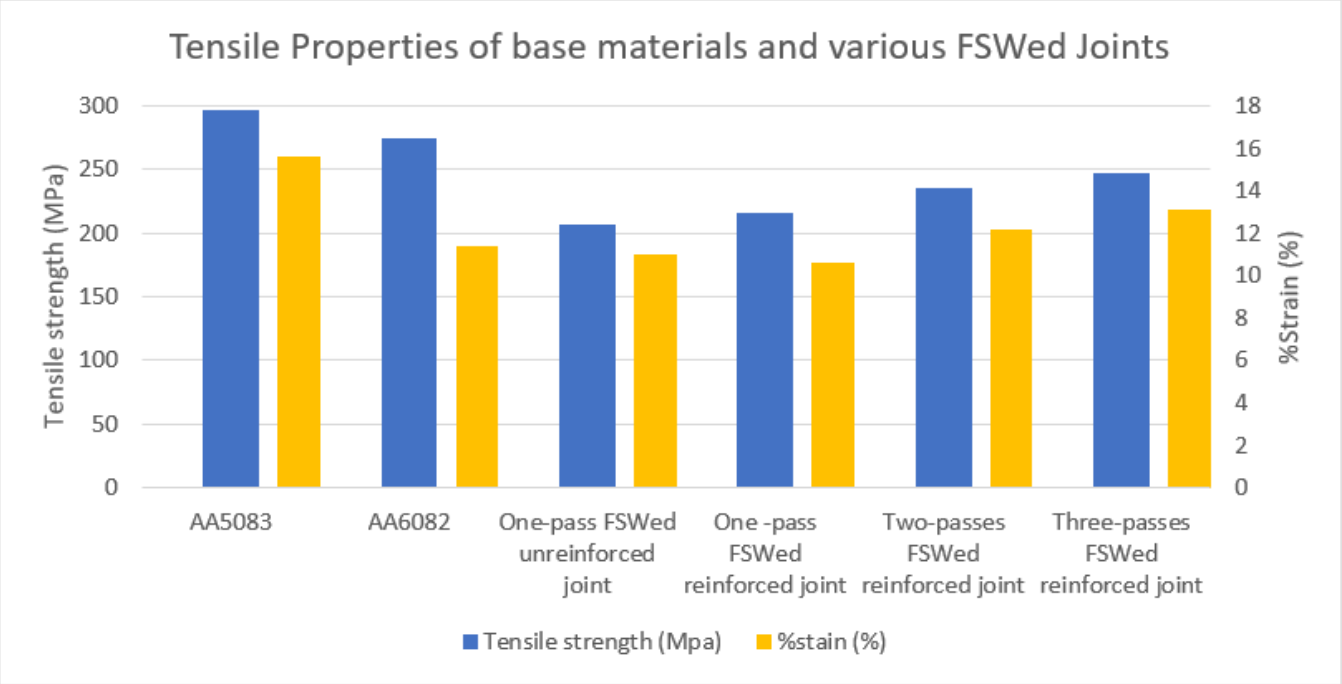


Figure 8

Tensile characteristics of various FSWed joints

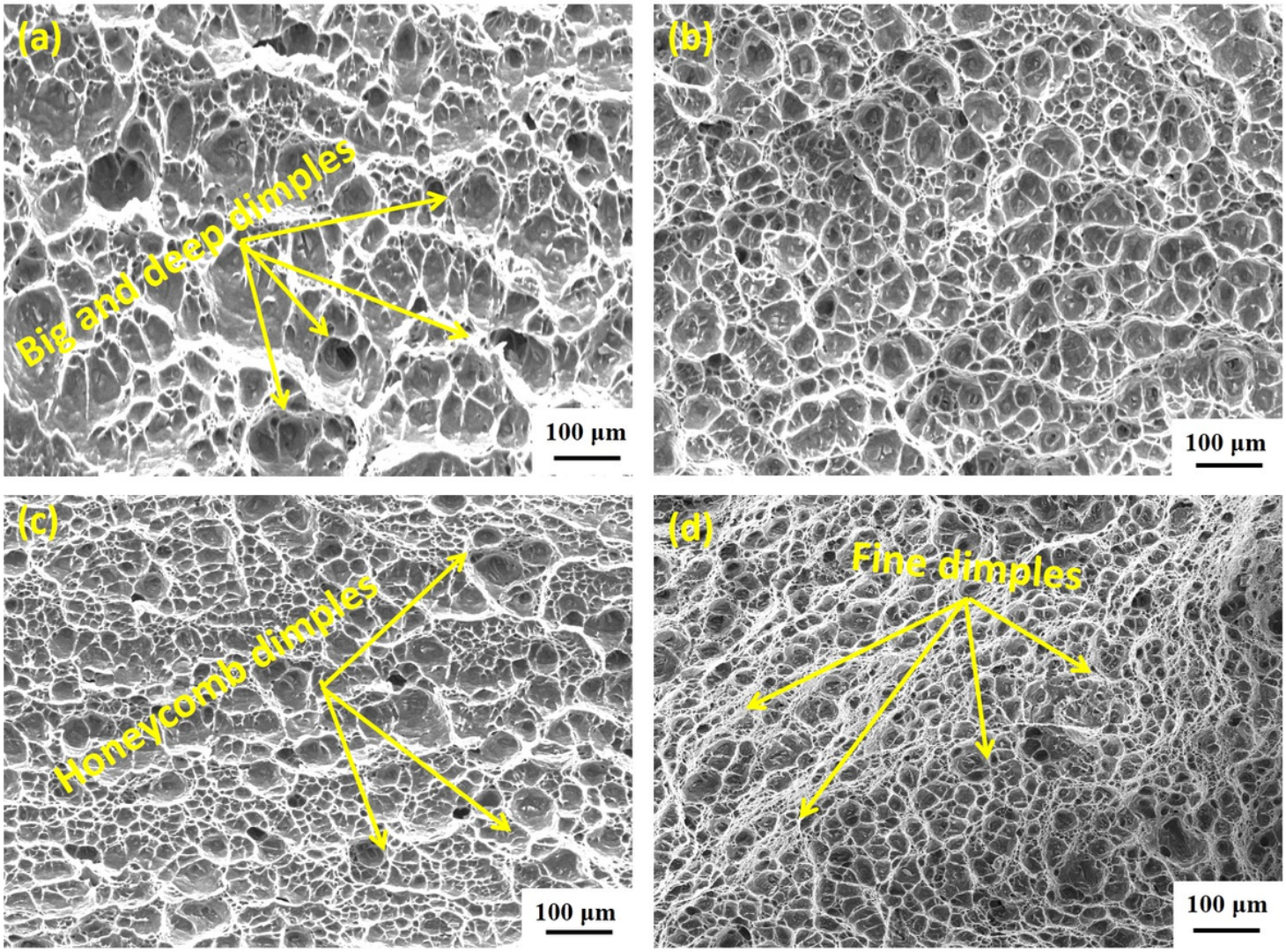


Figure 9

Fractography of fractured tensile samples; (a) FSWed unreinforced joint, (b-d) multi-pass FSWed reinforced joints, (b) one-pass (c) two-passes, (d) three-passes

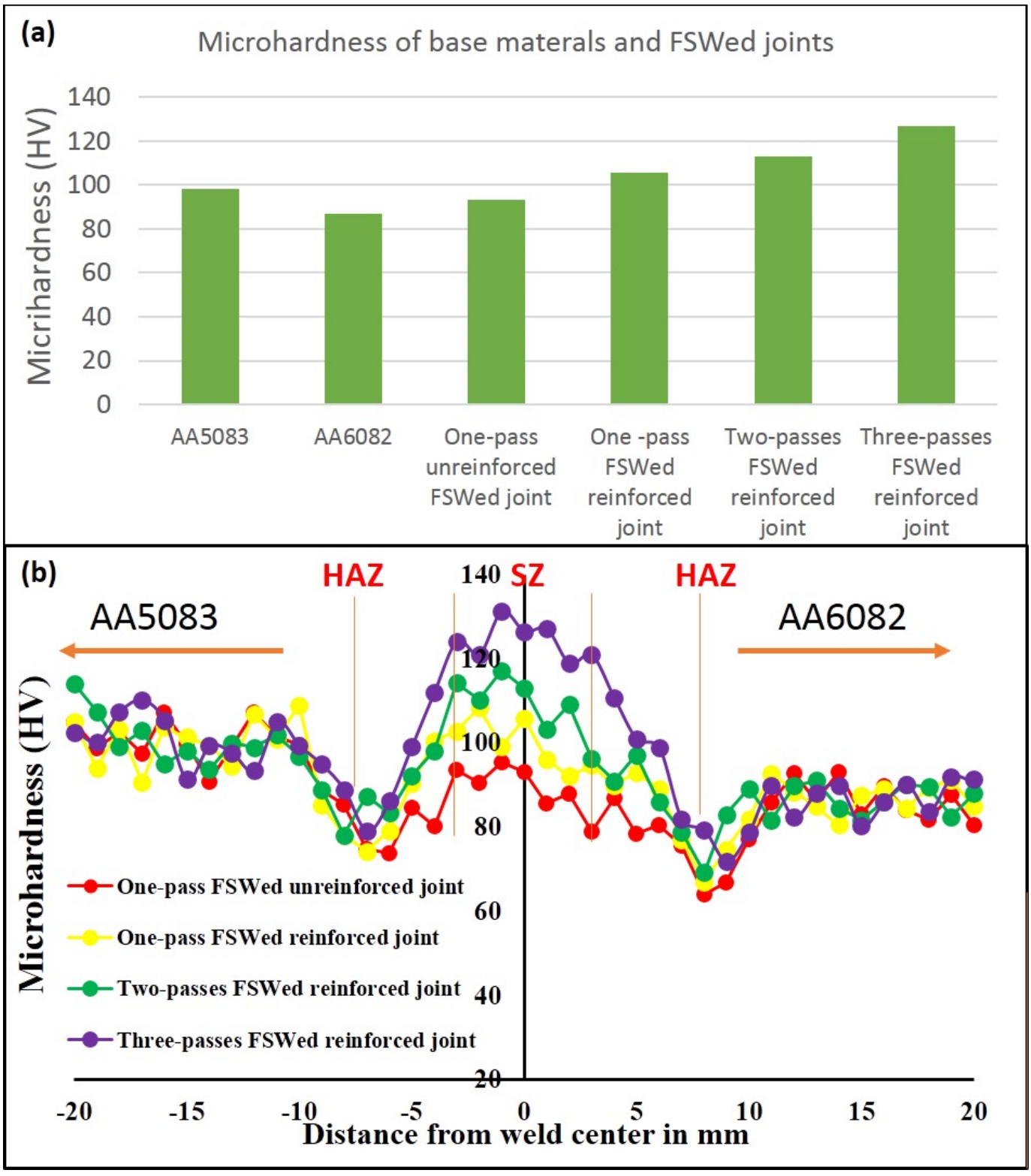


Figure 10

(a-b) Microhardness of various reinforced FSWed joints (a) Mean microhardness at the weld center; (b) Microhardness distribution in the weld zone of the

membranes

MDPI

Article

Integration of Pre-Treatment with UF/RO Membrane Process for Waste Water Recovery and Reuse in Agro-Based Pulp and Paper Industry

Sumit Dagar [1,*], Santosh Kumar Singh [1] and Manoj Kumar Gupta [2,*]

[1] Department of Environmental Engineering, Delhi Technological University, New Delhi 110042, India
[2] Environmental Management Division, Central Pulp and Paper Research Institute, Saharanpur 247001, India
* Correspondence: dagarsumit95@gmail.com (S.D.); mkg_cppri@rediffmail.com (M.K.G.)

Abstract: This recent study aims to evaluate the efficacy of membrane filtration on recovery of water resource from agro-waste such as bagasse, crop-based pulp and paper mill waste. A mini pilot scale membrane system having a combination of pre-treatment filter unit (pre-filter, sediment filter and pre-carbon filter), ultra-filtration and reverse osmosis with spiral wound configuration were employed to evaluate the water reuse efficacy of effluent coming from the secondary clarifier of the conventional treatment plant of the mill. The operational conditions were optimized using Taguchi method at pH 8, temperature 32 °C, and pressure 2 bar and a flow rate of 60 l/hr. The qualities of the effluent from the secondary clarifier, and the permeate from both the combination, viz. Combination 1 (pre-treatment + ultra-filtration) and Combination 2 (pre-treatment + ultra-filtration+ reverse osmosis) were analyzed and the percentage reduction in pH, TDS, TSS, BOD, COD, Color, Lignin, Potassium and Sodium were calculated. The elimination of TDS, COD and BOD with Combination 1 was not promising (<22%). However, the installation of a RO membrane greatly reduced (>88%) the contaminants in both paper mill effluents. The obtained qualities of water from all the combinations were compared with the tolerance standard for reuse as process water. The quality of effluent from the secondary clarifier did not agree with any class of water quality. The permeate from the combination of pre-treatment and UF sufficiently reduced the TSS to reach the requirement. However, the combination of (pre-treatment + UF + RO) adequately complied with the quality standard required for reuse in the making of all grades of paper.

Keywords: membrane filtration; secondary clarifier; wastewater effluent; organic pollutants; membrane fouling

check for updates

Citation: Dagar, S.; Singh, S.K.; Gupta, M.K. Integration of Pre-Treatment with UF/RO Membrane Process for Waste Water Recovery and Reuse in Agro-Based Pulp and Paper Industry. *Membranes* **2023**, *13*, 199. https://doi.org/10.3390/membranes13020199

Academic Editor: Frank Lipnizki

Received: 3 January 2023
Revised: 30 January 2023
Accepted: 1 February 2023
Published: 6 February 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Water is a pervious natural resource and the greatest gift to human civilization. India is one of the richest countries owing to possessing approximately 4% of the world's total water resources with 1869.35 billion cubic meters (BCM) of average surface water [1], and is one of the most populous countries supporting around 17.1% of the total world population. In the last few decades, a remarkable growth in the economy has been observed in India. Overgrowing population, rapid urbanization and a booming economy along with changes in lifestyle and land use patterns have considerably impacted on India's demand for water. Being an agrarian country, agriculture is the largest user of the total water reserve with a usage of 78% (71 BCM), followed by the domestic and industrial sectors utilizing 6% and 5% of the water, respectively [2]. Rapid demographic and economic growth has increased the water demands in all sectors. The projected industrial water demand is 161 BCM by 2050 [3]. Studies have projected the per capita water availability to be around 1191 m^3 by 2050, making India a water-scarce country as per the International Standard of 1700 m^3 [4].

The paper and pulp industry is one of the oldest and most economically important industries to both developing and developed nations [5]. It has eminent value in the

commercial and financial expansion of the nation. Because of the crucial role of paper as an important medium of record over a long time, its production is increasing day by day. This continuous increasing demand of water among the different sectors will severely affect water intensive industries such as the paper and pulp industry. Pulp and paper manufacturing units are the third largest water-consuming industries in the country. As per the present manufacturing process available in the country, the water consumption by Indian industries is 200 m^3–250 m^3 per ton in agro and large pulp and paper manufacturing units which is higher than the World Bank prescribed usage of 55 m^3 per ton of pulp and paper produced. Recycled water and effluents can significantly reduce this freshwater consumption. The waste water generated through these pulp and paper industries is comprised of suspended solids, color, inorganic compounds such as carbonates, bicarbonates, chlorides, sulphate and various toxic chemicals. The effluent has high amounts of biological oxygen demand (BOD), chemical oxygen demand (COD), lignin and its derivatives which damages the receiving water bodies. Papermaking requires proper treatment of these water fractions before reusing them. The wastewater generated must be treated using selected treatment methods such as coagulation/flocculation, flotation, biological processes, adsorption and membrane filtration before being discharged to the environment [6]. Most of the mills are still using the conventional methods that are not effective to remove the entire pollution loads. In order to sustain these industries, the use of technologies that can help to remove the maximum pollutants and make the water worth reusing can be the only solution. Recently, due to their reduced chemical utilization, short processing steps, energy saving and being able to be incorporated into the existing operating treatment plants, membrane technology has received great attention [7–10]. Various studies investigating the treatment of effluent waste water are illustrated in Table 1. A significant reduction in the pollutants such as COD, BOD, grease, fats, oils, suspended particles, etc., has been observed in these studies. However, due to the presence of a huge amount of organic pollutants and chemicals, the complexity increases for utilizing membrane filtration for paper and pulp industries. The organic materials present in the effluent deposit on the surface of the membrane and block the membrane known as fouling, which subsequently reduces the flux. The fouling of the membrane reduces its active surface area and hence decreases its efficiency; thus, it leads to a reduction in the membrane's lifespan, higher operating costs and membrane replacement expenditure [7,11]. Significant research has been devoted to address the issue of membrane fouling focusing on the fouling characterization [11], fouling mechanisms [12], pre-treatment methods [13] and fouling prevention and cleaning regimes [14–16].

Table 1. Physico-chemical parameters for analysis.

S.No	Physico-Chemical Parameters	Method Used
1	pH	Potentiometer
2	Chemical oxygen demand (COD)	Open efflux (Potassium Dichromate Method)
3	Biochemical Oxygen demand (BOD)	5 Days Incubation at 25 °C
4	Total dissolved solids (TDS)	IS: 3025 (Part 16)
5	Total suspended solids (TSS)	Gravimetric Method
6	Color	Spectrophotometric Method
7	Lignin	APHA method 5550 B
8	Electrical Conductivity	Conductivity Meter
9	Sodium and Potassium	Flame Photometer

This current study was conducted to test the feasibility of utilizing membrane filtration viz., ultra-filtration (UF) and reverse osmosis (RO) to reduce the pollution load from the effluent of pulp and paper industry. The membranes were used in combination with a pre-treatment filter and the operating conditions of the membrane were optimized to reduce the membrane fouling and obtain the maximum possible reduction in pollution.

2. Materials and Methods

2.1. Sample Collection

In the present work, an agro-waste-based pulp and paper manufacturing mill, i.e., Bindal Paper Mills Ltd., located in Muzaffarnagar, Uttar Pradesh, India was selected for effluent collection. Grab samples of effluent were collected from the secondary outlet (final treated effluent) of the effluent treatment plant and stored by refrigerating at 4 °C and then brought to room temperature prior to experimentation.

2.2. Experiment Set Up

The present study was undertaken to test the efficacy of assembled effluent treatment technologies. Three membrane technologies viz., pre-filtration unit (combination of pre-filter, sediment filter and pre-carbon filter), ultra-filtration (UF) membrane and reverse osmosis (RO) membrane were set up in two different assembly combinations. The first combination was the series assembly of pre-filtration and UF membrane (Figure 1), while the second combination was the series assembly of pre-filtration, UF membrane and RO membrane (Figure 2). These combinations were employed with a diaphragm-type pump that circulated the effluent from the pre-filter set up to the UF membrane. A flow meter was connected to collect the permeate (i.e., the treated water) from the experimentation unit. The sediment filter and pre-carbon filter with a max flow rate of 100 gpm and pressure $\leq$8 bar and temperature of 37 °C were used. The ultra-filtration membrane made up of polyether sulphone (PES) with an area of 81.073 cm^2, average pore size of 0.1 µ and molecular weight cut-off of 30 kDa having the capacity to withstand a pH range of 0–14, trans-membrane pressure and temperature $\leq$40 bars and 45 °C, respectively, was used for the experiment. A reverse osmosis membrane made of polypropylene with a capacity of 15 LPH to withstand a trans-membrane pressure of 30 Bar and temperature up to 40 °C was used.

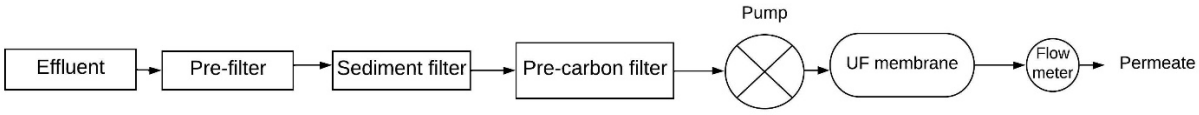


Figure 1. Flowchart of experiment set up (Combination 1: pre-filtration and UF membrane installed in series).

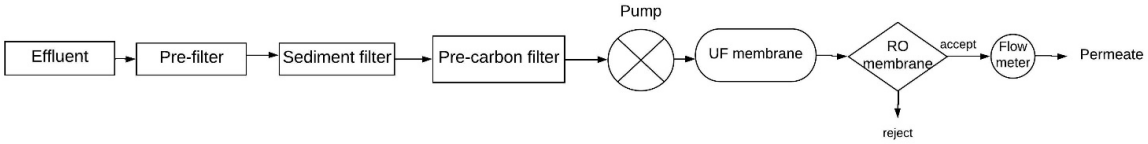


Figure 2. Flowchart of experiment set up (Combination 2: pre-filtration, UF membrane and RO membrane in series).

2.3. Operating Conditions

The Taguchi method was used to design the experiment because of its systematic, simple and efficient approach for optimization of parameters. The Taguchi method applies fractional experimental designs called orthogonal arrays (OA) to reduce the number of experimental required to determine the optimum conditions on the results. Four different factors and three levels were selected. The selected factors and their ranges were pH: 6–8; temperature: 20–32 °C; trans-membrane pressure: 2–4 bar and flow rate 30–60 l/hr. Flux decline caused by fouling and COD rejection rate were chosen as the response parameter to evaluate the membrane fouling. For the experimental design with four factors and three levels for each factor, an L9 (3^4) orthogonal array with eight degrees of freedom was selected. The optimal operating conditions based on Taguchi and the Grey Relation method were found to be at pH 8, temperature 32 °C, and pressure 2 bar and a flow rate of 60 l/hr.

2.4. Sample Analysis

The samples were analyzed for the pH, conductivity, TDS, TSS, BOD, COD, lignin, color, K^+, Na^+ in triplicate. All the analyses were performed as per the standard methods given provided by APHA, 2005 as provided in Table 1.

The percentage removal of various pollutants was calculated using Equation (1):

$$\text{Rej } (\%) = \left(1 - \frac{\text{Xp}}{\text{Xf}}\right) \times 100 \quad (1)$$

Rej (%) = Percentage Rejection; X_p = Concentration in permeate; X_f = Concentration in feed.

3. Results and Discussion

The effluent from the secondary clarifier after biological treatment, taken from agro-based pulp and paper mills located in Saharanpur, Uttar Pradesh, producing writing and printing grades of paper, were collected for further treatment through an advanced treatment system, i.e., laboratory-scale membrane filtration system comprised of pre-filtration, ultra-filtration (UF), membrane and reverse osmosis (RO) membrane to check the water reuse potential.

3.1. Characterization of Effluent Secondary Clarifier

In India, most integrated paper mills use agro waste such as bagasse raw materials. A region's climatic conditions and geography determine the physicochemical properties of wood. Untreated waste fluids might have significant biological oxygen demand (BOD), chemical oxygen demand (COD), suspended particles (primarily fibers), fatty acids, tannins, resin acids and lignin and its derivatives, depending on the raw material and the procedure involved [17]. The characteristics of the waste water effluent from the secondary clarifier are shown in Table 2. All other parameters are falling within the prescribed limits of water discharge given by CPCB.

Table 2. Physico-chemical properties of effluent from Bindal paper mill.

Parameter	Values	CPCB Limits
pH	7.46 ± 0.44	5.5–9.0
TSS mg/L	56 ± 0.96	100
TDS, mg/L	2215 ± 56.45	2100
COD, mg/L	216 ± 9.11	250
BOD, mg/L	27 ± 0.56	30
Color, PCU	412 ± 21.1	500
Lignin, mg/L	126 ± 7.71	-
Conductivity, µS/cm	2630 ± 81.21	-
Sodium (Na), mg/L	49 ± 2.1	
Potassium (K), mg/L	7.2 ± 0.17	-

3.2. Performance of Membrane Filtration Unit

Effluent samples after secondary clarifier were passed through two different combinations viz. (1). Pre-filtration and UF membrane installed in series, (2) Pre-filtration, UF Membrane and RO Membrane in series.

As shown in Figure 3, the pH of the effluent reduced to 7.19 ± 0.12 with a combination of pre-filter and ultra-filtration. On the other hand, the pH further reduced to 5.74 ± 0.51 with the addition of the RO membrane. The change in the TSS level with both combinations is shown in Figure 4, both the combinations of membrane filtration configuration were able to entirely reduce the TSS. However, the removal of TDS is not satisfactory with pre-treatment and UF, as only 7.7% (2215 ± 56.45 mg/L to 1951 ± 41.23 mg/L) of TDS could be reduced. The integration of the RO membrane with UF filtration and pre-filtration could, however, reduce 88.26% of the TDS (Figure 5). Similar results have been observed by

the authors of [18,19] on wastewater from poultry slaughter houses and gray water where the application of UF reduced the TSS to approximately 98% and 100%, respectively.

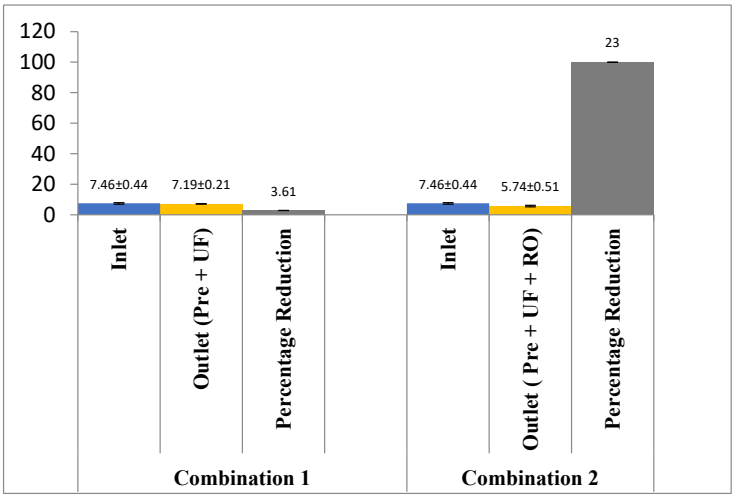


Figure 3. Effect of different treatment setups on the pH rejection.

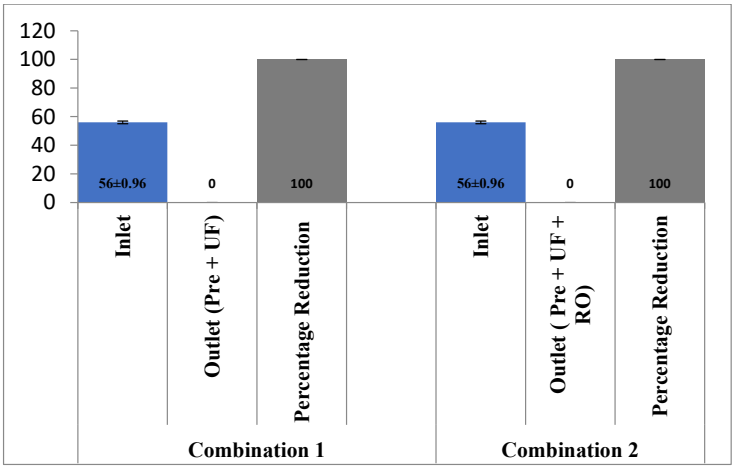


Figure 4. Effect of different treatment setups on the TSS rejection.

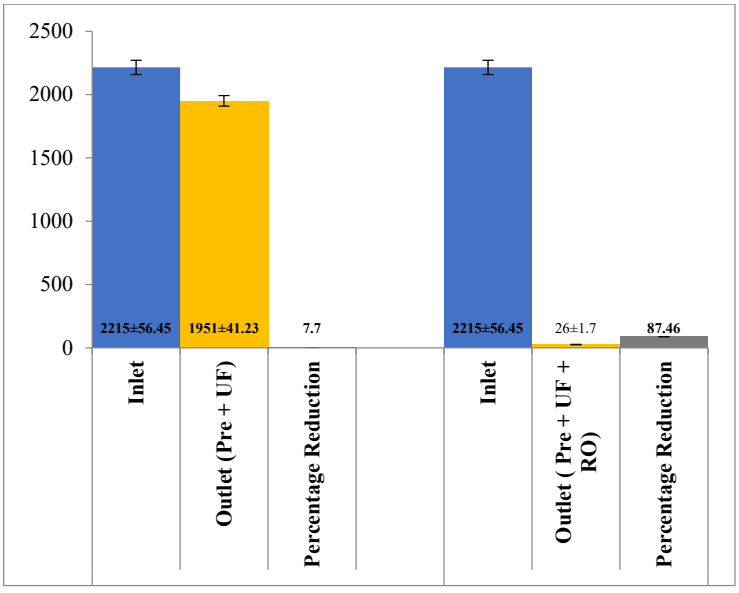


Figure 5. Effect of different treatment setups on the TDS rejection.

In terms of COD and BOD, the reduction potential using pre-filtration and UF was 17.12% (216 ± 9.11 mg/L to 179 ± 10.76 mg/L) and 22% (27 ± 0.56 mg/L to 21 ± 0.71 mg/L), respectively. However, the addition of the RO membrane along with pre-filtration and UF appeared to reduce 100% of COD and BOD from the effluent (Figures 6 and 7). These results are in agreement with the study conducted by [20] on phenolic wastewater generated from a paper mill and with the application of a combination of UF-NF/RO a reduction of 95.5% in COD has been observed. Yordanov [18] had also observed a reduction of more than 94% in BOD of the wastewater from the slaughter house with the application of ultra-filtration.

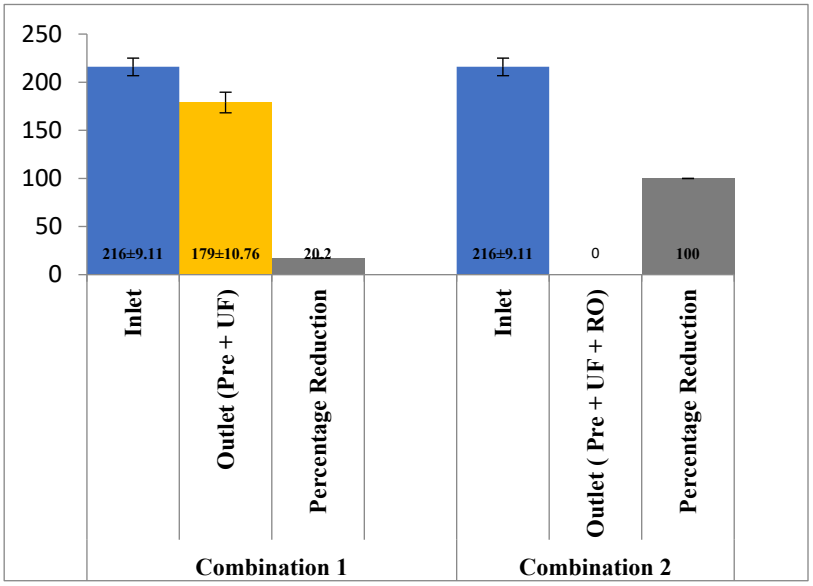


Figure 6. Effect of different treatment setups on the COD rejection.

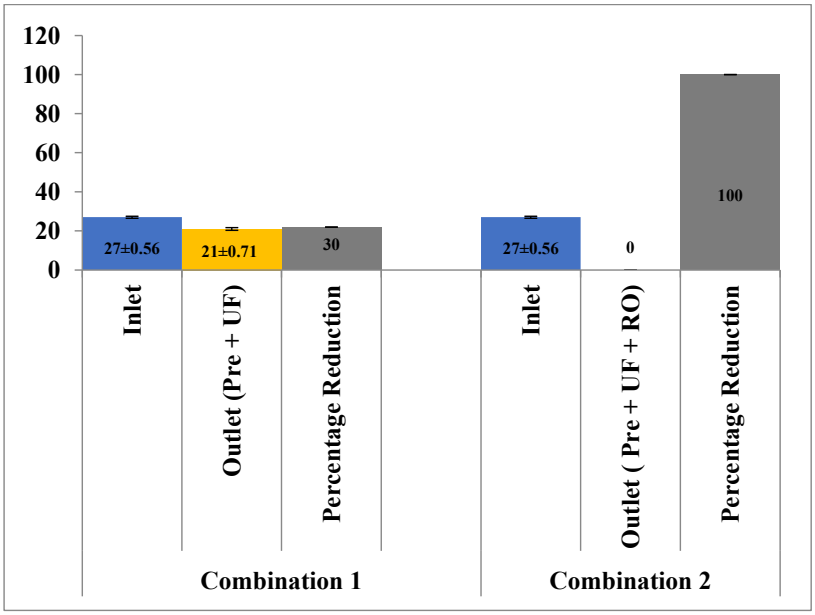


Figure 7. Effect of different treatment setups on the BOD rejection.

Pre-filtration and UF in series could reduce the color from 412 ± 21.10 PCU to 108 ± 10.2 PCU, i.e., by 73.78% (Figure 8). Lignin is the main component of color addition in the paper industry. In terms of lignin removal from the effluent, the pre-filtration and UF membrane fared significantly better as they were able to reduce 83% (126 ± 7.71 mg/L to 22 ± 1.01 mg/L) of the lignin although the additional RO membrane could effectively reduce the entire lignin content (100%) from the effluent (Figure 9). The application of

nanofiltration by the authors of [21] on textile industry wastewater reduced the color by 100%. Similarly, the use of ultra-filtration alone rejected 98% of the synthetic dye from synthetic dye wastewater [22].

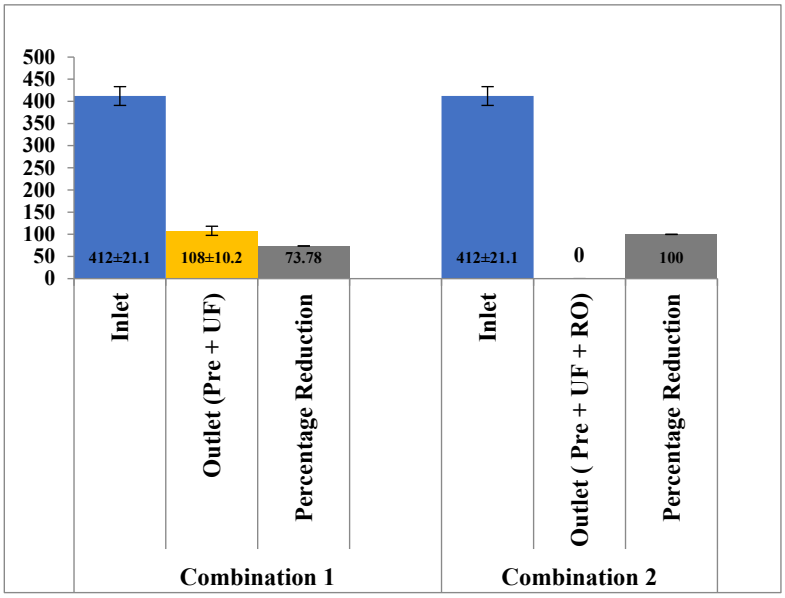


Figure 8. Effect of different treatment setups on the color rejection.

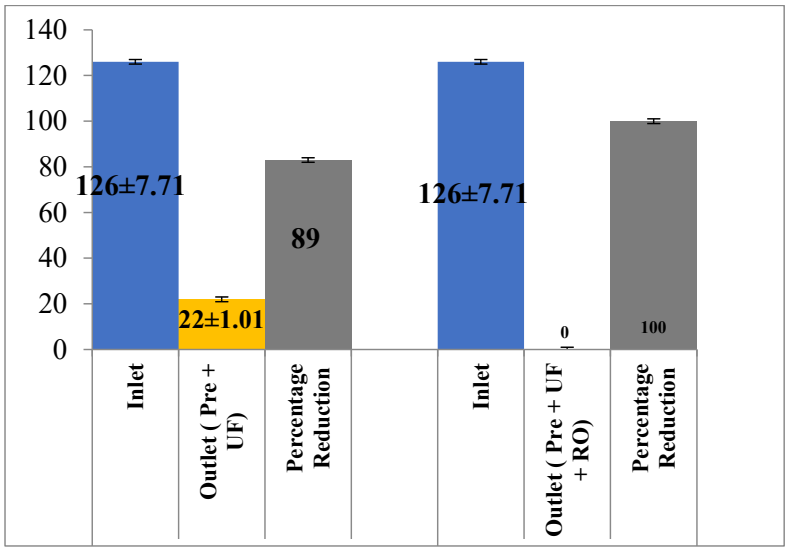


Figure 9. Effect of different treatment setups on the lignin removal.

In terms of conductivity, the reduction potential using pre-filtration and UF was 31% (1365 ± 81.21 mg/L to 942.9 ± 79.37 mg/L). However, the addition of the RO membrane along with pre-filtration and UF further reduced the conductivity to 98.16% from the effluent (Figure 10).

In terms of potassium and sodium, the reduction potential using pre-filtration and UF was 14.28% (4.2 ± 0.17 mg/L to 3.6 ± 1.1 mg/L) and 19% (28 ± 2.1 mg/L to 22.70 ± 1.72 mg/L), respectively. However, the addition of the RO membrane along with pre-filtration and UF further reduced the potassium and sodium ions to 98% and 87% from the effluent (Figures 11 and 12). Reverse osmosis is more effective than ultra-filtration in removal of ions such as potassium and sodium from the waste water.

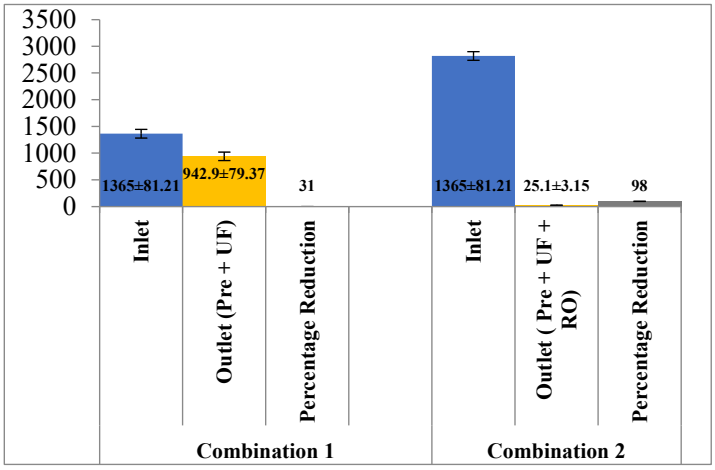


Figure 10. Effect of different treatment setups on the conductivity.

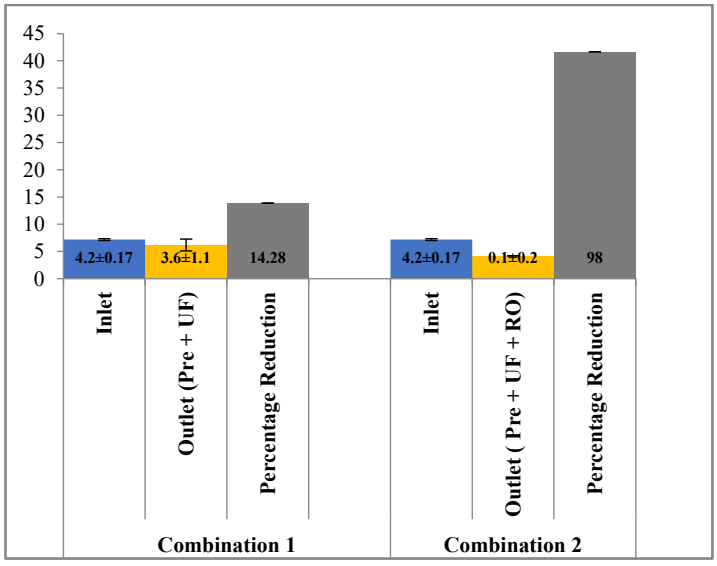


Figure 11. Effect of different treatment setups on removal of potassium.

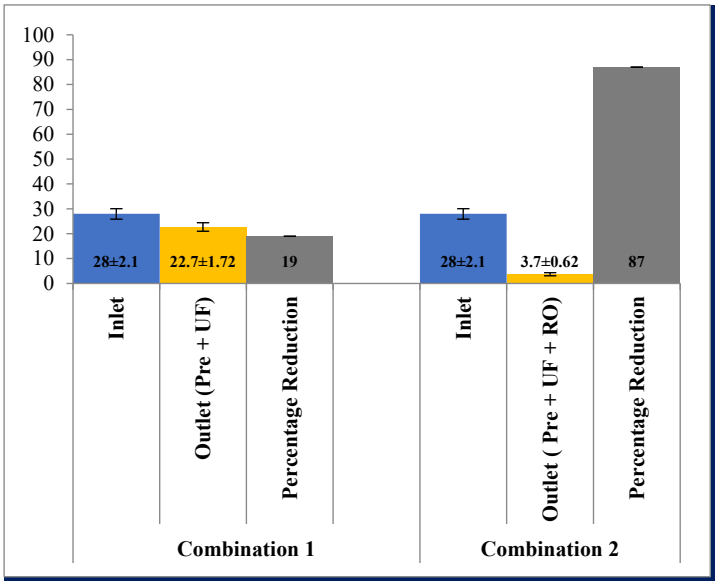


Figure 12. Effect of different treatment setups on removal of sodium.

The results obtained from our study are in good agreement of various previous studies to remove pollutants from waste water using both RO and UF membranes. There has been a significant decrease ($p < 0.05$) in the concentration of all pollution in both the treatment methods (Table 3). Membrane filtration at low pressure is reported to be best for bleach effluent treatment in pulp and paper mills [23]. Likewise, it has also been reported that the performance of using RO in the reduction in pulp and paper industry effluent is higher than UF and nanofiltration [24]. In addition, the authors of [25] have recommended the combined RO-UF membrane process for the treatment of wastewater.

Table 3. Change in the various water parameters with different treatment membrane setups.

	TSS (mg/L)	TDS (mg/L)	COD (mg/L)	BOD (mg/L)	Color (PCU)	Lignin (mg/L)	pH	Conductivity (µs)	Potassium (mg/L)	Na (mg/L)
Inlet (Secondary Clarifier)	56 ± 0.96	2215 ± 88.26	216 ± 9.11	27 ± 0.56	412 ± 21.1	126 ± 7.71	7.46 ± 0.44	1365 ± 81.21	4.2 ± 0.17	28 ± 2.1
Outlet (Pre-treatment+ UF)	0 *	1951 ± 56.45 *	179 ± 10.76 *	21 ± 0.71 *	108 ± 10.2 *	22 ± 1.01 *	7.19 ± 0.12 *	942.9 ± 79.37 *	3.6 ± 1.1 *	22.7 ± 1.72 *
Outlet (Pre-treatment+ UF + RO)	0 *	26 ± 1.70 *	0 *	0 *	0 *	0 *	5.74 ± 0 *	25.1 ± 3.15 *	0.1 ± 0 *	3.7 ± 0.62 *

* Significant at $p < 0.05$, Two-tailed independent *t*-test.

The permeate of both the combinations was compared with the feed water characteristics required by the pulp and paper industry listed in Table 4. The effluent from the secondary clarifier cannot be reused for the production of paper since none of the parameters fall under the prescribed tolerance limits. The application of pre-filtration and ultra-filtration however reduced the pollution load but, except for TSS, none of the parameters lay in the tolerance regime. However, the permeate from the series combination of pre-filtration, ultra-filtration and the RO unit was able to achieve the tolerance limits for color, TDS and TSS to be reused in the production of all the grades of paper.

Table 4. Water Quality requirements for paper and pulp industries in India.

	Effluent from Different Processes			Tolerance for Water for Pulp and Paper Industry (BIS)			
Water Quality Parameter	**Effluent from Secondary Clarifier**	**Permeate from Pre-Filtration and Ultra-Filtration Unit**	**Permeate from Pre-Filtration, Ultra-Filtration and RO Unit**	**Ground Wood Paper**	**Kraft Paper Bleached**	**Soda and Sulfite Paper**	**High Grade Paper**
Color	615	176	0	20	15	10	5
TDS	1933	1784	0	500	300	300	300
TSS	36	0	0	25	25	25	10
COD	183	146	0	NS	NS	NS	NS

NS—Not specified.

4. Conclusions

The purpose of the study was to evaluate the efficacy of membrane technology for the removal of pollutants from an agro residue-based paper mill secondary clarifier effluent and to check the potential for reuse of the treated wastewater for paper production. For this, a tertiary pilot system incorporating a membrane filtration unit, i.e., pre-filtration, ultra-filtration (UF), and reverse osmosis (RO) membrane were used. Two combinations of membrane filtration were used viz., Combination 1 comprised of series assembly of pre-filtration and UF whereas the RO membrane is added with pre-filtration and UF in the Combination 2. The operating conditions were optimized in order to minimize the membrane fouling and maximize the pollution rejection.

The combination of the pre-filtration filter and ultra-filtration could satisfactorily reduce the TSS, color and lignin from both the paper mill effluents. The placement of the pre-filtration filter helps to reduce the pollutant load on the membrane hence decreasing the fouling of the membrane. The pre-filtration filter also reduces the economic cost of the membrane cleaning and replacement due to fouling. However, the removal of TDS, COD and BOD using pre-treatment and ultra-filtration was not satisfactory (<30%). This may be attributed to the high concentration of low molecular weight compounds in the effluent. However, the addition of the RO membrane could significantly (>98%) reduce the pollutants from the paper mill effluent. RO membrane exhibited the highest pollutant removal ability when used in conjunction with pre-filtration and UF membrane. So, the application of pre-treatment, ultra-filtration and reverse osmosis in series could be an effective measure to help in the recovery of the vital and finite water resource and promote its reuse in the industries.

Author Contributions: Conceptualization, supervision: S.K.S.; Methodology, Software, Formal Analysis, Investigation, Writing: S.D.; Review, Proof Reading: M.K.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors are thankful to Delhi Technological University and Environmental Management Division, Central Pulp and Paper Research Institute for providing the resources to carry out the research work.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ministry of Water Resources, Government of India. *Central Water Commission Annual Report, 2010*; Ministry of Water Resources, Government of India: New Delhi, India, 2010.
2. Ministry of Water Resources, Government of India. *Public Information Bureau, 2013*; Ministry of Water Resources, Government of India: New Delhi, India, 2013.
3. Government of India. *Infrastructural Statisctics-2010, A Report by Ministry of Statistics and Programme Implementation*; Government of India: New Delhi, India, 2010.
4. Gangwar, A.; Prakash, R.; Nagaraja, T.P. Applicability of Microbial Xylanases in Paper and Pulp Bleaching: A Review. *Bioresources* **2014**, *9*, 3733–3754. [CrossRef]
5. Roy, T.K. Overview of Indian Paper Industry. In Proceedings of the Workshop on Adoption of Energy Efficient Process Technologies and Energy Management Practices in Pulp and Paper Sector under Energy Conservation Act, Saharanpur, India, 14 February 2007.
6. Putatunda, S.; Bhattachrya, S.; Sen, D.; Bhattacharjee, C. A review on application of different treatment process for emulsified oily wastewater. *Int. J. Environ. Sci. Technol.* **2019**, *16*, 2525–2536. [CrossRef]
7. He, Y.; Bagley, D.M.; Leung, K.T.; Liss, S.N.; Liao, B.A. Recent advances in membrane technologies for biorefining and bioenergy production. *Bitechnol. Adv.* **2012**, *30*, 817–858. [CrossRef]
8. Asif, M.B.; Zhang, Z. Ceramic membrane technology for water and wastewater treatment: A critical review of performance, ful-scale applications, membrane fouling and prospects. *Chem. Eng. J.* **2021**, *418*, 129481. [CrossRef]
9. Martini, S.; Haming, A. Hybrid TiO2/UV/PVDF ultrafiltation membrane for anola oil wastewater treatment. *Desalination Water Treat.* **2019**, *148*, 51–59. [CrossRef]
10. Yadav, M.K.; Saidulu, D.; Gupta, A.K.; Ghosal, A.M. Status and management of arsenic pollution in groundwater: A comprehensive appraisal of recent global scenario, human health impacts, sustainable filed scale treatment technologies. *J. Envion. Chem. Eng.* **2021**, *9*, 105203. [CrossRef]
11. Puro, L.; Kallioinen, M.; Mänttäri, M.; Nyström, M. Evaluation of behavior and fouling potential of wood extractives in ultrafiltration of pulp and paper mill process water. *J. Memb. Sci.* **2011**, *368*, 150–158. [CrossRef]
12. Puro, L.; Tanninen, J.; Nyström, M. Analyses of organic foulants in membranes fouled by pulp and paper mill effluent using solid-liquid extraction. *Desalination* **2002**, *143*, 1–9. [CrossRef]
13. Koivula, E.; Kallioinen, M.; Preis, S.; Testova, L.; Sixta, H.; Mänttäri, M. Evaluation of various pre-treatment methods to manage fouling in ultrafiltration of wood hydrolysates. *Sep. Purif. Technol.* **2011**, *83*, 50–56. [CrossRef]

14. Maartens, A.; Jacobs, E.P.; Swart, P. UFof pulp and paper effluent: Membrane fouling-prevention and cleaning. *J. Membr. Sci.* **2002**, *209*, 81–92. [CrossRef]
15. Li, J.; Sanderson, R.D.; Jacobs, E.P. Ultrasonic cleaning of nylon microfiltration membranes fouled by Kraft paper mill effluent. *J. Membr. Sci.* **2002**, *205*, 247–257. [CrossRef]
16. Weis, A.; Bird, M.R.; Nystrom, M. The chemical cleaning of polymeric UF membrane fouled with spent sulphate liquor over multiple operational cycles. *J. Membr. Sci.* **2003**, *216*, 67–79. [CrossRef]
17. Kamali, M.; Khodaparast, Z. Review on recent developments on pulp and paper mill wastewater treatment. *Ecotoxicol. Environ. Saf.* **2015**, *114*, 326–342. [CrossRef]
18. Yordanov, D. Preliminary study of the efficiency of ultrafiltration treatment of poultry slaughterhouse wastewater. *Bulgatian J. Agric. Sci.* **2010**, *16*, 700–704.
19. Oh, K.S.; Poh, P.E.; Chong, M.N.; Chan, E.S.; Lau, E.V.; Saint, C.P. Bathroom grey water recycling using polyelectrolyte-complex bilayer membrane: Advanced study of membrane structure and treatment efficiency. *Carbohydate Polym.* **2016**, *148*, 161–170. [CrossRef]
20. Sun, X.; Wang, C.; Li, Y.; Wang, W.; Wei, J. Treatment of phenolic wastewater by combined UF and NF/RO processes. *Desalination* **2015**, *355*, 68–74. [CrossRef]
21. Ellouze, E.; Tahri, N.; Amar, R.B. Enhancement of textile wastewater treatment process using Nanofiltration. *Desalination* **2012**, *286*, 16–23. [CrossRef]
22. Lin, J.; Ye, W.; Baltaru, M.C.; Tang, Y.P.; Bernstein, N.J.; Gao, P.; Balta, S.; Vlad, M.; Volodin, A.; Sotto, A.; et al. Tight ultrafiltration membranes for enhanced separation of dyes and Na2SO during textile wastewater treatment. *J. Membr. Sci.* **2016**, *514*, 217–228. [CrossRef]
23. Shukla, S.K.; Kumar, V.; Kim, T.; Bansal, M.C. Membrane filtration of chlorination and extraction stage bleach plant effluent in Indian paper Industry. *Clean Technol. Environ. Policy* **2013**, *15*, 235–243. [CrossRef]
24. Shukla, S.K.; Kumar, V.; Bansal, M.C. Treatment of combined bleaching effluent by membrane filtration technology for system closure in paper industry. *Desalination Water Treat.* **2010**, *13*, 464–470. [CrossRef]
25. Li, S.Z.; Li, X.Y.; Wang, D.Z. Membrane (RO-UF) filtration for antibiotic wastewater treatment and recovery of antibiotics. *Sep. Purif. Technol.* **2004**, *34*, 109–114. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

I. J. Education and Management Engineering, **2023, 1, 29-34**
Published Online on February 8, 2023 by MECS Press (http://www.mecs-press.org/)
DOI: 10.5815/ijeme.2023.01.04

Intelligent Model for Smartphone Addiction Assessment in University Students using Android Application and Smartphone Addiction Scale

Anshika Arora*
Department of Computer Science and Engineering, Netaji Subhas University of Technology, New Delhi-110078, India
Email: anshika.arora.trf19@nsut.ac.in
ORCID iD: https://orcid.org/0000-0003-0414-1874
*Corresponding author

Pinaki Chakraborty
Department of Computer Science and Engineering, Netaji Subhas University of Technology, New Delhi-110078, India
Email: pinaki_chakraborty_163@yahoo.com
ORCID iD: https://orcid.org/0000-0002-2010-8022

M.P.S. Bhatia
Department of Computer Science and Engineering, Netaji Subhas University of Technology, New Delhi-110078, India
Email: bhatia.mps@gmail.com
ORCID iD: https://orcid.org/0000-0001-7190-9770

Aditya Puri
Department of Electronics and Communication Engineering, Netaji Subhas University of Technology, New Delhi-110078, India
Email: adityapuri8888@gmail.com

Received: 17 July, 2022; Revised: 06 September, 2022; Accepted: 16 October, 2022; Published: 08 February, 2023

Abstract: Smartphones have been owned and used ubiquitously in all facets of society utilized for a wide number of tasks such as calling and messaging, social media, surfing as well as for entertainment. Spending a large amount of time on smartphone might lead to a dependence on it for a variety of purposes. This study uses objective measures of real time smartphone usage features to assess smartphone addiction. A purpose built android application to collect real time smartphone usage has been developed and linear classification models namely Support Vector Machine and Logistic Regression are used to predict smartphone addiction among university students. Furthermore, correlation and information gain measures are used to identify most vital features of smartphone usage which contribute maximum in assessment of smartphone addiction. It has been observed that both the linear models give worthy performance with more than 80% of accuracy. Also, the most important technical features impacting smartphone addiction are longest session spent for entertainment, total time used for communication, longest session spent for communication, longest session spent for work, total time used for entertainment, longest session for news and surfing, and data usage in other activities.

Index Terms: Smartphone addiction, machine learning, linear classification, android application.

1. Introduction

Smartphone usage has become inevitable with phones becoming an integral part of lifestyle serving multiple purposes such as communications, social networking, media streaming, information retrieval, etc. Spending significant amounts of time using smartphones for numerous activities potentially develop dependency on them for numerous purposes which may lead to addictive usage and development of associated psychological disorders [1,2]. There are many impairments associated with excessive and addictive used of smartphones such as increase in the risk of sleeping disturbance and stress, sensation of being less physically active, fall in mental health and even depression [3]. Novel

This work is open access and licensed under the Creative Commons CC BY License.

psychological variables such as nomophobia and Fear of Missing Out (FOMO) arising from smartphone addiction have also been studied [2,4]. Researchers suggest that smartphone use can be detrimental if it forms an addiction via abuse of technology.

Smartphone addiction is conventionally assessed via smartphone addiction scales, most of which rely on self-reported data obtained through questionnaires [5]. However, a research restriction is that self-report based smartphone addiction scales cannot reliably measure the actual phone usage. [5]. Moreover, this pairs with other shortcomings of these type of scales such as unconscientious or missing responses, and wrong interpretations of questions which lead to inaccurate information. Furthermore, filling of questionnaires is a cumbersome task for participants due to a large number of questions. Purpose-built smartphone applications that can track phone usage attributes in order to provide accurate measures of real-time usage patterns can provide an intervention for this. Subjective data from questionnaires can be replaced by application extracted objective measures of smartphone usage features.

Motivated by this, the objective of this study is to use real time smartphone usage data to predict smartphone addiction using supervised machine learning. A purpose built android application namely "UsageStats" has been developed with the capability to track real time smartphone usage patterns among multiple attributes such as duration of smartphone use and internet data used over the past 30 days. The applications installed in the smartphone are categorized into nine types namely Social Media, Communication, Entertainment, Productivity, News & Surfing, Gaming, Work, Photos & Camera, and Others. Patterns of usage of various categories of applications are attributed in features such as category total time used, category sessions, category longest session, and data used in that category. Smartphone addiction is labeled using the measures of most commonly used smartphone addiction scale, i.e. Smartphone Addiction Scale-Short Version (SAS-SV) designed by Known et al. [6]. Linear classifiers including Support Vector Machine (SVM) and Logistic Regression (LR) are used to predict smartphone addiction. Also, the most contributing attributes of smartphone usage are identified using correlation attribute evaluator and Information Gain (IG) attribute evaluator using predefined thresholds.

2. Related Work

The conventional approach for assessment of smartphone addiction is based on questionnaire based smartphone addiction scales which have been discussed in existing studies [2,5]. Soft computing techniques have been applied to these self-reported data by a number of researchers to provide a technically sophisticated diagnosis of smartphone addiction. Shin et al. [7] used a variety of mobile phone usage data to discover a number of variables in order to build machine learning methods for automated prediction of problematic smartphone use, such as Nave Bayes, Support Vector Machine, and AdaBoost. Supervised machine learning models based on K-Nearest Neighbour, Decision Tree, Nave Bayes, and Support Vector Machine were proposed for smartphone addiction evaluation [8]. However the accuracy and reliability of these results is subject to biasness in information filled.

Recent studies have suggested successful interventions of real time smartphone usage tracking for analysis of behavioral patterns and assessment of problematic use [1,9]. Soft computing methodologies are being used on these real-time smartphone extracted data to provide an accurate diagnosis of smartphone addiction. Ellis et al. [10] implemented an unsupervised machine learning algorithm, i.e. K-Means clustering to group the users with similar smartphone usage patterns. They used the usage behavior pattern retrieved from Apple's Screen Time application which logs a series of behavioral screen time metrics. Other psychological variables arising from smartphone addiction have also been assessed by researchers using real time smartphone extracted usage patterns and soft computing techniques. Arora et al. [2] utilized Gaussian mixture clustering algorithm to identify nomophobic behavior in users with similar smartphone usage pattern retrieved via smartphone application. In another study, Arora et al. [11] extracted real time smartphone data via an android application for prediction of nomophobia severity using supervised machine learning algorithms such as Random Forest, Decision Tree, Support Vector Machines, Naïve Bayes And K-Nearest Neighbor. Elhai et al. [12] implemented supervised machine learning algorithms in order to detect problematic smartphone use severity among undergraduate students in China.

3. Materials and Methods

The materials and methods used in this study are described in the following subsections.

3.1 Android Application: UsageStats

In order to collect real time smartphone usage data of participants, an android application namely "UsageStats" has been developed which automatically collects data from the smartphones of user. Fig. 1 presents the UsageStats screen.

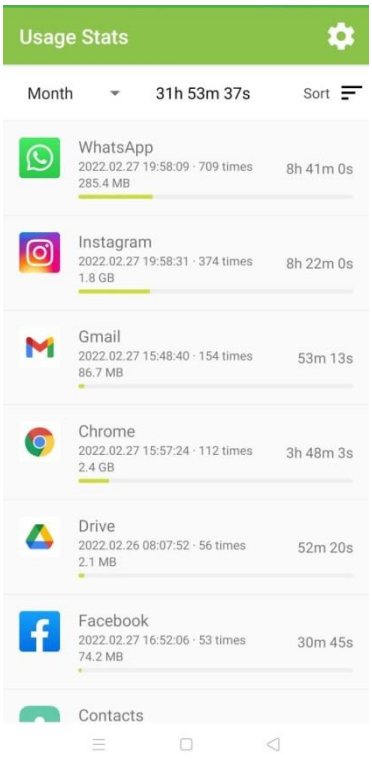


Fig. 1. Smartphone usage during the month

The application tracks overall smartphone usage as well as usage of various smartphone applications including the in-built applications and third party applications. The features of smartphone usage directly collected by UsgeStats are listed in Table 1.

Table 1. Features Extracted by Smartphone Application

S. No.	Attribute	Description
1.	Total_time	Total duration of smartphone use for the last 30 days
2.	Total_data_used	Sum of internet data used by all apps during the smartphone use in last 30 days
For each application		
1.	Time_used	Total duration of the application usage for last 30 days
2.	Times_open (sessions)	Total number of sessions of the application in last 30 days
3.	Longest_session	Longest duration in last 30 days for which the application is used in one launch
4.	Data_used	Sum of mobile and Wi-Fi internet data that has been used during the application usage in last 30 days
5.	Tag	Category to which the application belongs

The application is sent to undergraduate students of Netaji Subhas University of Technology and the records are maintained in an online database.

3.1.1 Smartphone Usage Attribute Set

After data collection using the UsageStats, the applications are categorized into nine tags viz. Social Meida, Communication, Entertainment, Productivity, News & Surfing, Gaming, Work, Photos & Camera, and Others. These tags are created by sorting and modifying the default tags extracted by the developed application. Applications are categorized as follows:

1. S*ocial Media* includes applications such as Twitter, Tumblr, Facebook, Instagram, Reddit.
2. C*ommunication* includes messaging applications and call applications like Snapchat, WhatsApp, Telegram, Truecaller, messenger, Discord and system applications such as contacts, phone, messaging, contacts and dialer etc.
3. *Entertainment* combines system applications such as movies & videos and music & audio as well as the third party video and music applications like Youtube, Netflix, Sonyliv, Prime video, Hotstar, Spotify, FM Radio Gaana, etc.
4. *Productivity* includes the applications like Paytm, Gpay, Docs, Sheets, etc.
5. *News & Surfing* includes various browsers such as Google, Chrome, Browser, Mi Browser, Internet and others as well as news and magazine applications like Inshorts, Quora, Google news, etc.
6. *Gaming* tag includes third party gaming applications like Battleground.

7. *Work* tag is maintained for all the work and study related applications including Gmail, Mail, LinkedIn, Classroom, Zoom, Unacademy, Udemy Meet, etc.
8. *Photos & Camera* include system applications like Albums, Camera, Photos, Gallery and third party applications like Picsart.
9. *Others* is the tag for applications belonging to none of the abovementioned categories. It includes applications like Files, numerous shopping applications, drive and maps applications and various food delivery applications.

After categorizing and sorting the applications into their respective tags, a dataset is created which for each user contains attributes of overall smartphone usage such as total_time in hours and total_data_used in GB and attributes for every tag including time_used in minutes, times_open (sessions), longest_session in minutes, and data_used in MB.

3.1.2 Normalization

The smartphone usage attribute set is normalized using the Min-Max normalization. For each attribute it transforms its values to a decimal between 0 and 1 which works as given in equation 1.

$$x'_a = \frac{x_a - min_a}{max_a - min_a} \quad (1)$$

where, x_a is a instance value belonging to attribute a,
min_a is the minimum value over attribute a
max_a is the maximum value over attribute a.

3.2 Labeling using SAS-SV

SAS-SV designed by by Known et al. [6] is the most commonly used scale for assessment of smartphone addiction [13]. This study uses SAS-SV to collect responses from students about their perception on their dependence on smartphones. SAS-SV contains 10 questions each of which has to be answered on a 6-point Likert scale. The sum of the answers is the indicator of presence of smartphone addiction with a cut off value of 33 for girls and 31 for boys. This is used for labeling students with smartphone addiction.

3.3 Classification

The dataset containing real time smartphone usage features with labels derived from SAS-SV is fed as input to two linear classification models namely, SVM and Logistic Regression. These models are chosen as they are the most suitable model for small datasets [14]. The performances of classification models are evaluated with accuracy and F1-score.

3.4 Identification of Most Contributing Smartphone Usage Attributes

This study uses statistical attribute evaluator methods which evaluate the relationship between each input variable and the target variable in order to use scoring measures as a basis to filter those input variables which contribute maximum for classification [15]. The methods used are Correlation attribute evaluator and Information Gain attribute evaluator. Correlation attribute evaluator evaluates the worth of an attribute by measuring the Pearson's correlation between it and the class [16,17]. IG attribute evaluator evaluates the worth of an attribute by measuring the information gain with respect to the class [16,18]. For each attribute evaluator, top 5 smartphone usage features are identified which contribute the most in assessment of smartphone addiction.

4. Results

The dataset is labeled using SAS-SV and evaluated using the real time attributes extracted via android application. Linear classification performance is evaluated using 5-fold cross validation.

Table 2. Classification Performance

	Accuracy	**F1 Score**
Linear SVM	81.33	0.810
Logistic Regression	82.67	0.827

The most important smartphone usage attributes contributing maximum in assessment of smartphone addiction are identified using correlation and information gain. Table 3 and Table 4 present the top five attributes extracted using the correlation evaluator and the information gain evaluator respectively.

Table 3. Top 5 Attributes Identified Using Correlation Evaluator

S. No.	Attributes	Correlation Score
1.	Entertainment Longest Session	0.4129
2.	Communication Total Time	0.3987
3.	Communication Longest Session	0.3496
4.	Work Longest Session	0.348
5.	Entertainment Total Time	0.3475

Table 4. Top 5 Attributes Identified Using Information Gain Evaluator

S. No.	Attributes	IG Score
1.	News and Surfing Longest Session	0.4129
2.	Communication Total Time	0.3987
3.	Communication Longest Session	0.3496
4.	Entertainment Longest Session	0.348
5.	Others Data Usage	0.3475

5. Conclusion

This study proposes a model for assessment of smartphone addiction in university students using the objective measures of real time smartphone usage and supervised machine learning. An android application is developed to collect real time smartphone usage features and linear classification models namely Support Vector Machine and Logistic regression are used to predict smartphone addiction among university students. Smartphone Addiction Scale-Short Version is used to label smartphone addiction among university students. Support Vector Machine predicts smartphone addiction with a superlative accuracy of 81.33% while Logistic Regression gives accuracy of 82.67%. Most contributing features of real time smartphone usage for assessment of smartphone addiction have also been identified using correlation and information gain measures.

The easy availability of smartphones and the Internet and the flexible schedules of millennials including students are leading them to have addictive smartphone usage behaviors [11]. Online education which involves use of digital technologies including smartphones has become useful, yet stressful practice for university students after the COVID-19 pandemic. However, overuse of smartphones and other digital technologies are known to cause fatigue [19] and affect academic performance of students [20]. Therefore, universities should make judicious use of these technologies.

References

[1] Tossell, C., Kortum, P., Shepard, C., Rahmati, A. and Zhong, L., 2015. Exploring smartphone addiction: insights from long-term telemetric behavioral measures. *International Journal of Interactive Mobile Technologies*, *9*(2).

[2] Arora, A., Chakraborty, P. and Bhatia, M.P.S., 2021. Problematic use of digital technologies and its impact on mental health during COVID-19 pandemic: assessment using machine learning. In *Emerging Technologies During the Era of COVID-19 Pandemic* (pp. 197-221). Springer, Cham.

[3] Vernon, L., Modecki, K.L. and Barber, B.L., 2018. Mobile phones in the bedroom: Trajectories of sleep habits and subsequent adolescent psychosocial development. *Child development*, *89*(1), pp.66-77.

[4] Arora, A. and Chakraborty, P., 2020. Diagnosis, prevalence and effects of nomophobia-A review. *Psychiatry Research*, *288*, pp.112975-112975.

[5] Harris, B., Regan, T., Schueler, J. and Fields, S.A., 2020. Problematic mobile phone and smartphone use scales: A systematic review. *Frontiers in Psychology*, *11*, p.672.

[6] Kwon, M., Kim, D.H., Cho, H. and Yang, S., 2013. SAS-SV (also known as: Smartphone Addiction Scale-Short Version)(appears in: The Smartphone Addiction Scale: Development and Validation of a Short Version for Adolescents.) Copyright: Creative Commons License.

[7] Shin, C. and Dey, A.K., 2013, September. Automatically detecting problematic use of smartphones. In *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing* (pp. 335-344).

[8] Lawanont, W. and Inoue, M., 2017, June. A development of classification model for smartphone addiction recognition system based on smartphone usage data. In *International Conference on Intelligent Decision Technologies* (pp. 3-12). Springer, Cham.

[9] Arora, A., Chakraborty, P. and Bhatia, M.P.S., 2021. Intervention of Wearables and Smartphones in Real Time Monitoring of Sleep and Behavioral Health: An Assessment Using Adaptive Neuro-Fuzzy Technique. *Arabian Journal for Science and Engineering*, pp.1-26.

[10] Ellis, D.A., Davidson, B.I., Shaw, H. and Geyer, K., 2019. Do smartphone usage scales predict behavior?. *International Journal of Human-Computer Studies*, *130*, pp.86-92.

[11] Arora, A., Chakraborty, P. and Bhatia, M.P.S., 2021. Real Time Smartphone Data for Prediction of Nomophobia Severity using Supervised Machine Learning. *AIJR Proceedings*, pp.78-83.

[12] Elhai, J.D., Yang, H., Rozgonjuk, D. and Montag, C., 2020. Using machine learning to model problematic smartphone use severity: The significant role of fear of missing out. *Addictive behaviors*, *103*, p.106261.

[13] Harris, B., McCredie, M. and Fields, S., 2020. Examining the psychometric properties of the smartphone addiction scale and its short version for use with emerging adults in the US. *Computers in Human Behavior Reports*, *1*, p.100011.

[14] Dreiseitl, S. and Ohno-Machado, L., 2002. Logistic regression and artificial neural network classification models: a methodology review. *Journal of biomedical informatics*, *35*(5-6), pp.352-359.
[15] Kumar, A. and Arora, A., 2019, February. A filter-wrapper based feature selection for optimized website quality prediction. In *2019 Amity International Conference on Artificial Intelligence (AICAI)* (pp. 284-291). IEEE.
[16] Gnanambal, S., Thangaraj, M., Meenatchi, V.T. and Gayathri, V., 2018. Classification algorithms with attribute selection: an evaluation study using WEKA. *International Journal of Advanced Networking and Applications*, *9*(6), pp.3640-3644.
[17] Sugianela, Y. and Ahmad, T., 2020, February. Pearson correlation attribute evaluation-based feature selection for intrusion detection system. In *2020 International Conference on Smart Technology and Applications (ICoSTA)* (pp. 1-5). IEEE.
[18] Win, T.Z. and Kham, N.S.M., 2019, February. Information gain measured feature selection to reduce high dimensional data. In *Seventeenth International Conference on Computer Applications (ICCA 2019)* (Vol. 68, No. 73, pp. 1-5).
[19] Shin, J. and Shin, M., 2016. To be connected or not to be connected? Mobile messenger overload, fatigue, and mobile shunning. *Cyberpsychology, Behavior, and Social Networking*, *19*(10), pp.579-586.
[20] Liu, X., Luo, Y., Liu, Z. Z., Yang, Y., Liu, J., & Jia, C. X. (2020). Prolonged mobile phone use is associated with poor academic performance in adolescents. *Cyberpsychology, Behavior, and Social Networking*, *23*(5), 303-311.

Authors' Profiles

Anshika Arora received her B.Tech. in computer science from the University of Delhi and M.Tech. in software engineering (gold medallist) from Delhi Technological University. She is currently pursuing Ph.D. at Netaji Subhas University of Technology, New Delhi. Her research interests include behavioral health analysis, web analytics using soft computing techniques.

Pinaki Chakraborty is an assistant professor at the Department of Computer Science and Engineering, Netaji Subhas University of Technology, New Delhi. He received his B.Tech. from Indraprastha University, and his M.Tech. and Ph.D. from Jawaharlal Nehru University. He has published more than 90 research papers in journals and conference of repute. His area or research includes systems software and educational and societal use of software.

M.P.S. Bhatia is professor in Department of Computer Science and Engineering, Netaji Subhas University of Technology. He is Ph.D. in computer engineering with more than twenty years of research experience. His area of research includes data mining, machine learning, cyber security, social network analysis, computer vision, smart health and big data analytics.

Aditya Puri is pursuing B.E. in electronics and communication engineering at Netaji Subhas University of Technology. His interests lie in android development and Java backend development.

How to cite this paper: Anshika Arora, Pinaki Chakraborty, M.P.S. Bhatia, Aditya Puri, "Intelligent Model for Smartphone Addiction Assessment in University Students using Android Application and Smartphone Addiction Scale", International Journal of Education and Management Engineering (IJEME), Vol.13, No.1, pp. 29-34, 2023. DOI:10.5815/ijeme.2023.01.04

Intrusion detection based on machine learning and statistical feature ranking techniques

Akhil Kumar
Department of Computer Science & Engineering
Delhi Technological University
Delhi, India
akhildtu7@gmail.com

Shailender Kumar
Department of Computer Science & Engineering
Delhi Technological University
Delhi, India
shailenderkumar@dce.ac.in

2023 13th International Conference on Cloud Computing, Data Science & Engineering (Confluence) | 978-1-6654-6263-1/23/$31.00 ©2023 IEEE | DOI: 10.1109/Confluence56041.2023.10048802

***Abstract*—Big data has made it easier for people to live an information-based Internet lifestyle, but it has also created several serious network security issues that make it difficult to use networks regularly. At the moment, intrusion detection systems are mostly used to identify aberrant network traffic. So, utilizing a Machine Learning (ML) technique, we suggested a statistical analysis-based intrusion detection system. In this paper, we suggested an intrusion detection system using the statistical analysis of the ranking method known as the T-test: two samples assuming unequal variances. A substantial amount of network traffic data that includes both malware data and normal traffic data is gathered to be able to identify the pattern of the malware data. Nine different traffic features for both intrusion and normal traffic are extracted and further ranked using the t-test, yielding nine "t" values from which various features were derived. The Naive Bayes machine learning algorithm will then be applied to the 9 features, deleting one feature at a time that has the lowest "t" value to provide 9 alternative accuracy values. After examining the accuracy value, we get to the conclusion that the two features with the lowest value are removed to attain the highest accuracy, i.e., 95.69% with the top 7 features rather than using all 9 features.**

***Index Terms*—Intrusion Detection, T-Test, Naive Bayes, Feature Ranking**

I. Introduction

Intrusion detection systems (IDS) are important for network security because they can identify and stop malicious activities. A key component of reliable services in extended network environments like the Internet of Things and big data is the network intrusion detection system, a key network security technology that watches packets for potentially dangerous events occurring in the network. Big data has made it easier for people to live an information-based Internet lifestyle, but it has also created several serious network security issues that make it difficult to use networks regularly. At the moment, intrusion detection systems are mostly used to identify aberrant network traffic. The academic community has been concentrating on identifying network assaults that target information and communication systems for years. A complex issue, network intrusion detection poses a variety of difficulties. While novel attacks appear as a result of the growth of connected devices and the advancement of communication technology, many current attacks go unnoticed. Systems for detecting network intrusions are crucial for safeguarding modern communication networks. However, a variety of out-of-date datasets and a wide range of different evaluation measures are employed in the literature to demonstrate algorithm effectiveness. In recent years, network dangers and hazards have been developing quickly. Networks are protected by a variety of technologies (including firewalls, anti-virus, anti-malware, and spam filters). A robust and efficient network security solution for spotting unwanted and unusual network traffic flow is an intrusion detection system (IDS). To lessen the impact of invaders, the study issue of intrusion detection and prevention system (IDPS) has attracted attention.

In the literature, several study projects have been suggested for network traffic-based intrusion detection such as [1], [2], [3], [4], [5], and many others. However, due to the inclusion of irrelevant features in their study, they report a significant false positive rate. In this research, we propose a Statistical Analysis Based Intrusion Detection System by employing Machine Learning (ML) approach to rank the traffic data and eliminate the irrelevant features, which may limit the detection accuracy. A substantial amount of network traffic data that includes both malware and normal data is gathered to identify the patterns of normal and malware data. Raw data is filtered to remove metadata and TCP flows are extracted from each captured file. Further, to rank the features, we applied the T-test with two samples assuming unequal variances on nine traffic features that comprise both malware and regular data, producing nine "t" values, one for each feature. We ranked the features based on this obtained "t" value. Such a ranking of traffic features helps us in eliminating irrelevant features while testing. Then Naive Bayes machine learning algorithm was applied to the testing data with 9 features by deleting each feature individually (removing from the last ranked feature), yielding nine distinct accuracy scores. After examining the accuracy value, we get to the conclusion that the three features with the lowest "t" value are removed to attain the highest accuracy, with the accuracy of the data being increased as we removed these lower-ranked features.

The remainder of this article is organized as follows. In section II, we review the pertinent research. With a preliminary review of the datasets under consideration, we give a thorough explanation of the suggested methodology in Section III. In Section IV, we present the results obtained with the proposed

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 01,2023 at 07:25:53 UTC from IEEE Xplore. Restrictions apply.

algorithm using the Naive Bayes machine learning algorithm performed on the data sets. Finally, in Section V, we conclude our paper with future work directions.

II. Related Work

This section examines the related work of intrusion detection. Several works have been reported in the literature for intrusion detection on mobile platforms such as [24], [25], [26], [27], [28], [29], and [30]. Out of these, works like [24], [26], and [31] focused on capturing network traffic of Android samples and found the distinguishing features, i.e., the features that can efficiently detect malicious Android network traffic. Authors in these works applied machine learning techniques on network traffic features to detect malicious network activity in smartphones. However, our work is based on desktop-based intrusion detection, hence, in this section; we focus in detail on desktop-based intrusion detection.

To ascertain the network's invasion behavior, the authors in [1] ran a sample test on KDDTest + dataset and their method had a high detection rate using LeNet-5 and DBN algorithms. The Energy-based Flow Classifier is a revolutionary technique that Pontes et al. [2] proposed in their work. Inferring a statistical model from labeled innocuous samples using inverse statistics is how this anomaly-based classifier works. The researchers demonstrated that EFC is more adaptable to a wide range of data distributions than conventional ML-based classifiers. A dataset gathered from 20 widely dispersed networks over the period of 60 days was utilized to compare FL with their [3] technique utilizing the weighted precision, retention, and F1 scores. They demonstrated that Segmented-FL outperforms in all three categories of intrusion detection when compared to Segmented-optimized FL. Bertoli et al [4] analyzed the AB-TRAP in both local (LAN) and global (internet) domains to prevent TCP port scanning attacks. For the LAN study scenario, a decision tree in kernel space with low CPU and RAM consumption achieved an area under the f1-score of 0.96. The authors in [5] discovered through sample studies that even at low sampling rates like 1/10 and 1/100, malicious flows with smaller sizes (i.e., number of packets) are likely to go undetected. Next, they looked at the effects of different sampling approaches on the NIDS detection rate.

Kim et al [6] demonstrated that network incursion outside the area of the learned data in the feature space can evade the ML-NIDS by examining the learning characteristics using representative features. Early classification of the active session by design, before it leaves the training dataset of the ML-detection NIDS's range, can successfully stop this from happening. By examining the link load behaviors of the Road Side Unit (RSU) in the IoV against various assaults that cause abnormal variations in traffic flows, the authors in [7] created data-driven IDS. To extract link load features and identify intrusions targeting RSUs, a deep learning architecture based on the Convolutional Neural Network (CNN) was used. He et al. [8] here proposed that the real procedure of the current traffic intrusion detection systems has to be improved due to their numerous flaws and high resource occupation rate. The classic traffic detection technique is improved in this study using deep learning and a clustering algorithm. The data were first preprocessed, then 62 features were extracted, and then Zhou et al [9] suggested a C4.5 divided algorithm to detect traffic. In the study by Ageyev et al. [10], datasets were categorized into categories such as network traffic-based feature sets, online traffic-based datasets, cloud-based network datasets, and IoT traffic-based datasets.

As part of the early development of an intrusion detection system (IDS), this study [11] used the Deep Learning Machine and Artificial Immune System to identify anomalies in the smart home system (AIS-ELM). The input parameters were optimized by AIS using the clonal method. Hashemi et al [12] provided Reconstruction from Partial Observation (RePO), a novel method for developing a network intrusion detection system (NIDS) that uses denoising autoencoders. To assess the model's efficacy, the authors [13] used the CSE-CIC-IDS2018 dataset. According to their experimental findings, the ACNNBN model can significantly improve detection accuracy. Raw traffic is transmitted to the machine learning classifier for examination and the detection of potential threats, according to Rose et al's [14] presentation. On the Cyber-Trust testbed, the effectiveness of the suggested methodology is evaluated using both legitimate and malicious network data. Raju et al. [15] conducted a review using three deep-learning models and a semi-balanced version of the CIC-IDS2017 and CSE-CIC-IDS2018 datasets. They designed two new assessment metrics—Matthews Correlation Coefficient and Cohen's Kappa Coefficient—to assess how well the used systems functioned.

The authors here [16] review the literature and offered additional context on either deep learning or machine learning algorithms on intrusion detection systems. Machine learning [ML] technologies were being used increasingly in circumstances where it is challenging to understand the process of interest logically, according to a study proposed by the authors [17]. A novel network-based intrusion detection model on convolution neural networks was provided by the authors in [18]. The ACNNBN has a novel feature called the Convolution Block Attention Module that allows cross-channel fusion with convolution modules. The primary purpose of this [19] work was to create anomaly-based intrusion detection systems using stacking and boosting ensemble approaches. Qadeer et al [20] here by using an existing IDS, SNORT, an efficient multicore approach for network monitoring has been put forth. Additionally, the suggested architecture used a PF ring for effective packet capturing and IP hash load balancing for network load balancing.

Baig et al [21] suggested a method for distinguishing between regular and abnormal network traffic based on GMDH. On the KDD 99 dataset, two variations of the technique—monolithic and ensemble-based—were put to the test. The dataset underwent pre-processing, and all features were ranked using the Information Gain, Gain Ratio, and GMDH by itself feature ranking approaches. Results showed that, when compared to previous intelligent classification algorithms for network intrusion detection, the proposed intrusion detection

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 01,2023 at 07:25:53 UTC from IEEE Xplore. Restrictions apply.

methodology yields high attack detection rates, close to 98 percent.

Finding the most relevant and practical characteristics that can be used as key features in a brand-new IDS dataset was the aim of this [22] study. To achieve the objective, a method for constructing an optimal IDS ensemble is created. The following six feature selection techniques are utilized and contrasted: Chi-Square, Relief-F (R-F), Symmetrical Uncertainty (SU), One-R (OR), and Gain Ratio (GR) are all metrics for information gain (CS). Collections of selected characteristics are produced by the feature selection methods. The best number of features from each feature selection methodology's feature ranking stage would be used to categorize assaults using the following four traditional classification methods: Decision Tree: J48, Naive Bayesian (NB), SOM, and Bayesian Network (BN). The best features from each feature selection methodology and the best features from each classification methodology are merged to create ensemble IDSs. Finally, the Hold-up, K-fold, F-Measure, and statistical validation procedures are used to analyze the ensemble IDSs. The efficient ensemble IDSs using (SU and BN), (CS and BN), (CS and SOM), (IG and NB), and (OR and BN) with respective ten, four, and seven better-selected features achieve 81.0316 percent, 85.2593 percent, and 80.8625 percent of accuracy, according to experimental results using Weka tools on the ITD-UTM dataset.

NSL-KDD was used to assess the machine learning techniques for intrusion detection [23]. Nevertheless, not all traits aid in efficiency in large datasets. A particular feature set may be lowered or selected in order to improve speed and accuracy. As a result, the feature selection process employs recursive feature elimination (RFE). The Intrusion Detection System (IDS), which employs the machine learning techniques Random Forest and Support Vector Machine, was thoroughly examined by them (SVM). They had demonstrated how it is possible to compare the model's performance before and after Random Forest and SVM feature selection. They had also offered the confusion matrices.

The authors in [32] proposed an anomaly detection method based on features importance. They applied Principal Component Analysis, and Clustering based approaches on the BETH dataset to efficiently detect anomalies. Raman et al. [33] applied several feature selection techniques such as Pearson Correlation, Information Gain, ExtraTreeClassifier, and Chi-Square tests to rank the features and further applied machine learning techniques for intrusion detection. Yash et al. [34] applied ANOVA and Chi-Square Test to rank the traffic features, however, the authors did not propose a detection model. In this work, we have applied Student's T-Test for feature ranking and we have also proposed an intrusion detection model. The details of the proposed work are presented in next section.

III. PROPOSED METHODOLOGY

Figure 1 summarizes the proposed methodology of this work. The work is divided in sub phases and we have discussed each phase in detail in the upcoming subsections.

A. Traffic Capturing

Using the Wireshark software that is installed on our PC, we were able to collect the typical normal network traffic. Additionally, the Canadian Institute for Cyber Security provided us with the intrusion traffic. We recovered TCP flows from each of the pcap files containing the traffic that was recorded. Our research study employed about 35,000 normal traffic flows and 30,000 malware traffic flows. Nine traffic features were taken from each traffic flow, and their specifics are provided in the following section.

B. Feature Extraction

Feature extraction is a process for transforming unprocessed data into usable numerical features while keeping the data from the original data set. It yields better results in comparison to performing machine learning on the raw data directly. The fundamental objective is to use fewer features to collect the same information. In order to lessen overfitting in the model, feature selection is a technique for decreasing the input variable for the model by selecting only pertinent data. Nine traffic features that we extracted are described in Table I below.

TABLE I
LIST OF TRAFFIC FEATURES USED ALONG WITH THEIR NOTATIONS

S. No.	Feature	Notation
1	Average packet size	F1
2	Average packet size sent	F2
3	Average packet size received	F3
4	Average time interval between packets sent	F4
5	Average time interval between packets received	F5
6	Ratio of incoming to outgoing packets	F6
7	Ratio of incoming to outgoing bytes	F7
8	Average flow duration	F8
9	Average number of destinations per connections	F9

C. Feature Ranking Using T-Test

On data sets from two separate populations with different variances, this program performs a two-sample student's t-test. Depending on whether we're testing whether the two population means differ or if one is higher than the other, this test can be either two-tailed or one-tailed. When the samples are normally distributed, the standard deviations of the two populations are unknown and assumed to be unequal, and the sample size is big enough, a two-sample T-test with unequal variance can be used (over 30). Hence, the t-test with unequal variance is calculated. In both t-tests, a "t" value and confidence interval are reported. The 10 different p values are obtained after the t-test on the 10 data set. The data set will be categorized in decreasing order after analysis to determine the p-value. The various data set is designated as F1 to F9. They will be sorted according to the "t" value obtained after applying the t-test, in descending order.

Suppose X and Y, as summarized in the table II, represent the values of any traffic feature "F" for normal and intrusion traffic respectively, then the Formula used for the T-test is given below:

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 01,2023 at 07:25:53 UTC from IEEE Xplore. Restrictions apply.

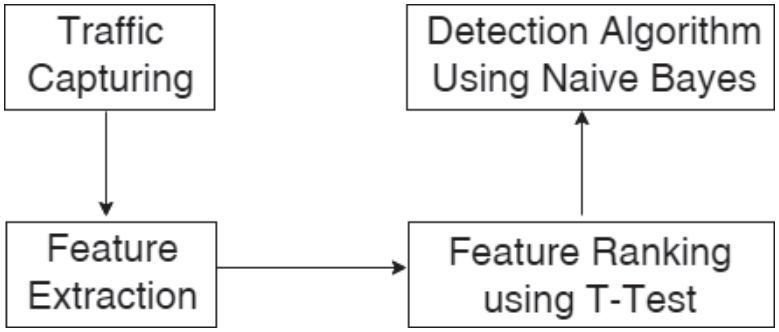


Fig. 1. Proposed Methodology

TABLE II
EXAMPLE OF TRAFFIC FEATURE VALUE

S. No.	X	Y
1	2	1
2	0	9
3	6	5

$$T = \frac{\frac{\sum D}{N}}{\sqrt{\frac{\sum D^2 - \frac{(\sum D)^2}{N}}{(N-1)(N)}}} \quad (1)$$

D. Detection Algorithm Using Naive Bayes

A probabilistic classifier is the NAÏVE BAYES algorithm for classification. It is based on probability models that make substantial assumptions about independence. Naive Bayes is a machine learning model that is used for vast amounts of data; it is the suggested strategy even when working with data that has millions of data records. It does NLP tasks like sentimental analysis with excellent results. On the 9 features data set that we recovered from the raw network traffic, we applied the NAÏVE BAYES algorithm. The nine features are subjected to the NAÏVE BAYES machine learning method, and each feature will be eliminated one at a time based on its ranking. We start removing the feature from the last, i.e., the lowest ranked feature will be eliminated and we test for detection accuracy. We repeat this procedure till we get the highest detection accuracy. We report the results obtained from the proposed approach in the next section.

IV. RESULTS AND DISCUSSION

In this section, we emphasize the outcomes of the proposed approach that was used to implement the statistical analysis-based intrusion detection system. We applied the T-test: two sample assuming unequal variances statistical analysis to rank the features. To identify the pattern of the malware data, a sizable volume of network traffic data that contains both malware data and regular data is collected. The t-test is run on 9 different types of data sets. Following the application of the t-test to nine distinct data sets that comprise both malware and normal data, nine t-values were created. Table III shows the feature ranking based on the decreasing "t" value. The higher the "t" value, the is better the feature with distinguishing capability. As can be seen from table III, the feature F3, i.e., Average packet size received is the most distinguishing feature between normal traffic and intrusion traffic. Next to F3 is the feature F2, i.e., the Average packet size sent is second in the ranking. The worst or most irrelevant feature turns out to be the Ratio of Incoming to Outgoing Bytes with the least "t' value.

TABLE III
FEATURE RANKING BASED ON T-VALUE OF FEATURES

S. No.	Feature	T-Value
1	F3	115.2452
2	F2	112.7775
3	F5	93.47604
4	F4	84.80577
5	F9	76.41087
6	F8	58.26489
7	F1	49.03203
8	F6	10.7755
9	F7	10.23431

TABLE IV
DETECTION ACCURACY AT DIFFERENT ITERATIONS

Feature Set	Naive Bayes Accuracy
All 9 features	0.7613846153846154
Top 8 features	0.7896410256410257
Top 7 features	0.9569230769230769
Top 6 features	0.9456410256410256
Top 5 features	0.9406666666666667
Top 4 features	0.9428717948717948
Top 3 features	0.9424102564102564
Top 2 features	0.9320512820512821
Top 1 features	0.9023589743589744

Further, we applied our proposed approach using the NAÏVE BAYES machine learning method to 9 data sets, it produces 9 different accuracy values by removing 1 feature at a time. We removed the last ranked feature, checked the accuracy, and if the accuracy increased, we permanently remove that feature. Then we proceed toward the next lower-ranked feature and the process continues. Table IV summarizes

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 01,2023 at 07:25:53 UTC from IEEE Xplore. Restrictions apply.

these detection results for different iterations. As can be seen from Table IV, the top 7 features give us the highest detection accuracy of around 95.69%. If we further remove the features, the accuracy decreases. Hence, we can say that we get the best accuracy of 95.69%. We can argue that feature ranking helps us in improving detection accuracy. As can be seen from Table IV, if do not apply feature ranking and use all 9 features for detection, we get an accuracy of around 76%. And it increases significantly to 95.69% on removing the lower two ranked features. Hence, feature ranking using the T-test helps us in improving detection accuracy. The accuracy number will be produced after the NAÏVE BAYES theorem has been applied to the data sets. After removing each feature individually, the various accuracy values are shown in table IV. So, as we can see in table IV when the NAÏVE BAYES algorithm is applied to all 9 characteristics, the accuracy value is lowest. The accuracy value decreases after 7 features, so based on this outcome, we can conclude that the maximum accuracy is provided when all 7 features are present.

A. Comparison with other works

In this subsection, we compare our proposed approach with other similar works in the field of network traffic based intrusion detection. Table V summarizes the comparison of our work with other similar works. As can be seen from the Table V, our work outperforms several other similar works proposed to detect intrusions based upon network traffic.

TABLE V
COMPARISON OF PROPOSED WORK WITH OTHER SIMILAR WORKS

Related Work	Detection Accuracy
Hsieh et al. [35]	79%
Chen et al. [36]	93.56%
Shah et al. [37]	95%
Chindove et al. [38]	90%
PROPOSED WORK	**95.69**%

V. CONCLUSION AND FUTURE WORK

In this work, we proposed an intrusion detection system based on the two sample assumption of unequal variances T test, a statistical analysis of ranking approach. In order to determine the pattern of the malware data, a sizable amount of network traffic data that consists of both intrusion and normal data was obtained. The T-test was used to score nine different traffic features that were retrieved from both intrusion and normal traffic files. Based upon the "t" value obtained from the T-test, features were ranked. The Naive Bayes machine learning algorithm was next used on the nine features, eliminating the ones with the lowest t-values one at a time to provide nine different accuracy values. After looking at the accuracy value, we got to the conclusion that in order to achieve the best accuracy, the two features with the lowest value should be removed. Instead of using all nine features, our work's accuracy rate of 95.69% was attained using the top 7 features. In our future work, we will aim to increase the number of traffic features for analysis and detection.

VI. ACKNOWLEDGEMENT

We would like to thank the team from Canadian Institute for Cyber Security, for providing us the intrusion network trafic for our research.

REFERENCES

[1] H. Yang and F. Wang, "Wireless Network Intrusion Detection Based on Improved Convolutional Neural Network," in IEEE Access, vol. 7, pp. 64366-64374, 2019

[2] C. F. T. Pontes, M. M. C. de Souza, J. J. C. Gondim, M. Bishop and M. A. Marotta, "A New Method for Flow-Based Network Intrusion Detection Using the Inverse Potts Model," in IEEE Transactions on Network and Service Management, vol. 18, no. 2, pp. 1125-1136, June 2021.

[3] Y. Sun, H. Esaki and H. Ochiai, "Adaptive Intrusion Detection in the Networking of Large-Scale LANs With Segmented Federated Learning," in IEEE Open Journal of the Communications Society, vol. 2, pp. 102-112, 2021

[4] G. De Carvalho Bertoli et al., "An End-to-End Framework for Machine Learning-Based Network Intrusion Detection System," in IEEE Access, vol. 9, pp. 106790-106805, 2021.

[5] J. Alikhanov, R. Jang, M. Abuhamad, D. Mohaisen, D. Nyang and Y. Noh, "Investigating the Effect of Traffic Sampling on Machine Learning-Based Network Intrusion Detection Approaches," in IEEE Access, vol. 10, pp. 5801-5823, 2022.

[6] T. Kim and W. Pak, "Robust Network Intrusion Detection System Based on Machine-Learning With Early Classification," in IEEE Access, vol. 10, pp. 10754-10767, 2022.

[7] L. Nie, Z. Ning, X. Wang, X. Hu, J. Cheng and Y. Li, "Data-Driven Intrusion Detection for Intelligent Internet of Vehicles: A Deep Convolutional Neural Network-Based Method," in IEEE Transactions on Network Science and Engineering, vol. 7, no. 4, pp. 2219-2230, 1 Oct.-Dec. 2020.

[8] Q. He, "Research on Network Traffic Anomaly Detection Based on Deep Learning," 2021 International Conference on Networking, Communications and Information Technology (NetCIT), 2021, pp. 50-53.

[9] J. Zhou, X. Jiang, C. Liu, J. Zhang, L. Liao and J. Lu, "Multi-Traffic Features Network Intrusion Detection Algorithm Based on C4.5," 2021 18th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP), 2021, pp. 548-552.

[10] D. Ageyev, T. Radivilova, O. Bondarenko and O. Mohammed, "Data Sets Selection for Distributed Infocommunication Networks Traffic Abnormality Detection," 2021 IEEE 8th International Conference on Problems of Infocommunications, Science and Technology (PIC S&T), 2021, pp. 635-638

[11] E. D. Alalade, "Intrusion Detection System in Smart Home Network Using Artificial Immune System and Extreme Learning Machine Hybrid Approach," 2020 IEEE 6th World Forum on Internet of Things (WF-IoT), 2020, pp. 1-2.

[12] M. J. Hashemi and E. Keller, "Enhancing Robustness Against Adversarial Examples in Network Intrusion Detection Systems," 2020 IEEE Conference on Network Function Virtualization and Software Defined Networks (NFV-SDN), 2020, pp. 37-43.

[13] L. Li, J. Mu, H. He and C. Liu, "An Attention-Based CNN with Batch Normalization Model for Network Intrusion Detection," 2021 China Automation Congress (CAC), 2021, pp. 3531-3536.

[14] J. R. Rose, M. Swann, G. Bendiab, S. Shiaeles and N. Kolokotronis, "Intrusion Detection using Network Traffic Profiling and Machine Learning for IoT," 2021 IEEE 7th International Conference on Network Softwarization (NetSoft), 2021, pp. 409-415.

[15] D. Raju, S. Sawai, S. Gavel and A. S. Raghuvanshi, "Development of Anomaly-Based Intrusion Detection Scheme Using Deep Learning in Data Network," 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT), 2021.

[16] J. A. Abraham and V. R. Bindu, "Intrusion Detection and Prevention in Networks Using Machine Learning and Deep Learning Approaches: A Review," 2021 International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA), 2021, pp. 1-4.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 01,2023 at 07:25:53 UTC from IEEE Xplore. Restrictions apply.

[17] S. Sridevi, R. Prabha, K. N. Reddy, K. M. Monica, G. A. Senthil and M. Razmah, "Network Intrusion Detection System using Supervised Learning based Voting Classifier," 2022 International Conference on Communication, Computing and Internet of Things (IC3IoT), 2022, pp. 01-06.
[18] L. Li, J. Mu, H. He and C. Liu, "An Attention-Based CNN with Batch Normalization Model for Network Intrusion Detection," 2021 China Automation Congress (CAC), 2021, pp. 3531-3536.
[19] V. Sidharth and C. R. Kavitha, "Network Intrusion Detection System Using Stacking and Boosting Ensemble Methods," 2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA), 2021, pp. 357-363.
[20] H. Qadeer, A. Talat, K. N. Qureshi, F. Bashir and N. Ul Islam, "Towards an Efficient Intrusion Detection System for High Speed Networks," 2020 17th International Bhurban Conference on Applied Sciences and Technology (IBCAST), 2020, pp. 428-433.
[21] Zubair A. Baig, Sadiq M. Sait, AbdulRahman Shaheen, "GMDH-based networks for intelligent intrusion detection", Engineering Applications of Artificial Intelligence, Volume 26, Issue 7,2013, Pages 1731-1740, ISSN 0952-1976.
[22] D. Stiawan et al., "An Approach for Optimizing Ensemble Intrusion Detection Systems," in IEEE Access, vol. 9, pp. 6930-6947, 2021.
[23] R. Patgiri, U. Varshney, T. Akutota and R. Kunde, "An Investigation on Intrusion Detection System Using Machine Learning," 2018 IEEE Symposium Series on Computational Intelligence (SSCI), 2018, pp. 1684-1691.
[24] A. Arora, S. Garg and S. K. Peddoju, "Malware Detection Using Network Traffic Analysis in Android Based Mobile Devices," Eighth International Conference on Next Generation Mobile Apps, Services and Technologies, pp. 66-71, 2014.
[25] A. Arora, S. K. Peddoju and M. Conti, "PermPair: Android Malware Detection Using Permission Pairs," in IEEE Transactions on Information Forensics and Security, vol. 15, pp. 1968-1982, 2020.
[26] A. Arora, and S.K. Peddoju, "Minimizing Network Traffic Features for Android Mobile Malware Detection", in Proceedings of the 18th International Conference on Distributed Computing and Networking (ICDCN), article 32, pp. 1-10, 2017.
[27] A. Arora and S. K. Peddoju, "NTPDroid: A Hybrid Android Malware Detector Using Network Traffic and System Permissions," 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/ 12th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE), pp. 808-813, 2018.
[28] A. Arora, S.K. Peddoju, V. Chouhan, and A. Chaudhary, "Hyrbid Android Malware Detection by Combining Supervised and Unsupervised Learning", in Proceedings of the 24th Annual International Conference on Mobile Computing and Networking, 2018.
[29] K. Khariwal, J. Singh and A. Arora, "IPDroid: Android Malware Detection using Intents and Permissions," Fourth World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4), 2020, pp. 197-202, 2020.
[30] M. Upadhayay, A. Sharma, G. Garg and A. Arora, "RPNDroid: Android Malware Detection using Ranked Permissions and Network Traffic," Fifth World Conference on Smart Trends in Systems Security and Sustainability (WorldS4), pp. 19-24, 2021.
[31] A. Shabtai, L. Tenenboim-Chekina, D. Mimran, L. Rokach, B. Shapira, and Y. Elovici, Mobile malware detection through analysis of deviations in application network behavior. Computers and Security, 43, pp. 1 – 18, 2014.
[32] N. Sushmakar, N. Oberoi, S. Gupta and A. Arora, "An Unsupervised Based Enhanced Anomaly Detection Model Using Features Importance," 2022 2nd International Conference on Intelligent Technologies (CONIT), 2022, pp. 1-7, doi: 10.1109/CONIT55038.2022.9848297.
[33] Raman, S. K. Jha and A. Arora, "An Enhanced Intrusion Detection System Using Combinational Feature Ranking and Machine Learning Algorithms," 2022 2nd International Conference on Intelligent Technologies (CONIT), 2022, pp. 1-8, doi: 10.1109/CONIT55038.2022.9847815.
[34] Y. Sharma, S. Sharma and A. Arora, "Feature Ranking using Statistical Techniques for Computer Networks Intrusion Detection," 2022 7th International Conference on Communication and Electronics Systems (ICCES), 2022, pp. 761-765, doi: 10.1109/ICCES54183.2022.9835831.
[35] C. -F. Hsieh and C. -M. Su, "MLNN: A Novel Network Intrusion Detection Based on Multilayer Neural Network," 2021 International Conference on Technologies and Applications of Artificial Intelligence (TAAI), 2021, pp. 43-48, doi: 10.1109/TAAI54685.2021.00017.
[36] J. Chen, S. Yin, S. Cai, L. Zhao and S. Wang, "L-KPCA: an efficient feature extraction method for network intrusion detection," 2021 17th International Conference on Mobility, Sensing and Networking (MSN), 2021, pp. 683-684, doi: 10.1109/MSN53354.2021.00104.
[37] A. Shah, S. Clachar, M. Minimair and D. Cook, "Building Multiclass Classification Baselines for Anomaly-based Network Intrusion Detection Systems," 2020 IEEE 7th International Conference on Data Science and Advanced Analytics (DSAA), 2020, pp. 759-760, doi: 10.1109/DSAA49011.2020.00102.
[38] H. Chindove and D. Brown, "Adaptive Machine Learning Based Network Intrusion Detection", in Proceedings of the International Conference on Artificial Intelligence and its Applications (icARTi '21). Association for Computing Machinery, New York, NY, USA, Article 15, 1–6. https://doi.org/10.1145/3487923.3487938, 2021.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 01,2023 at 07:25:53 UTC from IEEE Xplore. Restrictions apply.

Wireless Personal Communications
https://doi.org/10.1007/s11277-023-10170-4

LDML: A Proposal to Reduce Leakage Power in DML Circuits

Neetika Yadav[1,2] · **Neeta Pandey**[1] · **Deva Nand**[1]

Accepted: 4 February 2023
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

Abstract
This paper puts forward a proposal to reduce leakage power in dual mode logic (DML) circuits by adapting LECTOR (LEakage Control TransistOR) technique and the circuits so obtained are referred to as L-DML circuits. The concept is elucidated through footed Type A and Type B DML based gates which are called LDML-TA and LDML-TB. The leakage power for two and four input NAND and NOR gates implemented through DML and LDML circuits is measured in static and dynamic mode through simulative investigations at 90 nm and 45 nm nodes using the Symica DE tool. The performance of the proposal is investigated using standard and high threshold transistor variants in LECTOR technique. The dependence of leakage power on temperature is also studied for both DML and LDML circuits. It is observed that the leakage power of the circuits exhibits an upward trend with lowering of technology node and increase in temperature irrespective of the mode of operation. In static mode, the maximum power saving for LDML-TA is 58.9% at 27 °C whereas the corresponding saving in LDML-TB is 66.6% for 2-input design. Similarly, the LDML-TA and LDML-TB show power saving of (28.5%, 58.9%) and (32.5%, 66.6%) in (pre-charge, evaluate) states respectively in dynamic mode.

Keywords Dual mode logic · Dynamic · LECTOR · LCT · Leakage · Static

1 Introduction

Technological advancements in process technology have led to steady and unparalleled developments in integrated circuit design in past decade [1–4]. The device density in ICs has increased multi-fold and portable systems, in particular, have limited power

✉ Deva Nand
devkamboj07@gmail.com

Neetika Yadav
yadav.neetika92@gmail.com

Neeta Pandey
n66pandey@rediffmail.com

[1] Department of Electronics and Communication Engineering, Delhi Technological University, Delhi 110042, India

[2] Department of Electronics and Communication Engineering, Amity School of Engineering and Technology (Affiliated to GGSIPU), Sector 125, Noida 201303, India

resources. Therefore, power consumption in such systems is of utmost importance and necessitate the efforts to curb it. Traditionally, the switching power dominates the overall power consumption of the circuit. At lower technology nodes, the leakage current which flows through off device show its dominance and makes leakage power significant. The researchers have proposed various techniques to address this issue in static and dynamic CMOS circuits [5–18]. The leakage reduction techniques can broadly be classified into three methods. The first methodology [5, 6] involves the identification of transistors for different threshold voltage assignment in the design. Transistors in non-critical path are assigned high-threshold voltage for leakage reduction and low threshold voltage is assigned to transistors in critical path. The second method [7–10] is based on the usage of sleep transistors with additional overhead of sleep signal generation and control. The third method [11–17] is based on insertion of extra transistors between PDN and PUN, thereby increasing stacking of transistors between supply and ground. It is self-controlled technique and does not require identification of critical path transistors for threshold assignment [5, 6] or additional control circuitry required in [7–10]. Dual mode logic (DML) is a promising logic family, proposed in the last decade, for energy-delay optimization at gate level [19]. The DML has two modes of operation, i.e. Static and Dynamic mode, that can be achieved through a mode signal input [19]. Static mode of operation exhibits low energy and high performance is achieved by the dynamic mode. It mainly consists of a static logic family with pre-charge/discharge transistor which could either be NMOS or PMOS depending on the type of topology [19]. The researchers have presented—analysis and design of DML [19–23], synthesis of DML [24], carry look-ahead adder [25], carry save multiplier [26], logical effort in DML circuits [27] and performance of DML in FD-SOI [28–30]. Further to enhance the energy saving capability of DML, leakage minimization techniques can be incorporated in DML. The power gating and multi-Vth techniques [31–35] are used in DML designs to reduce leakage power. The power gating techniques namely sleep, sleepy stack and dual sleep are applied in [31–34] wherein external signal(s) is used to reduce power. The multi-Vth technique [35], on the other hand relies on using multi-threshold (different threshold transistors) design to achieve power minimization. It is worth mentioning that the available leakage reduction techniques either use an external signal [31–34] that will lead to increase routing or require search to identify transistors which will be subjected to threshold modification [35]. The objective of this paper is to incorporate a leakage reduction technique in DML that does not require external signal or search to identify transistor for threshold modification. The leakage control technique, LECTOR (LEakage Control TransistOR) is a self-controlled technique [11] and has been adapted in DML.

The paper is organised as follows: Sect. 2 discusses the architecture and functionality of dual mode logic designs and concept of leakage in DML logic is also included in this section. LECTOR incorporated DML logic (LDML) is proposed in Sect. 3 which first discusses LECTOR operation and the concept of LECTOR in DML is further discussed in both static and dynamic mode of the design. Simulation results for 2-input LECTOR incorporated DML Type A NOR gate (LDML-TA-NOR2) and NAND gate (LDML-TA-NAND2) and 2-input LECTOR incorporated DML Type B NOR gate (LDML-TB-NOR2) and NAND gate (LDML-TB-NAND2) are placed in Sect. 4 which further show simulated results of four-input NAND and NOR DML and LECTOR incorporated (LDML-TA-NAND4, LDML-TA-NOR4, LDML-TB-NAND4, LDML-TB-NOR4) designs. Finally, the paper is concluded in Sect. 5.

2 DML Logic Operation

DML logic family has the advantage of allowing on the fly change between two modes-static mode and dynamic mode [19]. DML has two types of configurations for its design-Unfooted and Footed DML gates. In both these variants, DML allows two types of topologies- Type A and Type B topology as shown in Fig. 1. In all topologies, a pre-charge (M_{pre})/pre-discharge (M_{eval}) transistor is added at the output to conventional CMOS design. In footed variants of DML, an additional NMOS transistor (M_{nfoot}) is put between Pull Down Network (PDN) and ground for type-A topology and for type-B topology, a PMOS transistor (M_{pfoot}) is placed between V_{DD} and Pull Up Network (PUN) [19]. In type A footed DML design, the MODE signal is provided with a constant HIGH to cut-off the pre-charge transistor (M_{pre}) in static mode and the output is measured using the inputs applied to the standard CMOS gate input terminals i.e. it behaves similar to a standard CMOS gate. In dynamic mode, the MODE signal is provided with a clock signal providing both pre-charge and evaluation phase. Type B footed topology achieves static mode of operation by using a constant LOW as MODE signal which cut-off the pre-discharge transistor (M_{eval}). For dynamic mode, a clock signal is supplied as MODE signal to have both pre-charge and evaluation phase. In pre-charge phase, for type-A footed DML, the output node is charged to V_{DD} through M_{pre} transistor and M_{nfoot} transistor is off. In type B footed DML, the output node is discharged through M_{eval} transistor and transistor M_{pfoot} is off [21]. The output is evaluated depending on the types of inputs applied to the DML gate in evaluation phase. The flexibility of the design to switch between static and dynamic mode makes these structures unique as compared to other logic families [19].

3 Leakage in DML

In this section, leakage phenomenon in DML gates is discussed by considering the example of NAND gate in footed type A and type B topology as depicted in Figs. 2 and 3. The PDN consists of stacked NMOS transistors (TN_1-TN_k) in series and PUN has PMOS transistors (TP_1-TP_k) with a pre-charge PMOS transistor (T_1) in parallel in Type A. In static mode, transistor (T_1) should ideally be off for all input combinations but some current persists which make up the leakage current. Consider the case when all inputs are LOW, PDN is off, due to leakage a path is formed between output to ground. Considering all inputs HIGH case, substantial amount of current flows through off

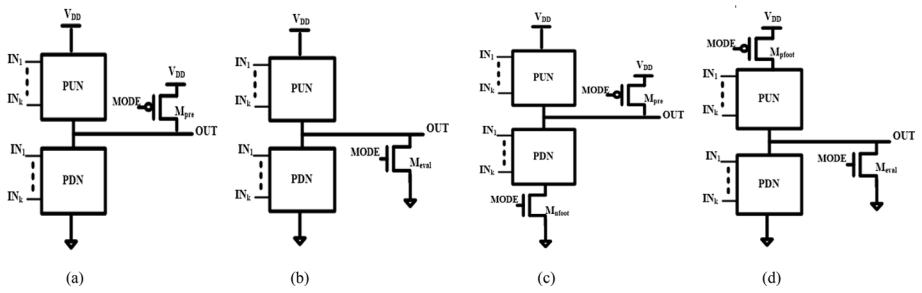


Fig. 1 Dual mode logic topologies: **a** Type A, **b** Type B, **c** footed Type A and **d** footed Type B [21]

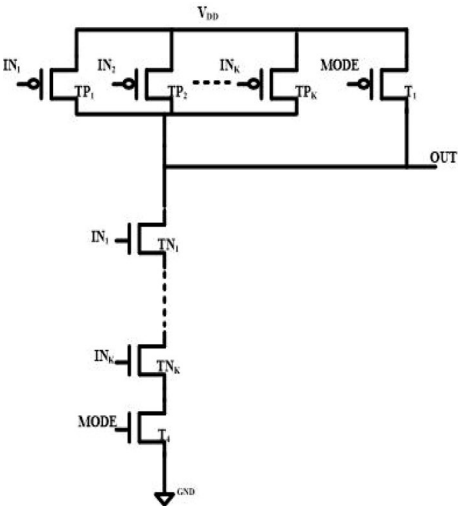


Fig. 2 DML footed Type A NAND

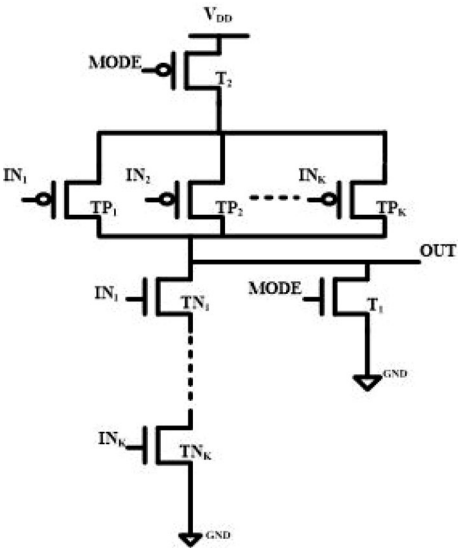


Fig. 3 DML footed Type B NAND

PMOS transistors. In its dynamic mode, during pre-charge phase, transistor T_1 is on and charges the output node to VDD. The footed transistor, T_4 is also off. Ideally the current through PDN should be zero but subthreshold leakage current through off transistors manage to flow. During evaluation, the output charging or discharging depends on the input vectors applied. The leakage current will be governed by the input vectors applied and the transistor, T_1.

Similarly, in type B footed NAND gate, in static mode, the pre-discharge transistor (T_1) along with off transistors due to inputs will incur leakage. In dynamic mode, the output remains at logic LOW during pre-discharge phase with MODE input HIGH irrespective of the input conditions since transistor T_1 is on and provides the path for

current to ground. Ideally the current through the PUN should be zero but the sub-threshold leakage current through the off transistors along the path manage to flow to the ground terminal. During evaluation phase with MODE input LOW, the output of DML design depends on the inputs applied. When all inputs are LOW, the PDN is off but due to leakage in off transistors (TN_1 to TN_k and transistor T_1) leakage current flows which contributes to leakage power in the evaluation mode of dynamic mode DML design. Hence, there exists leakage current through off transistors in the evaluation network depending on the input combinations applied.

4 LECTOR with DML Logic

4.1 LECTOR Operation

LECTOR was proposed as a technique to achieve power reduction in various circuits by using the concept of stacking of transistors in the path from supply voltage to ground [11]. It is based on the concept that "more leakage reduction is achieved if more number of OFF transistors are present in the path from VDD to ground" [11]. So, two leakage control transistors referred to as LCTs are introduced in each CMOS gate such that one of the transistors is always near its cut-off region of operation. Due to which the effective resistance between VDD and ground increases resulting in reduced leakage. Incorporation of high threshold transistors as LCTs further improves the design.

4.2 LECTOR in DML Logic

The proposed circuit technique incorporates the concept of leakage reduction by LECTOR technique in DML footed design as illustrated in Fig. 4. Consider the working of DML footed type A NAND gate using the proposed approach, transistor LCT1 (PMOS) and LCT2 (NMOS) are introduced between the PUN and PDN and the gates of these transistors are controlled by each other's source. The drain nodes of LCT1 and LCT2 are

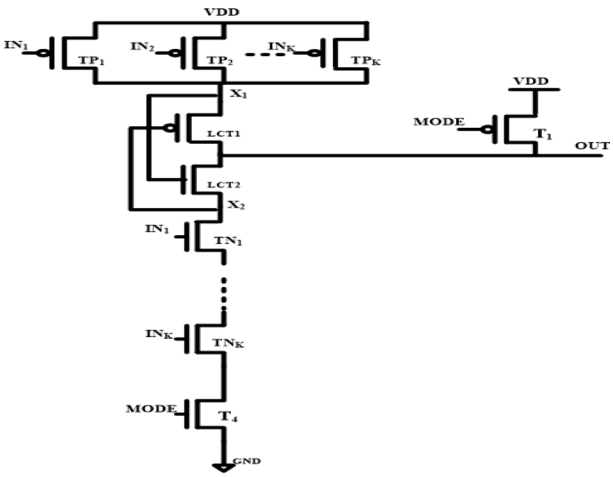


Fig. 4 DML footed Type A NAND gate using LECTOR approach (LDML-TA-NAND)

connected to form the output node to which the pre-charge transistor, T_1 (PMOS) is connected to have the DML design. Switching of transistors, LCT1 and LCT2 will depend on the voltage potential at nodes X2 and X1, respectively. In static mode of DML, the MODE signal is set to constant HIGH making transistor T_1 to cut-off. This MODE signal would also turn on transistor T_4. The output in this case depends on the inputs applied at input terminals of NAND gate. If all the inputs are HIGH, the PDN is on which sets output to logic LOW but at the same time PUN is off and some amount of leakage exists in the off transistors (TP_1-TP_k) and the off transistor T_1. This results in flow of leakage current from supply to ground. Due to the presence of LCTs in the path the effective resistance increases along the path which results in less leakage current. One of the LCTs would be in cut-off depending on the inputs applied. When all inputs are LOW, the PUN is on and PDN is off driving output level to HIGH, leakage current flows through off transistors and additional transistor but due to presence of LCTs, leakage current decreases. During pre-charge phase in dynamic mode, the MODE signal is set to LOW, the output node is charged HIGH through PUN transistors; it is independent of inputs applied but the leakage current is governed by the inputs applied. The transistor LCT1 offers more resistance along the path which leads to a reduction in leakage current. When MODE signal is set to HIGH, transistor T_1 is off and footed transistor, T_4 is on, this phase is referred to as evaluation phase, and depending on the inputs, the output node is charged or discharged. If all inputs are set to HIGH level, the output node would be discharged through PDN transistors. Leakage current flows through off PUN transistors but with incorporation of LCTs, the path from VDD to ground becomes more resistive towards the flow of current. Similar analysis goes for the case when all inputs are LOW in evaluation phase.

5 Simulation Results

All the circuits are designed with footed DML and proposed LECTOR incorporated footed DML technique. The circuits are simulated in a 45 nm and 90 nm PTM models for CMOS technology ($V_{DD} = 1.2$ V, $C = 5$ fF) with 2-input footed type A and type B DML NAND and NOR gates by the Symica DE tool. A comparison is performed for leakage power in all 2-input circuits by both techniques during static and dynamic modes at three different temperatures i.e. − 25 °C, 27 °C and 100 °C. For 2-input designs, average leakage power over all input combinations is considered. The proposed topology is also extended to four input NAND (NAND4) and NOR (NOR4) designs in static and dynamic mode. Effect of standard and high threshold voltage LCTs is examined in all the proposed designs.

5.1 Leakage Power in Static Mode of DML

In static mode, the average leakage power of the footed DML circuits at 90 nm and 45 nm for − 25 °C, 27 °C and 100 °C is shown in Figs. 5 and 6 for footed type A and type B topology respectively. The results demonstrate that the average leakage power in LECTOR incorporated footed DML circuits is reduced as compared with that for the footed DML circuits.

The percentage power saving achieved by the proposed circuit technique is shown in Table 1. Percentage saving obtained in leakage varies from 53 to 70% and 28% to 50% for 45 nm and 90 nm respectively at − 25 °C; at 27 °C, leakage power decreases by 48

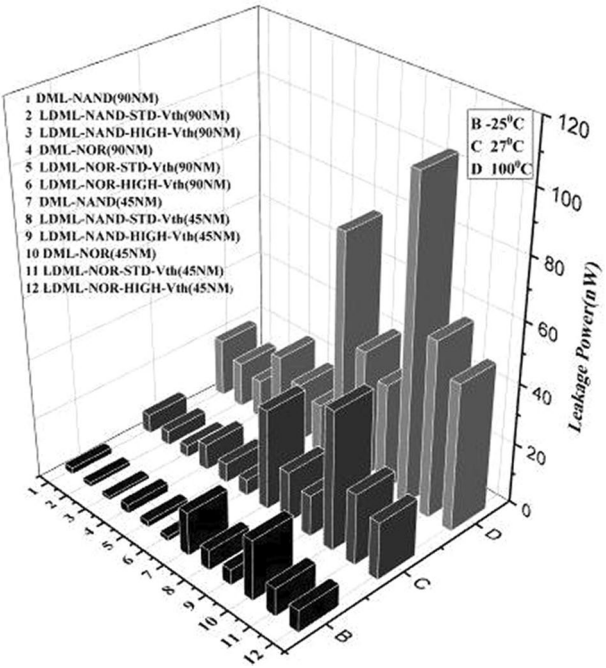


Fig. 5 Leakage power comparison of static DML-TA and static LDML-TA technique at − 25 °C, 27 °C and 100 °C at 90 nm and 45 nm technology node

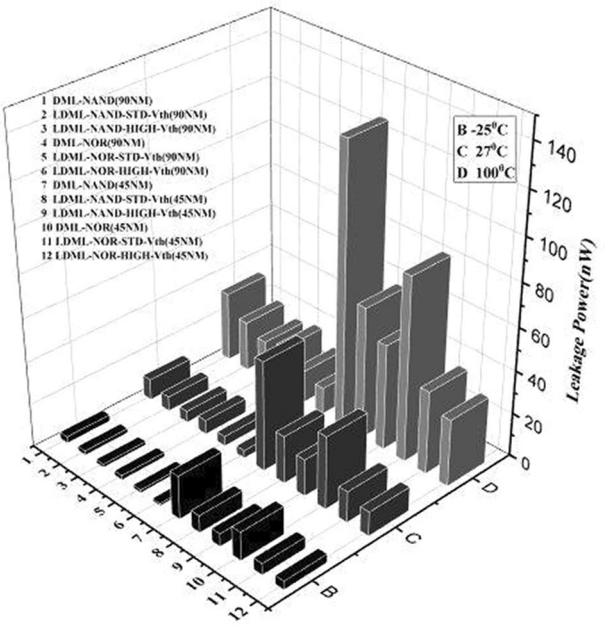


Fig. 6 Leakage power comparison of static DML-TB and static LDML-TB at − 25 °C, 27 °C and 100 °C at 90 nm and 45 nm technology node

to 66% for 45 nm and by 24 to 44% for 90 nm; and at 100 °C, power saving is 44 to 63% for 45 nm and 21to 41% for 90 nm.

Table 1 Percentage leakage power saving for two input NAND and NOR type A and type B topology in static mode for different temperature at 45 nm and 90 nm

	Power saving at − 25 °C		Power saving at 27 °C		Power saving at 100 °C	
	Standard V_{th}	High V_{th}	Standard V_{th}	High V_{th}	Standard V_{th}	High V_{th}
45 nm process technology, supply voltage = 1.2 V, static mode of DML						
LDML-TA-NAND2	54.09	64.18	48.99	58.93	44.53	53.59
LLDML-TB-NAND2	60.38	70.37	56.75	66.43	53.27	62.35
LLDML-TA-NOR2	53.41	63.15	48.79	58.49	44.59	53.54
LLDML-TB-NOR2	59.79	69.55	57.32	66.6	54.73	63.36
90 nm process technology, supply voltage = 1.2 V, Static mode of DML						
LLDML-TA-NAND2	28.56	44.68	24.37	38.36	21.66	33.49
LLDML-TB-NAND2	34.06	50.58	30.19	44.84	27.98	40.78
LLDML-TA-NOR2	28.52	44.28	24.31	38.38	21.98	33.72
LLDML-TB-NOR2	33.34	49.46	30.48	44.82	29.29	41.79

The observations from Table 1 are further summarized as:

- Higher power saving is achieved at 45 nm technology node in comparison to the 90 nm counterpart.
- Approximately 10% higher power saving is achieved with High V_{th} LCTs with respect to Standard V_{th} LCTs.
- The power saving shows downward trend with increase in temperature.

The power saving is also examined for four input NAND and NOR type A and type B topology in static mode at 45 nm and 90 nm at 27 °C. The power saving is observed for two conditions—all inputs as LOW, and all inputs as HIGH. The findings are listed in Table 2 which are similar to those obtained from Table 1.

Table 2 Percentage leakage power saving for four input NAND and NOR type A and type B topology in static mode at 45 nm and 90 nm at 27 °C

	Standard V_{th} (90 nm)	High V_{th} (90 nm)	Standard V_{th} (45 nm)	High V_{th} (45 nm)
Power saving at supply voltage = 1.2 V, static mode of DML				
LDML-TA-NAND4 (LOW inputs)	18	38.23	36.01	49.78
LDML-TA-NAND4 (HIGH inputs)	41.19	55.28	60.42	68.12
LDML-TA-NOR4 (LOW inputs)	36.86	49.93	62.35	70.61
LDML-TA-NOR4 (HIGH inputs)	24.11	41.37	53.5	65.04
LDML-TB-NAND4 (LOW inputs)	29.45	44.18	55	64.81
LDML-TB-NAND4 (HIGH inputs)	39.72	54.11	69.65	77.58
LDML-TB-NOR4 (LOW inputs)	37.64	50.39	62.83	70.9
LDML-TB-NOR4 (HIGH inputs)	0.26	1.02	0.23	0.69

5.2 Leakage Power in Pre-charge Phase of Dynamic Mode of DML

This subsection discusses the leakage power by incorporating LECTOR technique in pre-charge phase of dynamic mode. The average leakage power of the DML circuits at 90 nm and 45 nm for three different temperatures i.e. − 25 °C, 27 °C and 100 °C is shown in Figs. 7 and 8 for footed type A and type B topology respectively. Table 3 enlists the percentage power saving for 2-input designs in dynamic mode. It can be seen that in pre-charge phase, the percentage power saving is 9–35% for 45 nm and 4–35% for 90 nm at − 25 °C; 8–32% for 45 nm and 4–26% for 90 nm at 27 °C and 6–30% for 45 nm and 3–23% for 90 nm at 100 °C. The overall trend is similar to the one observed in static mode counterparts. The power saving is also examined for four input NAND and NOR type A and type B topology in static mode at 45 nm and 90 nm at 27 °C. The power saving is observed for two conditions—all inputs as LOW, and all inputs as HIGH. The findings are listed in Table 3 which are similar to those obtained from Table 1.

5.3 Leakage Power in Evaluation Phase of Dynamic Mode of DML

Leakage Power reduction achieved by the proposed technique in evaluation phase of dynamic mode is discussed in this subsection. For evaluation phase, the average leakage power comparison of 2-input designs is depicted in Figs. 9 and 10 for footed type A and footed type B topology respectively at two different technology nodes- 45 nm and 90 nm with temperature − 25 °C, 27 °C and 100 °C.

As observed from Table 3, at − 25 °C, power saving is 53–70% for 45 nm and 28–50% for 90 nm; at 27 °C, 48–66% for 45 nm and 24–44% for 90 nm; and at 100 °C, 44–63% for 45 nm and 21–41% for 90 nm. The power saving is also examined for four input NAND and NOR type A and type B topology in static mode at 45 nm and 90 nm at 27 °C. The power saving is observed for two conditions—all inputs as LOW, and all inputs as HIGH. The findings are listed in Table 4 which are similar to those obtained from Table 1.

Few of the observations which have been made from Tables 1, 2, 3 and 4 are:

1. As the design progresses towards lower technology node i.e. from 90 to 45 nm leakage power of the design increases.
2. The percentage power saving of the design increases with technology scaling i.e. the efficacy of the proposed technique to combat leakage power increases
3. An increase in the operating temperature of the design leads to an increase in the leakage power.
4. When LECTOR incorporated design is implemented with high threshold voltage for LCTs then there is an increase of power saving of the design.

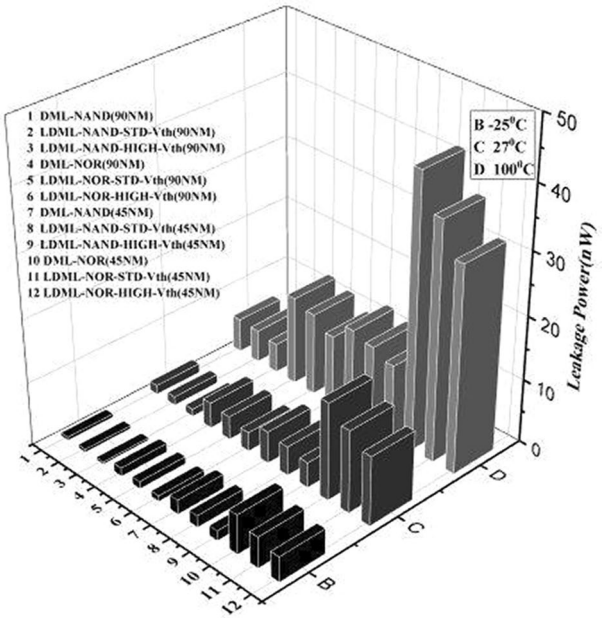


Fig. 7 Leakage power comparison of pre-charge phase of dynamic DML-TA and LDML-TA at − 25 °C, 27 °C and 100 °C at 90 nm and 45 nm technology node

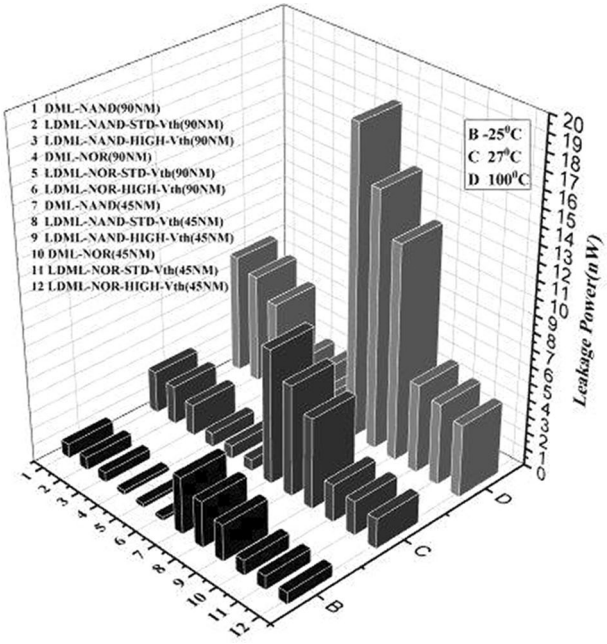


Fig. 8 Leakage power comparison of pre-charge phase of dynamic DML-TB and LDML-TB at − 25 °C, 27 °C and 100 °C at 90 nm and 45 nm technology node

Table 3 Percentage leakage power saving for 2-input type A and type B DML NAND and NOR in dynamic mode for different temperature at 45 nm and 90 nm

	Power saving at − 25 °C		Power saving at 27 °C		Power saving at 100 °C	
	Standard V_{th}	High V_{th}	Standard V_{th}	High V_{th}	Standard V_{th}	High V_{th}
45 nm process technology, supply voltage = 1.2 V, dynamic mode of DML						
Pre-charge phase						
LDML-TA-NAND2	16.69	29.66	9.03	20.41	6.41	15.37
LDML-TB-NAND2	18.71	35.65	17.55	32.53	16.81	30.07
LDML-TA-NOR2	21.24	35.68	15.26	28.51	12.82	24.27
LDML-TB-NOR2	9.08	20.96	8.28	19.36	8.14	17.82
Evaluation phase						
LDML-TA-NAND2	54.09	64.18	49.01	58.93	44.54	53.6
LDML-TB-NAND2	60.38	70.37	56.49	66.43	53.27	62.36
LDML-TA-NOR2	53.4	63.15	48.79	58.48	44.59	53.54
LDML-TB-NOR2	59.84	69.61	57.32	66.61	54.72	63.36
90 nm process technology, supply voltage = 1.2 V, dynamic mode of DML						
Pre-charge phase						
LDML-TA-NAND2	10.39	30.86	4.51	19.24	3.09	12.98
LDML-TB-NAND2	9.03	29.7	8.55	26.26	8.27	23.27
LDML-TA-NOR2	14.75	35.16	8.79	25.24	6.82	19.89
LDML-TB-NOR2	4.12	20.71	4.38	17.67	3.85	14.79
Evaluation phase						
LDML-TA-NAND2	28.58	44.71	24.36	38.36	21.68	33.49
LDML-TB-NAND2	34.08	50.53	30.19	44.84	27.97	40.77
LDML-TA-NOR2	28.51	44.29	24.36	38.39	21.95	33.67
LDML-TB-NOR2	33.25	49.39	30.49	44.83	29.29	41.79

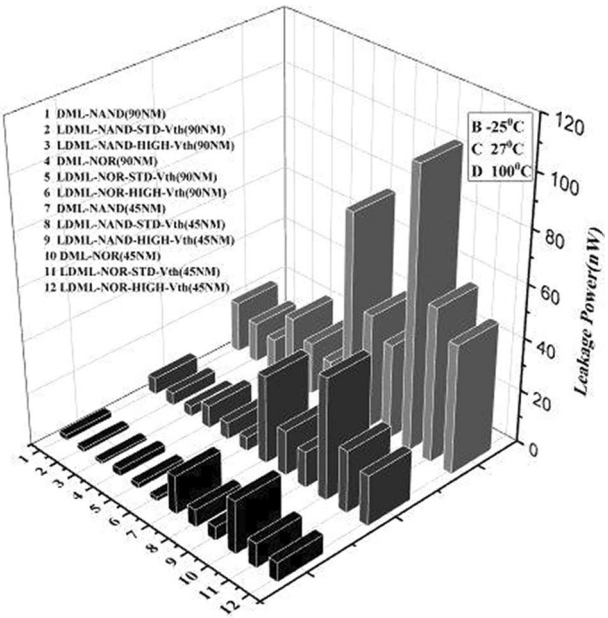


Fig. 9 Leakage power comparison of evaluation phase of dynamic DML-TA and LDML-TA at − 25 °C, 27 °C and 100 °C at 90 nm and 45 nm technology node

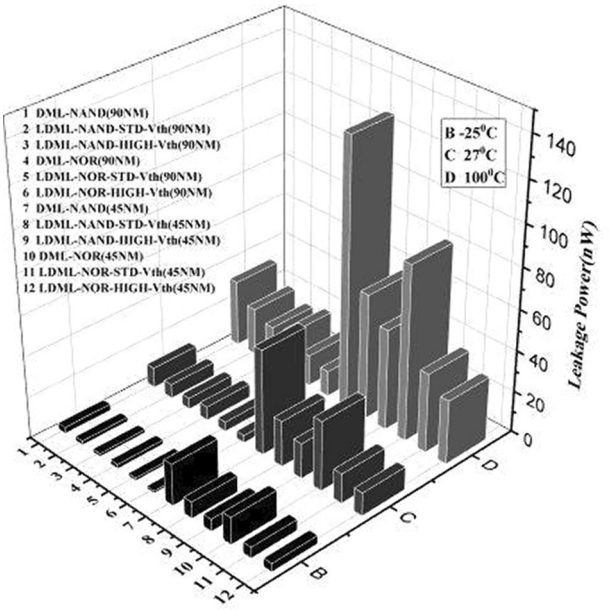


Fig. 10 Leakage power comparison of evaluation phase of dynamic DML-TB and LDML-TB at − 25 °C, 27 °C and 100 °C at 90 nm and 45 nm technology node

Table 4 Percentage leakage power saving for 4-input type A and type B NAND and NOR in dynamic mode at 45 nm and 90 nm at 27 °C

	Standard V_{th} (90 nm)	High V_{th} (90 nm)	Standard V_{th} (45 nm)	High V_{th} (45 nm)
Power saving at supply voltage = 1.2 V, dynamic mode of DML				
Pre-charge phase				
LDML-TA-NAND4 (low inputs)	18.29	39.82	23.36	39.71
LDML-TA-NAND4 (high inputs)	4.27	18.52	3.31	10.62
LDML-TA-NOR4 (low inputs)	21.8	43.08	37.91	51.88
LDML-TA-NOR4 (high inputs)	12.02	28.79	19.13	34.13
LDML-TB-NAND4 (low inputs)	14.29	33.12	24.25	41.35
LDML-TB-NAND4 (high inputs)	1.47	4.41	3.73	8.21
LDML-TB-NOR4 (low inputs)	2.59	14.22	4.51	12.56
LDML-TB-NOR4 (high inputs)	0.37	0.73	0.58	1.16
Evaluation phase				
LDML-TA-NAND4 (low inputs)	18	38.23	36.01	49.78
LDML-TA-NAND4 (high inputs)	41.19	55.28	60.42	68.12
LDML-TA-NOR4 (low inputs)	36.91	49.93	62.35	70.61
LDML-TA-NOR4 (high inputs)	24.11	41.37	53.54	65.04
LDML-TB-NAND4 (low inputs)	29.45	44.18	55	64.81
LDML-TB-NAND4 (high inputs)	39.72	54.18	69.65	77.58
LDML-TB-NOR4 (low inputs)	37.64	50.39	62.83	70.9
LDML-TB-NOR4 (high inputs)	0.15	0.87	0.23	0.69

6 Conclusion

In this paper, LECTOR based leakage reduction technique is incorporated in footed type A and type B DML circuits in static and dynamic mode. The performance of the proposal is examined at 45 nm and 90 nm technology nodes. At 27°, the simulation results at 90 nm show 24.37–44.84% leakage power saving in LDML-TA and LDML-TB in comparison to the standard footed DML logic counterpart. The saving increases to 48.79–66.6% at 45 nm under same conditions. For the same logic gates at 27 °C, power saving of 4.38–26.26% at 90 nm and 8.28–32.53% at 45 nm in pre-charge phase of dynamic mode. During evaluation phase, power saving of 24.36–44.84% and 48.79–66.61% at 90 nm and 45 nm respectively, when compared to standard footed DML logic circuits. The proposed technique can be extended to low power applications where DML based designs can be used as a power efficient alternative to existing design.

Funding The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Data Availability The authors confirm that the data supporting the findings of this study are available within the article, its supplementary materials or mentioned references in the article itself.

Declarations

Conflict of interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Zhang, L. (2014). Silicon process and manufacturing technology evolution: An overview of advancements in chip making. *IEEE Consumer Electronics Magazine, 3*(3), 44–48. https://doi.org/10.1109/MCE.2014.2317896
2. Xiu, L. (2019). Time Moore: Exploiting Moore's law from the perspective of time. *IEEE Solid State Circuits Magazine, 11*(1), 39–55. https://doi.org/10.1109/MSSC.2018.2882285
3. Pedram, M., & Rabaey, J. (2002). *Power aware design methodologies*. Kluwer Academic Publishers. https://doi.org/10.1007/b101914
4. Rabaey, J. (2009). *Low power design essentials*. Berlin: Springer. https://doi.org/10.1007/978-0-387-71713-5
5. Wei, L., Chen, Z., Roy, K., Johnson, M. C., Ye, Y., & De, V. K. (1999). Design and optimization of dual-threshold circuits for low-voltage low-power applications. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems, 7*(1), 16–24. https://doi.org/10.1109/92.748196
6. Wei, L., Chen, Z., Roy, K., Ye, Y., & De, V. (1999). Mixed-Vth (MVT) CMOS circuit design methodology for low power applications. In *Proceedings of the 36th annual ACM/IEEE design automation conference, New Orleans, LA, USA* (pp. 430–435). https://doi.org/10.1109/DAC.1999.781355
7. Park, J. C., & Mooney, V. J., III. (2006). Sleepy stack leakage reduction. *IEEE Transactions on Very Large Scale Integration (VLSI) systems, 14*(11), 1250–1263. https://doi.org/10.1109/TVLSI.2006.886398
8. Islam, M. S., Nasrin, M. S., Mansur, N., & Tasneem, N. (2010). Dual stack method: A novel approach to low leakage and speed power product VLSI design. In *International conference on electrical & computer engineering (ICECE 2010)* (pp. 89–92). IEEE, Dhaka, Bangladesh. https://doi.org/10.1109/ICELCE.2010.5700560
9. Lakshmikanthan, P., & Nunez, A. (2007). VCLEARIT: A VLSI CMOS circuit leakage reduction technique for nanoscale technologies. *ACM Sigarch Computer Architecture News, 35*(5), 10–16. https://doi.org/10.1145/1360464.1360471

10. Anjana, R., & Somkuwar, A. (2014). FELERION: A new approach for leakage power reduction. *Journal of Semiconductors, 35*(12), 1–5. https://doi.org/10.1088/1674-4926/35/12/125001
11. 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
12. Gupta, T. K., & Khare, K. (2013). LECTOR with footed-diode inverter: A technique for leakage reduction in domino circuits. *Circuits, Systems and Signal Processing, 32*(6), 2707–2722. https://doi.org/10.1007/s00034-013-9615-2
13. Katrue, S., & Kudithipudi, D. (2008). GALEOR: Leakage reduction for CMOS circuits. In *15th IEEE international conference on electronics, circuits and systems* (pp. 574–577), IEEE, St. Julien's. https://doi.org/10.1109/ICECS.2008.4674918
14. Sharma, V. K., Pattanaik, M., & Raj, B. (2014). ONOFIC approach: Low power high speed nanoscale VLSI circuits design. *International Journal of Electronics, 101*(1), 61–73. https://doi.org/10.1080/00207217.2013.769186
15. Sharma, V. K., Pattanaik, M., & Raj, B. (2015). INDEP approach for leakage reduction in nanoscale CMOS circuits. *International Journal of Electronics, 102*(2), 200–215. https://doi.org/10.1080/00207217.2014.896042
16. Bajpai, P., Pandey, N., Gupta, K., Panda, J., Bajpai, P., et al. (2017). On improving the performance of dynamic DCVSL circuits. *Journal of Electrical and Computer Engineering, 2017*, 1–11. https://doi.org/10.1155/2017/8207104
17. Bajpai, P., Pandey, N., Gupta, K., & Panda, J. (2019). LECTOR incorporated Differential Cascode Voltage Swing Logic (L-DCVSL). *Analog Integrated Circuits and Signal Processing, 100*(1), 221–234. https://doi.org/10.1007/s10470-019-01466-2
18. Lorenzo, R., & Chaudhury, S. (2017). Review of circuit level leakage minimization techniques in CMOS VLSI circuits. *IETE Technical Review, 34*(2), 165–187. https://doi.org/10.1080/02564602.2016.1162116
19. Kaizerman, A., Fisher, S., & Fish, A. (2013). Subthreshold dual mode logic. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems, 21*(5), 979–983. https://doi.org/10.1109/TVLSI.2012.2198678
20. Levi, I., & Fish, A. (2013). Dual mode logic—Design for energy efficiency and high performance. *IEEE Access, 1*, 258–265. https://doi.org/10.1109/ACCESS.2013.2262015
21. Levi, I., Kaizerman, A., & Fish, A. (2013). Low voltage dual mode logic: Model analysis and parameter extraction. *Microelectronics journal, 44*(6), 553–560. https://doi.org/10.1016/j.mejo.2013.03.005
22. Yuzhaninov, V., Levi, I., & Fish, A. (2016). Design flow and characterization methodology for dual mode logic. *IEEE Access, 3*, 3089–3101. https://doi.org/10.1109/ACCESS.2016.2514398
23. Bikki, P. (2019). Analysis of high-performance near-threshold dual mode logic design. *International Journal of Electronics and Telecommunications, 65*(4), 723–729.
24. Moyal, L., Levi, I., Teman, A., & Fish, A. (2016). Synthesis of dual mode logic. *Integration, the VLSI Journal, 55*, 246–253. https://doi.org/10.1016/j.vlsi.2016.07.004
25. Levi, I., Bass, O., Kaizerman, A., Belenky, A., & Fish, A. (2012). High speed dual mode logic carry look ahead adder. In *2012 ieee international symposium on circuits and systems (ISCAS)* (pp. 3037–3040). IEEE, Seoul, South Korea. https://doi.org/10.1109/ISCAS.2012.6271959
26. Rose, R. D., Romero, P., & Lanuzza, M. (2019). Double-precision dual mode logic carry-save multiplier. *Integration, the VLSI Jouirnal, 64*, 71–77. https://doi.org/10.1016/j.vlsi.2018.08.003
27. Levi, I., Belenky, A., & Fish, A. (2013). Logical effort for CMOS-based dual mode logic gates. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems, 22*(5), 1042–1053. https://doi.org/10.1109/TVLSI.2013.2257902
28. Taco, R., Levi, I., Lanuzza, M., & Fish, A. (2017). Evaluation of dual mode logic in 28 nm FD-SOI technology. In *2017 IEEE international symposium on circuits and systems (ISCAS)* (pp. 1–4). IEEE, MD, USA. https://doi.org/10.1109/ISCAS.2017.8050998
29. Taco, R., Levi, I., Lanuzza, M., & Fish, A. (2017). Energy-delay tradeoffs of low-voltage dual mode logic in 28nm FD-SOI. In *2017 IEEE SOI-3D-subthreshold microelectronics technology unified conference (S3S)* (pp. 1–3). IEEE, CA, USA. https://doi.org/10.1109/S3S.2017.8309250
30. Taco, R., Levi, I., Lanuzza, M., & Fish, A. (2019). Live demo: An 88fJ/40 MHz [0.4 V]–0.61 pj/1GHz [0.9 V] dual mode logic 8× 8-bit multiplier accumulator with a self-adjustment mechanism in 28 nm FD-SOI. In *2019 IEEE international symposium on circuits and systems (ISCAS)* (p. 1). IEEE, Sapporo, Japan. https://doi.org/10.1109/ISCAS.2019.8702170
31. Jose, C. E., & Kousalya, B. (2015). Power reduction in CMOS sub-threshold dual mode logic circuits by power gating. *IOSR Journal of VLSI and Signal Processing, 5*(2), 60–67.

32. Lakshmisree, P. V., & Raghu, M. C. (2014). Design of subthreshold DML logic gates with power gating techniques. *International Journal of Research in Engineering and Technology, 3*(4), 174–180. https://doi.org/10.15623/ijret.2014.0304032
33. Singh, S. N., & Madhu, R. (2015). Power analysis for CMOS based dual mode logic gates using power gating techniques. *International Journal of Science, Engineering and Technology Research, 4*(12), 4067–4072.
34. Patil, S. S., Pathak, S. S., Kathar, R. R., & Patil, D. S. (2017). Low power based dual mode logic gates using power gating technique. *International Research Journal of Engineering and Technology, 4*(5), 1462–1467.
35. Bikki, P., & Karuppanan, P. (2016). Low power and high performance multi-Vth dual mode logic design. In *11th international conference on industrial and information systems, Roorkee, India* (pp. 463–468). https://doi.org/10.1109/ICIINFS.2016.8262985

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.

Neetika Yadav received her B.Tech. in Electronics and Communication Engineering from Guru Gobind Singh Indraprastha University (GGSIPU), Delhi, India and M.Tech. in VLSI Design from Guru Gobind Singh Indraprastha University (GGSIPU), Delhi, India. She is currently pursuing Ph.D. from Delhi Technological University (DTU), Delhi, India and working as assistant professor in department of ECE, Amity School of Engineering and Technology, affiliated to GGSIPU, Amity University, Noida. Her area of research is low power VLSI.

Neeta Pandey received her M.E. in Microelectronics from Birla Institute of Technology and Sciences, Pilani and Ph.D. from Guru Gobind Singh Indraprastha University Delhi. She has served in Central Electronics Engineering Research Institute, Pilani, Indian Institute of Technology, Delhi, Priyadarshini College of Computer Science, Noida and Bharati Vidyapeeth's College of Engineering, Delhi in various capacities. At present, she is Professor in ECE Department, Delhi Technological University. A life member of ISTE, and senior member of IEEE, USA. She has published papers in international, national journals of repute and conferences. Her research interests are in analog and digital VLSI Design.

Deva Nand received his B.Tech., M.Tech. degrees in Electronic and Communication Engineering from Kurukshetra University, Kurukshetra, India. He received his Ph.D. from Delhi Technological University, Delhi, India. At present he is Assistant Professor in Department of ECE, DTU, Delhi, India. A life member of ISTE, member of IAENG and IEEE. He has published papers in International, National Journals of repute and conferences. His research interests include analog and digital VLSI design.

Physics of Plasmas

Localization and turbulence of beam-driven whistler wave with magnetosonic wave in magnetopause

Cite as: Phys. Plasmas **30**, 022904 (2023); https://doi.org/10.1063/5.0134920
Submitted: 15 November 2022 • Accepted: 20 January 2023 • Published Online: 17 February 2023

Jyoti, Suresh C. Sharma and R. P. Sharma

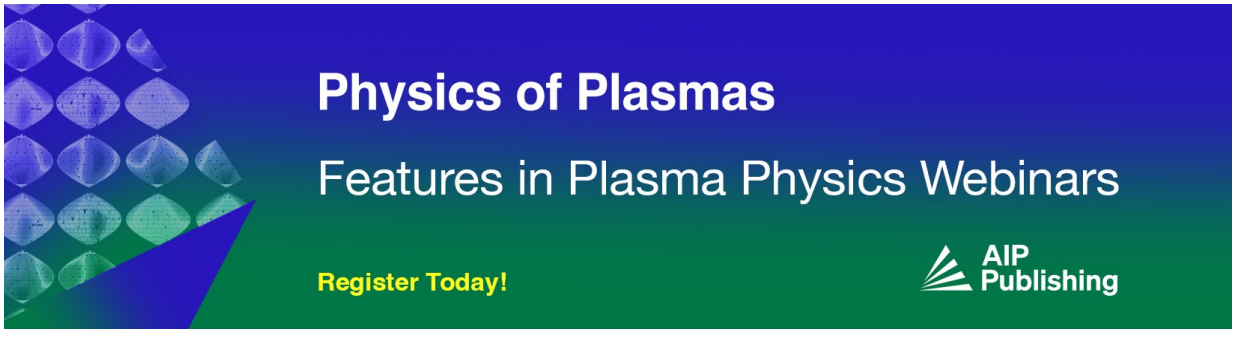
View Online Export Citation CrossMark

ARTICLES YOU MAY BE INTERESTED IN

The electron diffusion region dominated by electromagnetic turbulence in the reconnection current layer
Physics of Plasmas **30**, 022106 (2023); https://doi.org/10.1063/5.0129591

Theory of ion holes in plasmas with flat-topped electron distributions
Physics of Plasmas **30**, 022903 (2023); https://doi.org/10.1063/5.0086613

Energy transfer of the solar wind turbulence based on Parker solar probe and other spacecraft observations
Physics of Plasmas **30**, 020501 (2023); https://doi.org/10.1063/5.0121140



AIP Publishing

© 2023 Author(s).

Physics of Plasmas ARTICLE scitation.org/journal/php

Localization and turbulence of beam-driven whistler wave with magnetosonic wave in magnetopause

Cite as: Phys. Plasmas **30**, 022904 (2023); doi: 10.1063/5.0134920
Submitted: 15 November 2022 · Accepted: 20 January 2023 ·
Published Online: 17 February 2023

View Online Export Citation CrossMark

Jyoti,[1,a] Suresh C. Sharma,[1,b] and R. P. Sharma[2,c]

AFFILIATIONS

[1]Plasma and Nano-Simulation Lab, Department of Applied Physics, Delhi Technological University, Shahbad Daulatpur, Bawana Road, New Delhi 110042, India
[2]Centre for Energy Studies, Indian Institute of Technology, Delhi, Hauz Khas, New Delhi 110016, India

[a)]Electronic mail: jyotirani1506@gmail.com
[b)]Author to whom correspondence should be addressed: Suresh321sharma@gmail.com
[c)]Electronic mail: rpsiitd@gmail.com

ABSTRACT

This study proposes a model to explain how energetic electron beams (produced by the magnetic reconnection process) cause whistler turbulence in the magnetic reconnection area of the magnetopause, as observed by the Magnetospheric Multiscale Mission. In this scenario, the energetic electron beam source has replaced the magnetic reconnection mechanism. We develop dynamic equations of the beam-driven whistler mode so that because of the large amplitude of the beam energy, it rises from the noise level. As a result, nonlinear effects follow due to ponderomotive force, which results in whistler wave localization; eventually, the turbulent state is achieved. A theoretical model is developed using the basic two-fluid equations and Maxwell's equations to study the dynamics of high-frequency whistler waves and low-frequency magnetosonic waves (MSWs). Then, using a pseudospectral approach and a finite difference method, a set of dimensionless equations for the whistler wave and MSWs was numerically solved. The outcomes of the numerical simulation show a localized structure and a turbulent power spectrum, which follow Kolmogorov scaling laws. It has also been clarified that the current investigations are pertinent to the most recent observations.

Published under an exclusive license by AIP Publishing. https://doi.org/10.1063/5.0134920

I. INTRODUCTION

In space plasma physics, heating and acceleration of plasma particles have been an active research era. Numerous suggestions have been put up to explain heating and acceleration of plasma particles in space plasmas, but the underlying processes are still not fully understood. It is considered that plasma turbulence is a leading contender for particle and energy transmission.[1,2] Turbulence is crucial in space plasmas, because it cascades energy from larger to smaller scales, ultimately leading to dissipation and particle heating.[3–7] In laboratory, space, and astrophysical plasmas, magnetic reconnection is thought to emerge in a variety of applications. Magnetic reconnection is a fundamental plasma process of an energy conversion mechanism, which involves changing the magnetic topology to convert magnetic energy into plasma kinetic and thermal energy. There is significant evidence that in magnetized plasmas, magnetic reconnection and turbulence are likely connected.[8–11]

The Earth's magnetosheath is the highly turbulent and heated region bounded by the bow shock and the magnetopause. Thus, it provides better conditions to investigate plasma turbulence, transport, and dissipation, which are the major unresolved issues in the modern space era. Several wave modes exist in the space plasmas,[12–17] and the nonlinear interaction of these wave modes contributes to the formation of turbulent cascade modes. Among these wave modes, it is tremendously assumed that in space and laboratory plasmas, a whistler wave[18] plays a vital role. Whistler waves are observed in the magnetotail, radiation belts, and dayside magnetopause, among other areas of the space plasma.[19–22] Along with whistler wave modes, the magnetosonic wave (MSW) modes are also anticipated to be a prominent candidate for heating and accelerating the particles in space plasmas.[23–26] Recently, the generation of low-frequency MSWs through wave-wave interactions has been reported in earth magnetosheath.[27] For more analysis, the 2D particle-in-cell (PIC) simulation studied the magnetosonic turbulence and observed the

Published under an exclusive license by AIP Publishing

cascading of energy, which might contribute to the localized steepening of magnetic fluctuation of the MSW mode.[28]

Recently, Zhao *et al.*[29] observed the beam-driven whistler wave in the magnetopause region. They observed the solitary structures of whistler exist at f_{ce} (where f_{ce} is the electron gyrofrequency). Krafft and Volokitin[30] have studied the beam-driven Langmuir turbulence involving density fluctuations. Reeves *et al.*[31] have experimented with beam-plasma interactions to study and characterize fundamental wave-particle interactions by generating whistler waves using a modulated electron beam. Narita *et al.*[32] observed whistler turbulence on the electron scale in the solar-wind region along with the magnetic power spectrum. The broadband spectrum of the whistler mode of magnetic field fluctuations has been observed in the vicinity of the magnetopause.[33] Karpman[34] studied the nonlinear theory of localized whistler wave structures by a non-linear Schrödinger system of equations (NLS), which leads to the formation of density enhancements that trap whistler waves. Das *et al.*[35] demonstrated the turbulent spectrum of the whistler wave via nonlinear coupling of whistler waves with magnetosonic perturbations, thereby leading to the formation of nonlinear coherent structures. This broadband spectrum in the magnetopause may result from a cascade of spectral power from high to low frequencies as the result of gradient-driven instabilities.

Biskamp *et al.*[36] studied the electron magnetohydrodynamic (EMHD) turbulence and the various wave modes and also observed that turbulent structures and the energy spectrum E_k follow the Kolmogorov-type scaling, $E_k \sim k^{-5/3}$ for $k\lambda_e > 1$ and $E_k \sim k^{-7/3}$ for $k\lambda_e < 1$ by high-resolution numerical simulations. After that, Das and Diamond[37] examined the EMHD turbulence and linear properties of whistler waves through numerical simulation.

It is evident from the literature that the energy cascade arises from turbulence and also a weak turbulence theory is mediated by the nonlinear interactions of whistler waves.[38] Langmuir turbulence is studied by the nonlinear Schr ö dinger equation (NLS) and Zakharov equations (ZSEs), numerically solved for an initial condition through modulational instability and finally saturates into the solitons structure.[39] Doolen *et al.*[40] used Zakharov system of equations (ZSE), and Sharma *et al.*[41] used Zakharov–Boussinesq's equations to study the Langmuir turbulence using cascading of energy and cavitation processes.

The non-linear whistler wave along with solitons, radiation, self-focusing, and modulational instability has been studied in past by Karpman.[34] Later on, a few scientists also worked on wave turbulence. Also, a lot of studies have been done on whistler waves as discussed above but to the best of our knowledge specific research on beam-driven whistler turbulence has not been revealed. Recently, Jyoti *et al.*[42] studied beam-driven whistler turbulence on account of nonlinear density perturbations by ion-acoustic waves (IAWs). They studied the nonlinear structures and turbulence by the cavitation process by the modified Zakharov system of equations. However, in the magnetopause reconnection region along with Whistler waves, magnetosonic waves (MSWs) are frequently observed, and these MSWs have magnetic perturbations along with density perturbations. Therefore, the work of Jyoti *et al.*[42] is not applicable in these realistic situations, because magnetic perturbations may provide significant contributions to nonlinear structure formation and turbulence generation in the presence of a beam.

Motivated from the previous studies, in the present work, we have provided a model to explain how energetic electron beams (generated by the magnetic reconnection process) lead to whistler turbulence in the magnetic reconnection region of the magnetopause, as seen by the Magnetospheric Multiscale Mission (MMS).[29] The intense electron beam source has been used in place of the magnetic reconnection mechanism in this instance.

Here, we have proposed a model in which the beam-driven whistler waves are studied. From the energy of the beam, the amplitude of beam-driven modes grows from the noise level to the higher values such that the ponderomotive force will create the nonlinear effects in the whistler waves, which leads to the localization and turbulence of whistler waves.

For this purpose, we have proposed that the number density of the pump whistler wave becomes disturbed in the presence of the MSWs, and the components of ponderomotive force of the whistler are taken into account in the MSWs dynamics. (Since MSWs are an electromagnetic wave, we have considered perturbation in density as well as in a magnetic field, i.e., $n = n_0 + \delta n$, $B = B_0 + \delta B$.) Then, the governing equations of the whistler wave and the MSWs are framed in the form of coupled nonlinear equations. For these nonlinear equations, numerical simulation is carried out using the finite difference with modified version of the predictor-corrector method for temporal integration and the pseudo-spectral method for spatial integration. The outcomes of the numerical simulation show the localization and turbulent state of the whistler wave. The objective of our current work is to examine the localized structure and turbulent spectra in the magnetopause region.

This paper is organized as follows: Sec. II presents the model equations, which present the derivation part of the dynamical equation of the whistler wave and the magnetosonic wave. Section III presents the numerical simulation. Section IV presents the semianalytical method. Section V concludes this paper with the results, conclusion, and future recommendations.

II. MODEL EQUATIONS

The proposed theoretical model describes whistler turbulence in a plasma near the magnetopause by taking background density fluctuations into account. In this work, whistler turbulence induced by electron beams was believed to be well-developed. We have taken into account low and high-frequency plasma responses, density fluctuations, and slow and fast wave dynamics, including other linear and non-linear processes like wave coupling and ponderomotive force.

The analysis consists of two parts. Initially, the dynamical equation of high-frequency whistler waves is solved. Later, the dynamical equations of MSWs and the relationship between low-frequency MSWs and high-frequency whistler waves are discussed. A 2D whistler wave system using an electron beam is characterized by a two-fluid model, considering the positive growth rate of the beam-driven whistler in our computation. Hence, phenomenologically, an instability growth rate $\gamma' > 0$ is into whistler dynamics.

A. 2D whistler dynamics

The whistler wave dynamical equation is obtained using a two-fluid model. The whistler wave is propagating in a magnetized plasma with the magnetic field along the z-axis and wave vector, $\vec{k} = k_{0x}\hat{x} + k_{0z}\hat{z}$. Here, $B\,(= B_0 + \delta B)$ is the magnetic field, B_0 is the background magnetic field, and δB is the perturbation in the magnetic field.

Published under an exclusive license by AIP Publishing

The governing dynamical equation of the whistler wave can be written as (see the Appendix for a detailed derivation of whistler's dispersion relation)

$$\begin{aligned}&\frac{-\partial^2 E_y}{\partial t^2}-\lambda_e^4\frac{\partial^6 E_y}{\partial x^4\partial t^2}-\lambda_e^4\frac{\partial^6 E_y}{\partial z^4\partial t^2}-2\lambda_e^4\frac{\partial^6 E_y}{\partial x^2\partial z^2\partial t^2}\\&\quad+2\lambda_e^2\frac{\partial^4 E_y}{\partial x^2\partial t^2}+2\lambda_e^2\frac{\partial^4 E_y}{\partial z^2\partial t^2}\\&\quad=\lambda_e^4\omega_{ce}^2\left(\frac{\partial^4 E_y}{\partial z^2\partial x^2}+\frac{\partial^4 E_y}{\partial z^4}\right)+\lambda_e^4\omega_{ce}^2\left(\frac{\partial^4 E_y}{\partial z^2\partial x^2}+\frac{\partial^4 E_y}{\partial z^4}\right)\frac{\delta B}{B_0},\end{aligned}\tag{1}$$

where $\lambda_e\left(=\sqrt{\frac{c^2 m_e}{4\pi n_0 e^2}}\right)$ is the electron skin depth and $\omega_{ce}\left(=\frac{eB_0}{m_e c}\right)$ is the electron gyrofrequency.

As beam-driven growth is also considered, $\frac{d}{dt}\to\left(\frac{d}{dt}+\gamma\right)$,

$$\begin{aligned}&-\left(\frac{\partial^2}{\partial t^2}+2\gamma\frac{\partial}{\partial t}\right)E_y-\lambda_e^4\frac{\partial^6 E_y}{\partial x^4\partial t^2}-\lambda_e^4\frac{\partial^6 E_y}{\partial z^4\partial t^2}-2\lambda_e^4\frac{\partial^6 E_y}{\partial x^2\partial z^2\partial t^2}\\&\quad+2\lambda_e^2\frac{\partial^4 E_y}{\partial x^2\partial t^2}+2\lambda_e^2\frac{\partial^4 E_y}{\partial z^2\partial t^2}\\&\quad=\lambda_e^4\omega_{ce}^2\left(\frac{\partial^4 E_y}{\partial z^2\partial x^2}+\frac{\partial^4 E_y}{\partial z^4}\right)+\lambda_e^4\omega_{ce}^2\left(\frac{\partial^4 E_y}{\partial z^2\partial x^2}+\frac{\partial^4 E_y}{\partial z^4}\right)\frac{\delta B}{B_0}.\end{aligned}\tag{2}$$

For this Eq. (2), the following envelope solution is assumed:

$$\tilde{E}_y=E_y(x,z,t)e^{i(k_{0x}\hat{x}+k_{0z}\hat{z}-\omega_0 t)}.\tag{3}$$

Since the nonlinear whistler wave dynamical equation is achieved by considering the effect of a strong whistler wave, by substituting Eq. (3) in Eq. (2), we get

$$\begin{aligned}&2i\omega_0(1+\lambda_e^4k_{0x}^4+\lambda_e^4k_{0z}^4+2\lambda_e^2k_{0x}^2+2\lambda_e^2k_{0z}^2)\frac{\partial E_y}{\partial t}\\&\quad+(-4\omega_0^2\lambda_e^4k_{0x}^2)\frac{\partial^2 E_y}{\partial x^2}+(4\lambda_i^2v_A^2k_{0z}^2-4\omega_0^2\lambda_e^4k_{0z}^2)\frac{\partial^2 E_y}{\partial z^2}\\&\quad+(4\lambda_i^2v_A^2k_{0x}k_{0z}-8\omega_0^2k_{0z}k_{0x}\lambda_e^4)\frac{\partial^2 E_y}{\partial x\partial z}\\&\quad+2i(-2\omega_0^2k_{0x}^3\lambda_e^4+2\omega_0^2k_{0x}k_{0z}^2\lambda_e^4+2\omega_0^2\lambda_e^2k_{0x}+k_{0x}k_{0z}^2\lambda_i^2v_A^2)\frac{\partial E_y}{\partial x}\\&\quad+2i(-2\omega_0^2k_{0z}^3\lambda_e^4+2\omega_0^2k_{0z}k_{0x}^2\lambda_e^4+2\omega_0^2\lambda_e^2k_{0z}\\&\quad+k_{0z}k_{0x}^2\lambda_i^2v_A^2+2k_{0z}^3\lambda_i^2v_A^2)\frac{\partial E_y}{\partial z}+(9\lambda_i^2v_A^2k^2k_{0z}^2)\frac{\delta n}{n_0}E_y\\&\quad+(-9k_{0z}^4\omega_{ce}^2\lambda_e^4-4k_{0x}^2k_{0z}^2\lambda_e^4\omega_{ce}^2)\frac{\delta B}{B_0}=0,\end{aligned}\tag{4}$$

where k_{0x} and k_{0z} are wave vector components related to the background whistler magnetic field.

Using the normalizing parameter, we get

$$t_n=1/\omega_0.$$
$$z_n=\frac{2i(-2\omega_0^2k_{0x}^3\lambda_e^4+2\omega_0^2k_{0x}k_{0z}^2\lambda_e^4+2\omega_0^2\lambda_e^2k_{0x}+k_{0x}k_{0z}^2\lambda_i^2v_A^2+2k_{0z}^3\lambda_i^2v_A^2)}{2\omega_0(1+\lambda_e^4k_{0x}^4+\lambda_e^4k_{0z}^4+2\lambda_e^2k_{0x}^2+2\lambda_e^2k_{0z}^2)}.$$

Thus, the normalized dimensionless form of Eq. (4) is given as

$$\begin{aligned}&i\frac{\partial E'}{\partial t'}+c_1\frac{\partial^2 E'}{\partial x'^2}+c_2\frac{\partial^2 E'}{\partial z'^2}+i\frac{\partial E'}{\partial x'}+i\frac{\partial E'}{\partial z'}+c_3\frac{\partial^2 E'}{\partial x'\partial z'}\\&\quad+c_4\left(\frac{\delta n}{n_0}\right)E'+c_5\left(\frac{\delta B}{B_0}\right)E'+2i\gamma' E'=0.\end{aligned}\tag{5}$$

Here, Equation (5) depicts the normalized dynamical equation for the beam-driven whistler wave, which leads to the turbulence. To describe the beam instability, we have phenomenologically incorporated γ', which represents the normalized growth rate and in our simulations $\gamma=0.5083\omega_{ce}$, which is consistent with the observations of Zhao *et al.*,[29]

$$\begin{aligned}\text{where}\quad c_1&=\frac{-4\omega_0^2\lambda_e^4k_{0x}^2}{2\omega_0(1+\lambda_e^4k_{0x}^4+\lambda_e^4k_{0z}^4+2\lambda_e^2k_{0x}^2+2\lambda_e^2k_{0z}^2)x_n^2},\\c_2&=\frac{(4\lambda_i^2v_A^2k_{0z}^2-4\omega_0^2\lambda_e^4k_{0z}^2)}{2\omega_0(1+\lambda_e^4k_{0x}^4+\lambda_e^4k_{0z}^4+2\lambda_e^2k_{0x}^2+2\lambda_e^2k_{0z}^2)z_n^2},\\c_3&=\frac{(4\lambda_i^2v_A^2k_{0x}k_{0z}-8\omega^2k_{0z}k_{0z}\lambda_e^4)}{2\omega_0(1+\lambda_e^4k_{0x}^4+\lambda_e^4k_{0z}^4+2\lambda_e^2k_{0x}^2+2\lambda_e^2k_{0z}^2)z_nx_n},\\c_4&=\frac{9\lambda_i^2v_A^2k^2k_{0z}^2}{2\omega_0(1+\lambda_e^4k_{0x}^4+\lambda_e^4k_{0z}^4+2\lambda_e^2k_{0x}^2+2\lambda_e^2k_{0z}^2)},\\c_5&=(-9k_{0z}^4\omega_{ce}^2\lambda_e^4-4k_{0x}^2k_{0z}^2\lambda_e^4\omega_{ce}^2).\end{aligned}$$

Also, $\gamma'=\frac{\gamma}{\omega_0}$, where $\gamma(=0.5083\omega_{ce})$ is the normalized growth rate, and $\gamma(=0.5083\omega_{ce})$ is the beam-driven whistler wave growth rate.

B. Magnetosonic waves (MSWs)

Consider a magnetosonic wave propagating in the x-direction, i.e., $\vec{k}=k_{0x}\hat{x}$ along with the background magnetic field in the z-direction, i.e., $\vec{B}=B_0\hat{z}$. The electric field is polarized in the y-direction, i.e., $\vec{E}=E_y\hat{y}$. The dynamical equation for the magnetosonic wave is obtained from the following basic equations:

Equation of motion:

$$m_j\left[\frac{\partial\vec{v}_j}{\partial t}+(\vec{v}_j.\nabla\vec{v}_j)\right]=q_j\vec{E}+\frac{q_j}{c}(\vec{v}_j\times\vec{B})+F_j.\tag{6}$$

Equation of continuity:

$$\frac{\partial n_j}{\partial t}+\vec{\nabla}.(n_j\vec{v}_j)=0.\tag{7}$$

Faraday's law:

$$\nabla\times\vec{E}=-\frac{1}{c}\frac{\partial\vec{B}}{\partial t}.\tag{8}$$

Ampère's–Maxwell's equation:

$$\nabla\times\vec{B}=\frac{4\pi}{c}\vec{J}+\frac{1}{c}\frac{\partial\vec{E}}{\partial t}.\tag{9}$$

Here j denotes the particle species ($j=i$ for ions and e for electrons) and m_j, v_j, q_j, T_j denotes the mass, velocity, charge, and temperature of the species, respectively. n_0 is the background density. F_j is the ponderomotive force due to the whistler wave given below:

$$\vec{F}_j=-m_j(\vec{v}_j*.\vec{\nabla})\vec{v}_j+\frac{q_j}{c}(\vec{v}_j*\times\vec{B}).\tag{10}$$

Published under an exclusive license by AIP Publishing

TABLE I. Testing of invariants with grid points 64.

Time steps (dt)	10^{-5}	10^{-6}	10^{-7}	10^{-8}
N	9.999 999 97	9.999 999 996	1.000 000 000 07	1.000 000 000 07

TABLE II. Testing of invariants with time steps $dt = 5 \times 10^{-5}$ (this step size $dt = 5 \times 10^{-5}$ is also used for the finite difference method to monitor the invariants of the NLS equation to desired accuracy).

Grid points	64	128	256	512
N	9.999 999 1	9.999 997	9.999 999 4	9.999 999 8

Using the velocity components of electrons and ions (see the Appendix for velocity components) in Eq. (10), and then using Eq. (7) with approximation $\omega \ll \omega_{ci}$, we obtained the following form of equation:

$$\left(\frac{\partial^2}{\partial t^2} - c_6 \frac{\partial^2}{\partial x^2}\right) \frac{n}{n_0} = -\frac{\partial}{\partial x}\left(\frac{F_{ey} + F_{iy}}{m_i}\right). \tag{11}$$

Furthermore, on putting the values of ponderomotive components in Eq. (11), we obtain the following equation:

$$\left(\frac{\partial^2}{\partial t^2} - c_6 \frac{\partial^2}{\partial x^2}\right) \frac{n}{n_0} = -c_7 \frac{\partial^2}{\partial x^2} |E_y^2|, \tag{12}$$

where c_6 and c_7 are constants, whose value is given below:

$$c_6 = c^2 \left(\frac{c_s^2 + v_A^2}{c^2 + v_A^2}\right)$$

and

$$c_7 = \left[\frac{e^2}{m_e m_i \omega_{ce}^2} - \frac{e^2 \omega_0}{m_e m_i \omega_{ce} \alpha_1}\left\{\left(1 + \frac{\omega_{ce}^2}{\alpha_1}\right) + \frac{\omega_0^3}{\alpha_1 \omega_{ce}}\right\} + \frac{e^2}{m_i^2 \omega_{ci}^2} - \frac{e^2 \omega_0}{m_e m_i \omega_{ce} \alpha_1}\left\{\left(1 + \frac{\omega_{ci}^2}{\alpha_1}\right) + \frac{\omega_0^3}{\alpha_1 \omega_{ci}}\right\}\right].$$

Also, $c_s (= \frac{v_{the}^2 + v_{thi}^2}{m_i})^{\frac{1}{2}}$ is the speed of the sound wave and $\alpha_1 = -k^2 \lambda_e^2 (\omega_0^2 - \omega_{ce}^2) - \omega_0^2$.

Now, using the same normalizing parameter as we have used in the dynamics of the whistler wave, i.e., x_n, z_n, and t_n and after taking the adiabatic response of Eq. (12), we obtain

$$E_n = \left[-\frac{1}{n_0}\left(\frac{c_6 x_n^2}{c_7 t_n^2}\right)\right]^{\frac{1}{2}},$$

and the following normalization equation:

$$n' = \frac{c_7}{c_6} |E_y|^2. \tag{13}$$

Substituting Eq. (13) in Eq. (5), we obtain

$$i\frac{\partial E'}{\partial t'} + c_1 \frac{\partial^2 E'}{\partial x'^2} + c_2 \frac{\partial^2 E'}{\partial z'^2} + i\frac{\partial E'}{\partial x'} + i\frac{\partial E'}{\partial z'} + c_3 \frac{\partial^2 E'}{\partial x' \partial z'} + \frac{c_7 c_4}{c_6} |E_y^2| E' + c_5 n E' + 2i\gamma' E' = 0. \tag{14}$$

III. NUMERICAL SIMULATION

Using the 2D pseudo-spectral approach, numerical simulation has been performed. We implemented a pseudo-spectral method for space integration (in x-direction) having periodic length $L_x = 2\pi/\alpha_x$ and a finite difference method with a modified version of the Gazdaz[43] predictor-corrector method (in the z-direction) for time evolution. For solving the system of dimensionless equations, the authors have studied the algorithm of the nonlinear Schr ö dinger (NLS) equation in following two steps:

(i) The testing of the invariants plasmon number $N = \sum_k |E_k|^2$ with grid points and time steps is given below (NLS) (Tables I and II).

(ii) Modified for the present case and applied to magnetopause parameters below:
For numerical simulation, equations have been solved in the periodic domain $(10\pi \times 10\pi)$ with the grid size (512×512). The initial conditions used for the numerical simulations are given in Eq. (15), and their parameters are listed in Table III

$$E_y(x, z) = E_0(1 + \beta \cos(\alpha_x x))(1 + \beta \cos(\alpha_z z)). \tag{15}$$

We have presented the results of the numerical simulation of dimensionless equations (5), (12), and (14) applicable to the magnetopause region. For application purposes, the typical parameters for numerical simulations are given in Table IV.

For solving the set of dimensionless Eqs. (5), (12), and (14), we have developed two models:

(a) modified nonlinear Schrödinger equation (MNLS model) and
(b) modified Zakharov system of equations (MZSE model).

After checking the accuracy of the algorithm, it has been modified for the set of dimensionless equations, which is used to study the nonlinear coupling of whistler and magnetosonic waves.

First, we present the results of the numerical simulation of the MNLS model. For the MNLS model, the numerical simulation of Eq. (14) was carried out. Figure 1 depicts the spatial evolution of the electric field intensity of the whistler in the x-z plane. The figure shows that at an initial time, the electron beam imparts energy to the system, and localized structures are formed with increasing amplitude.

TABLE III. Parameters used for initial conditions in numerical simulations.

Parameters	Initial amplitude E_0	Perturbed electric field β	Perturbation in wavenumbers	
			α_x (Normalized by x_n^{-1})	α_z (Normalized by z_n^{-1})
Values	0.5	0.1	0.2	0.2

Published under an exclusive license by AIP Publishing

TABLE IV. Various parameters used for numerical simulations.

Parameters	Values
Background magnetic field, B_0	4.5×10^{-5} G
Ion temperature (T_i) and electron temperature (T_e)	$T_i = 235$ and $T_e = 37$ eV
Background density, n_0	$n_0 = 21\ \text{cm}^{-3}$
Normalized magnetic field B_n, and electric field E_n	$B_n = 1.60 \times 10^{-6}$ G and $E_n = 6.8 \times 10^{-7}$ StatV/cm
x_n, z_n, and t_n	$x_n = 1.05 \times 10^5$ cm, $z_n = 1.54 \times 10^5$ cm, $t_n = 6.31 \times 10^{-4}$ s
Constants	$c_1 = -0.2681$, $c_2 = 1.0080$, $c_3 = 2.4398$, $c_4 = 1.1698$, $c_n = 0.1775$, $c_6 = 1.5413 \times 10^5$, $c_7 = 5.4 \times 10^5$

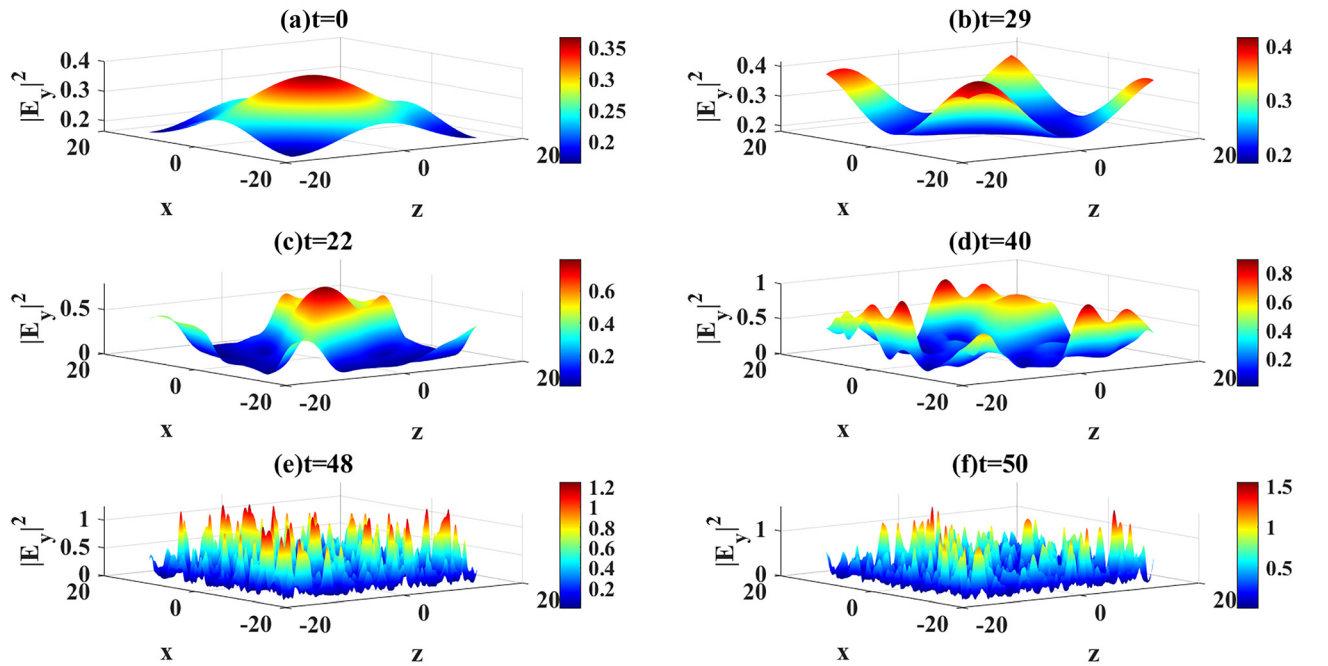


FIG. 1. The normalized spatial evolution of whistler's electric field in 3D in the magnetopause by the MNLS model at (a) t = 0, (b) t = 29, (c) t = 22, (d) t = 40, (e) t = 48, and (f) t = 50.

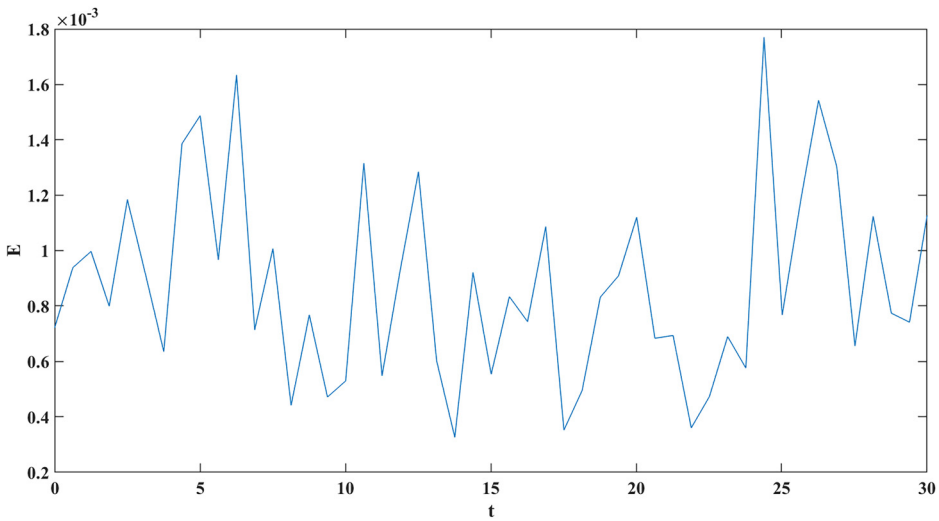


FIG. 2. Evolution of electric field fluctuations with time by the MNLS model.

Published under an exclusive license by AIP Publishing

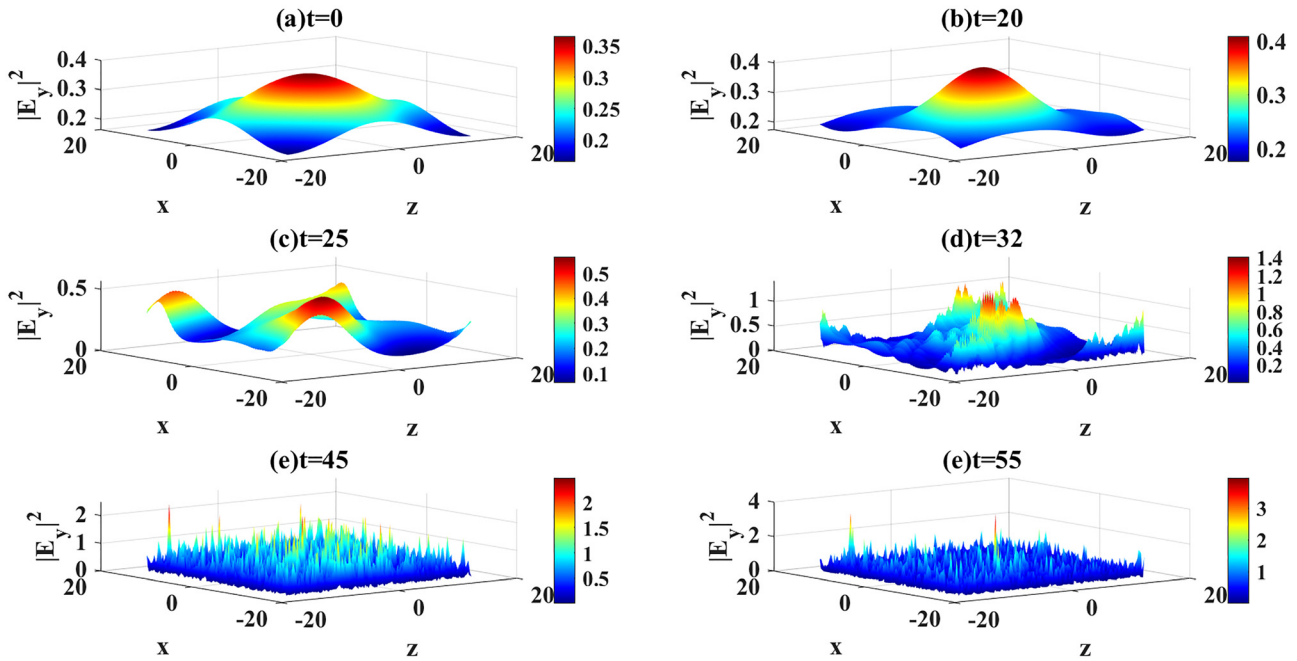


FIG. 3. The normalized spatial evolution of whistler's electric field in 3D in the magnetopause by the MZSE model at (a) t = 0, (b) t = 20, (c) t = 25, (d) t = 32, (e) t = 45, and (f) t = 55.

However, at later stage when the whistler amplitude reaches its peak value then the system transit to the turbulent regime, and we get chaotic structures. These localized structures are the results of the ponderomotive force nonlinearity-driven density modification. Figure 2 depicts the fluctuations of the electric field with time. It shows the turbulent whistler and for the present computation, the whistler fluctuations peak value is around 1.8 mV/m for the present parameters. We have also checked that most of the power is near the pump frequency of the whistler in the frequency band $0.1f_{ce} - 0.5f_{ce}$ as observed by Zhao *et al.*[29]

Figure 3 depicts the spatial evolution of the electric field intensity of the whistler wave for the MZSE model. Initially, the electron beam imparts energy to the system and at later stage when the whistler attains the peak value of its amplitude then the system transit to the turbulent regime, and corresponding localized structures are formed. Figure 4 depicts the 3D spatial evolution of the density profile at normalized time t = 50. It shows the density cavities and humps in the real space associated with the MSW induced by ponderomotive nonlinearity of whistler waves. The low-frequency MSWs' perturbation couples nonlinearly with the pump whistler wave as a result of ponderomotive force. The whistler wave contributes energy to the perturbation as a result of this nonlinear interaction. Due to the low-frequency MSW's disturbance of the background density, the whistler's phase velocity is altered. Whistler waves get focused and defocused as a result of phase velocity variation, which localizes the wave field.

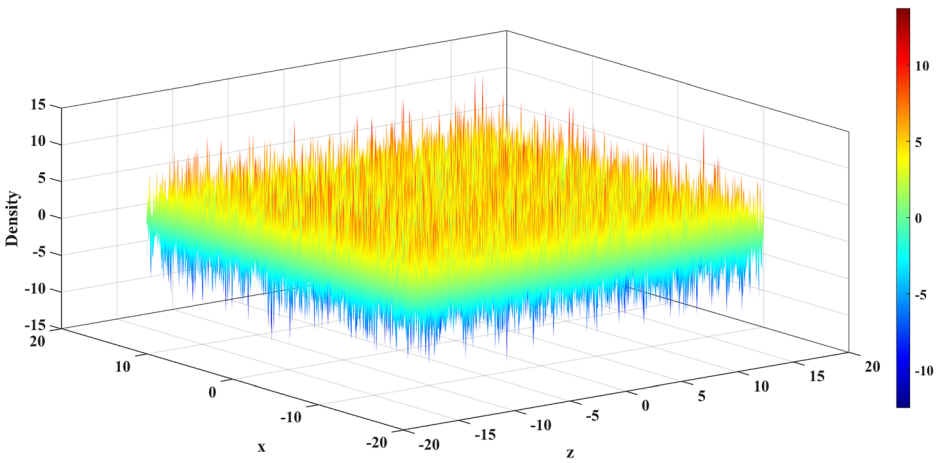


FIG. 4. 3D spatial plot of density cavitation at t = 40 with accumulation and depletion regions as acquired by numerical simulation.

Published under an exclusive license by AIP Publishing

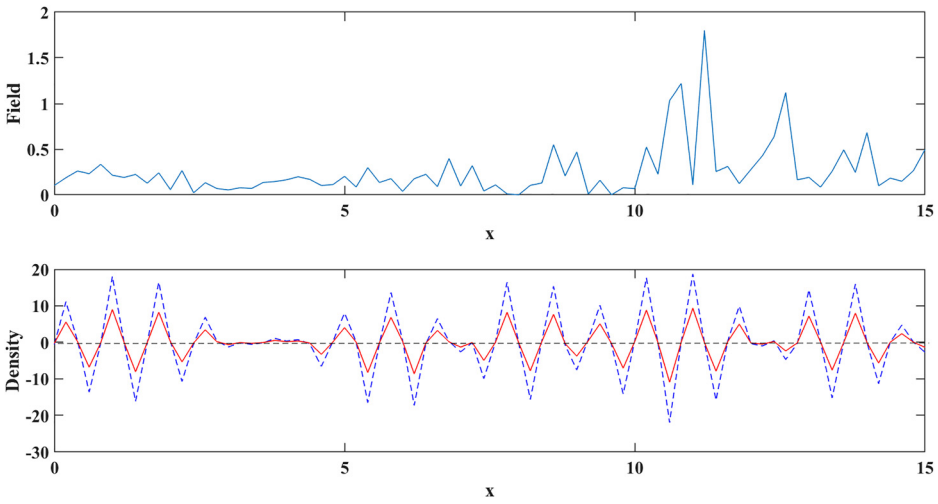


FIG. 5. 2D spatial plot of density cavitation and electric field fluctuations for MZSE at t = 40.

Consequently, the whistler breaks down into filamentary structures. This localization process reveals energy exchange between the pump whistler and magnetosonic waves. Figure 5 illustrates the density fluctuations and amplitude variation of the electric field for the MZSE. As a consequence of the ponderomotive force, the electric field gets trapped in the density cavity region. The modified Zakharov equation's electric fluctuations are shown in the top panel of Fig. 5, and density fluctuations (density cavitation and humps) for MZSE and MNLS are shown in the bottom panel (red and blue dotted lines correspond to MNLS and MZSE, respectively). This graphic demonstrates that MZSE and MNLS have slightly different density values, as would be expected. It should be noted that Eq. (12) was used to plot density fluctuations for MNLS. Figures 6(a) and 6(b) depict the wavenumber spectrum on the (512×512) grid size for MNLS and MZSE models for time ensembled average T = 50–60 and at T = 53, respectively. In the current analysis, the constant growth rate of the beam-driven mode $(\gamma = 0.5083\omega_{ce})$ has been taken into account, and the quasi-steady state has not yet been achieved and will not be for some time. Therefore, it is unnecessary to explain the power spectra, but we are planning to do our next problem by overcoming this shortcoming by considering the intensity-dependent growth rate. However, after the whistler wave's amplitude is large enough to be seen, the beam-driven mode's growth rate is switched-off. Then we remodel Eq. (14) from 2D to 1D in the following form:

$$i\frac{\partial E'}{\partial t'} + c_2\frac{\partial^2 E'}{\partial z'^2} + i\frac{\partial E'}{\partial z'} + \frac{c_7 c_4}{c_6}|E_y^2|E' + c_5 nE' = 0. \tag{16}$$

Then we simulated Eq. (16) for grid size (4096×4096). The power spectra [Fig. 7(a)] of whistler waves have been studied to investigate the energy cascading from larger to lower scales due to the localization process. The ensemble average for several electric field power spectra at the dynamic time range is implemented graphically for the quasi-steady state. The plot of the ensembled averaged $|E_k|^2$ against k has been considered in this study, which is an indicative of the power possessed by spatial modes between normalized time t = 65–80, consisting of 15 spectra [Fig. 7(a)]. Figure 7(b) depicts the localized structure of whistler waves with high amplitude.

IV. SEMIANALYTICAL MODEL

We have created a semianalytical model to help us better grasp the physics underlying this localization of field. Also, we have determined the characteristic transverse scale size of the localized structure using this model.

For this steady-state model, the whistler and MSWs propagation in the x-z plane must be considered. In this case, we deduced the following form from Eq. (5):

$$a_1\frac{\partial E_y}{\partial z} + a_2\frac{\partial^2 E_y}{\partial x^2} + a_3\left(\frac{\delta n}{n_0}\right)E_y + a_4\left(\frac{\delta B}{B_0}\right)E_y + E_y\exp(k_i z) = 0, \tag{17}$$

where a_1, a_2, a_3, and a_4 are constants defined as

$$a_1 = 2ik_{0z}\frac{\left(-2\omega_0^2 k_{0z}^2\lambda_e^4 + 2\omega_0^2 k_{0x}^2\lambda_e^4 + 2\omega_0^2\lambda_e^2 + k_{0x}^2\lambda_i^2 v_A^2 + 2k_{0z}^2\lambda_i^2 v_A^2\right)}{\left(9\lambda_i^2 v_A^2 k_{0z}^2 k^2\right)},$$

$$a_2 = -\frac{4\omega_0^2\lambda_e^4 k_{0x}^2}{9\lambda_i^2 v_A^2 k_{0z}^2 k^2}. \quad a_3 = -\frac{\left(9k_{0z}^4\omega_{ce}^2\lambda_e^4 + 4k_{0x}^2 k_{0z}^2\lambda_e^4\omega_{ce}^2\right)}{9\lambda_i^2 v_A^2 k_{0z}^2 k^2},$$

$$\text{and} \quad a_4 = 1 + \frac{-4k_{0z}^4\omega_{ce}^2\lambda_e^4 - 2k_{0x}k_{0z}^2\lambda_e^4\omega_{ce}^2 - 2k_{0z}k_{0z}^2\lambda_e^4\omega_{ce}^2 - 4k_{0z}k_{0x}\omega_{ce}^2\lambda_e^4 - k_{0z}^4\lambda_e^4\omega_{ce}^2}{9\lambda_i^2 v_A^2 k_0^2 k_{0z}^2}.$$

Published under an exclusive license by AIP Publishing

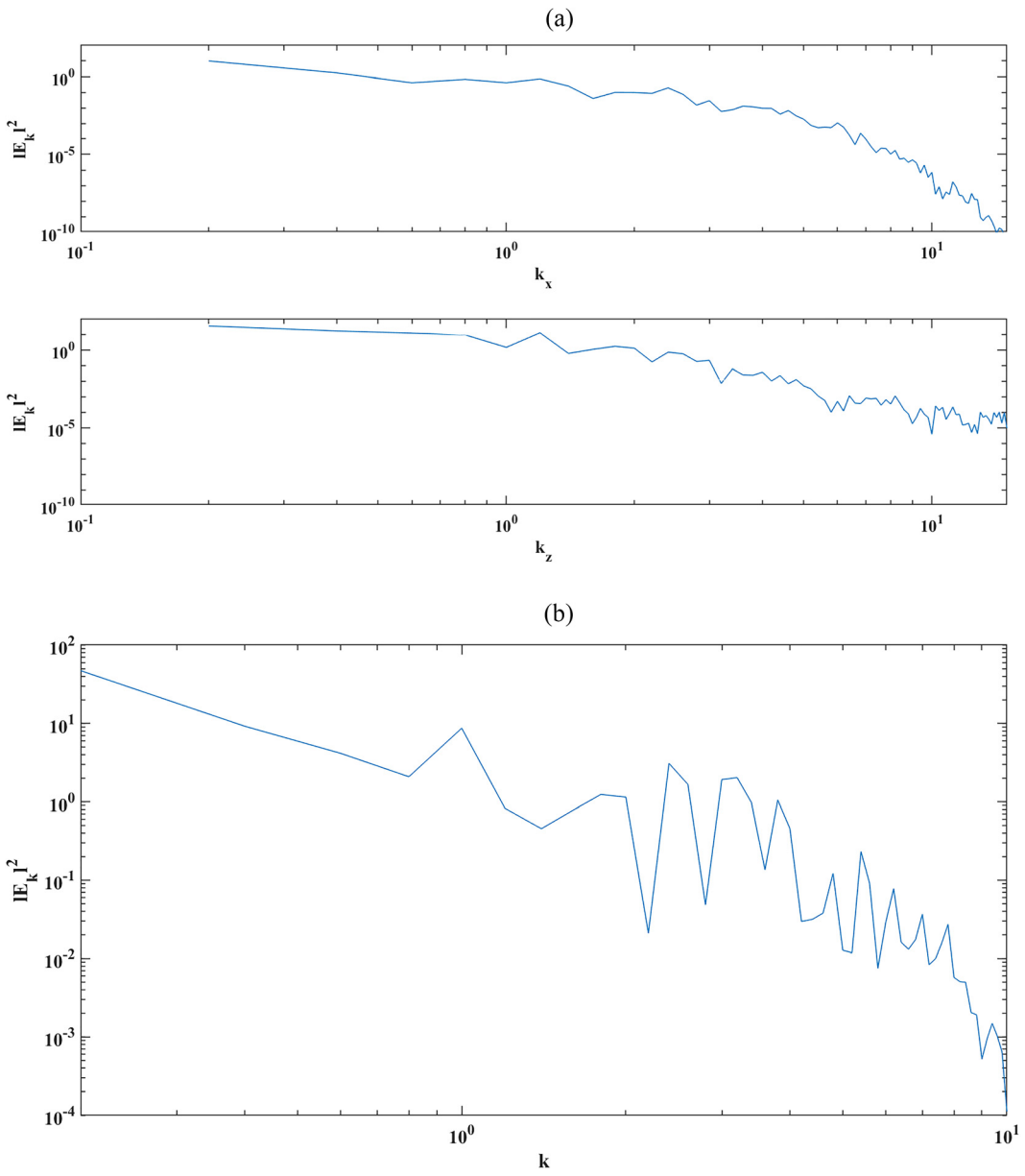


FIG. 6. (a) Evolution of electric field fluctuations' power spectra at t = 59 by the MNLS model. (b) Evolution of electric field fluctuations' power spectra at t = 53 by the MZSE model.

Self-focusing of the beam is described in Eq. (17).

Taking the variation in the electric field, $\tilde{E}_y$ in terms of function eikonal "s" is as follows:

$$\tilde{E}_y = E_0(x,z)e^{ik_{0z}s(x,z)}. \tag{18}$$

Now, using Eq. (18) into Eq. (17) and equating the real and imaginary parts, we obtain

$$-ia_1\frac{\partial S}{\partial z}E_0 + a_2\frac{\partial^2 E_0}{\partial x^2} + a_2k_{0z}^2\left(\frac{\partial S}{\partial x}\right)^2 E_0 + \left[\exp\left(\alpha'|E_x|^2\right) - 1\right]E_0 = 0, \tag{19}$$

and

$$\frac{ia_1}{k_{0z}}\frac{\partial E_0}{\partial z} + 2a_2k_{0z}\frac{\partial S}{\partial x}\frac{\partial E_0}{\partial x} + a_2k_{0z}\frac{\partial^2 S}{\partial x^2}E_0 + E_0\exp\left(k_iz\right) = 0. \tag{20}$$

Additionally, assuming the solution of Eqs. (19) and (20) are of the Gaussian beam profile, we have

$$E_0^2 = \frac{E_{00}^2}{f_1}\exp\left(\frac{-x^2}{r_0^2f_1^2} + k_iz\right), \tag{21}$$

$$S = \beta_1(z)\frac{x^2}{2} + \phi(z). \tag{22}$$

Published under an exclusive license by AIP Publishing

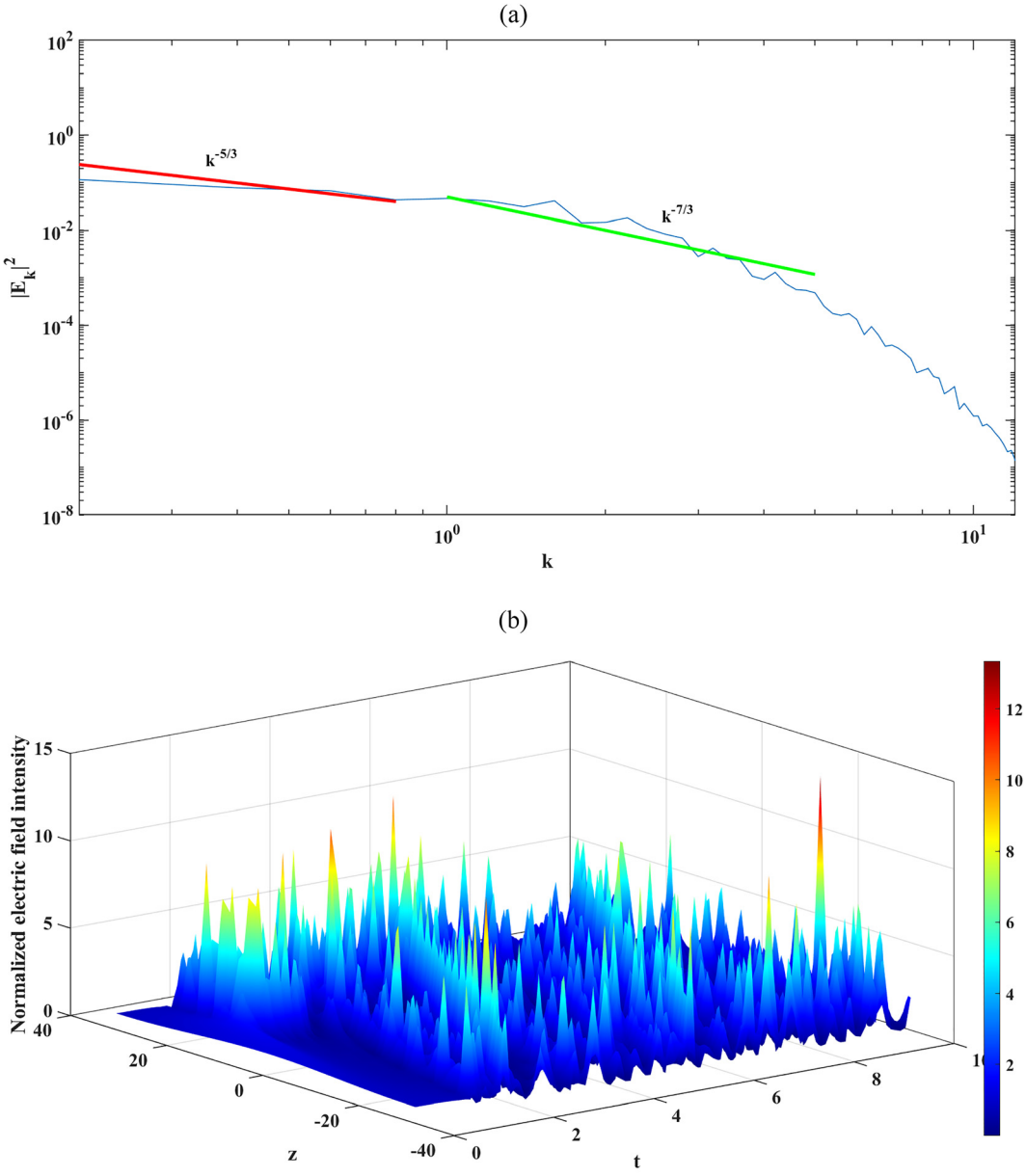


FIG. 7. (a) Power spectra of the electric field intensity of the whistler wave $|E_k|^2$ against normalized wave vector k for the modified nonlinear Schrödinger equation in quasi-steady state (in normalized units) on 4096 × 4096 grid size. (b) The evolution of the electric field $|E_y(t,z)|^2$ at different times in the t-z plane for the modified nonlinear Schrödinger equation in quasi-steady state (in normalized units).

Substituting Eqs. (21) and (22) into Eq. (20), we obtain

$$\beta_1 = \frac{a}{f_1}\frac{df_1}{dz},$$
$$a = -\frac{(2\omega_0^2 k_{0z}^2\lambda_e^4 + 2\omega_0^2 k_{0x}^2\lambda_e^4 + 2\omega_0^2\lambda_e^2 + k_{0x}^2\lambda_i^2 v_A^2 + 2k_{0z}^2\lambda_i^2 v_A^2)}{4\omega_0^2\lambda_e^4 k_{0x}^2}. \quad (23)$$

Here, β_1 is the slowly increasing function of z, r_0 is the characteristic transverse scale size of the whistler wave along the x-direction, and f_1 is the beam width parameter of the wave.

Substituting Eqs. (21) and (22) into Eq. (19) and equating coefficients of x^2, we obtain the dimensionless form beam width parameter

$$\frac{d^2 f_1}{d\xi^2} = 4\left(\frac{a_2}{a_1}\right)^2 \frac{R_d}{r_0^2}\left[\frac{2}{r_0^2 f_1^3} - \frac{1}{a_2}\frac{\alpha E_{00}^2}{f_1^2}\exp\left(\frac{\alpha E_{00}^2}{f_1} + k_i z\right)\right], \quad (24)$$

where $R_d (= k_{0z} r_0^2)$ is the diffraction length and $\xi \left(= \frac{z}{R_d}\right)$ is the distance of propagation

Published under an exclusive license by AIP Publishing

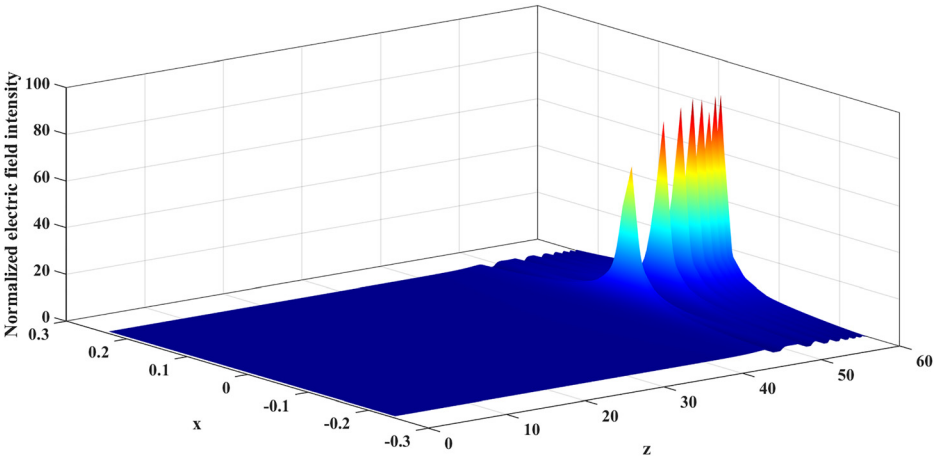


FIG. 8. The normalized electric field intensity variation of the pump wave in the *x-z* plane obtained by the semi-analytical model after incorporating the density harmonics from simulation results.

$$a_1 = \frac{2}{k_{0z}}\left(\frac{-2\omega_0^2 k_{0z}^2\lambda_e^4 + 2\omega_0^2 k_{0x}^2\lambda_e^4 + 2\omega_0^2\lambda_e^2 + k_{0x}^2\lambda_i^2 v_A^2 + 2k_{0z}^2\lambda_i^2 v_A^2}{9\lambda_i^2 v_A^2 k^2}\right),$$

$$\text{and} \quad a_2 = \frac{4\omega_0^2\lambda_e^4 k_x^2}{9\lambda_i^2 v_A^2 k^2 k_{0z}^2}.$$

Equation (24) depicts the variation of f_1 (beam width parameter) and ξ (the distance of propagation). Using the Runge–Kutta method with the boundary conditions: $f_1|_{z=0} = 1$ and $\frac{df_1}{d\xi}|_{z=0} = 0$, Eq. (24) is solved. On the right side, there are two words with opposing signs. First, there is one for the diverging term, and then there is one for the nonlinear term. The fact that both components in this equation have opposing signs indicates that they compete. The first term, which causes the divergence, is a diffracting term. The second term, which results from the ponderomotive force acting on the electrons, is nonlinear. Figure 8 depicts the electric field intensity distribution of a whistler wave in the *x-z* plane. Here, $r_{01} = 3.41 \times 10^7$ cm $(200\lambda_e)$. The electric field is very low $\approx 10^{-3}$ StatV/cm due to which power is also decreasing so the diffraction term dominates till $f \gg 1$, as z propagates, the intensity increases of term $\exp(k_i z)$ and again f decreases less than one and the intensity increases as a result focusing increases.

V. SUMMARY AND DISCUSSION

In the current analysis, we have studied the MNLS and MZSE model involving the nonlinear interactions of a high-frequency whistler wave and low-frequency MSWs, where the coupling among these waves is due to ponderomotive nonlinear force. However, in this scenario, the energetic electron beam source has replaced the magnetic reconnection mechanism. As a result, the dynamics of the beam-driven whistler-mode have been set up with the reckoning that it will mount up from the noise level due to beam energy to large amplitude such that nonlinear effects due to nonlinear ponderomotive force will lead to the localization of the whistler wave, and ultimately, to the turbulent state. On the account of this nonlinear ponderomotive force, the whistler wave gets localized and density cavities and humps are formed. Our attention will be drawn to three key aspects of plasma turbulence: its existence, scaling rules, the contribution of the small-scale coherent structure to plasma heating, and the effect of density fluctuations on the rate at which turbulent energy cascades. In our current analysis, the cavitation and renucleation process generates turbulence by using the MNLS model established by the low-frequency MSW and high-frequency whistler. Numerical simulation has been done to study all these aspects. Also, quasi-steady state has not been achieved due to the constant growth rate of the beam; so to study the quasi-steady state of the system, we remodel the simulation equation [Eq. (14)] into the *t-z* domain for grid size (4096×4096). The consequences of this numerical simulation manifest the presence of turbulence. We have also examined the power spectra of the whistler wave [Fig. 7(a)] in 1D for grid size (4096×4096). For $k\lambda_e < 1$, the power spectrum follows scaling $\sim k^{-5/3}$, and for $k\lambda_e > 1$, it follows scaling $\sim k^{-7/3}$, known as Kolmogorov's scaling as reported in the literature, which may result in the particle acceleration and heating of particles. We have also devised a semianalytical model and determined the transverse scale size of these structures to help us comprehend the physics that underlies localizations. Consequently, owing to nonlinear interactions, the power-law fluctuation of the electric field spectrum gives a huge impact that this noise is a well-grown turbulence process with electric field energy cascading to higher frequencies. The depth of this is in the range of few-electron skin $(0.19\lambda_e)$. The outcomes of the study show the whistler turbulence in the magnetic reconnection sites created by the electron beam and show localized structures and whistler fluctuations, which are to the observations of Zhao *et al.*[29]

The effects of magnetic islands and current sheet generation on whistler turbulence due to magnetic reconnection have not been accounted for in the existing model. We want to carry out this task in the future. In conclusion, our research will help us to improve the quantitative knowledge of whistler-mode waves generated by beams of light in turbulent environments.

ACKNOWLEDGMENTS

Jyoti would like to acknowledge the financial assistance (Award No. 515500) provided by the University Grant Commission (UGC), Government of India.

The content of the paper is original, has independent intellectual property rights, has not been publicly published at home and abroad, has no plagiarism, forgery, and multiple submissions per manuscript, and does not involve confidentiality and other copyright-related infringement issues.

Published under an exclusive license by AIP Publishing

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Jyoti: Conceptualization (equal); Data curation (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Resources (equal); Supervision (equal); Validation (equal); Visualization (equal); Writing – original draft (equal); Writing – review & editing (equal). **Suresh C. Sharma:** Conceptualization (equal); Formal analysis (equal); Resources (equal); Supervision (equal); Validation (equal). **Rampal Sharma:** Conceptualization (equal); Data curation (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Resources (equal); Supervision (equal); Validation (equal); Visualization (equal); Writing – original draft (equal); Writing – review & editing (equal).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

APPENDIX: DISPERSION RELATION OF 2D WHISTLER WAVES

The whistler wave dynamical equation is obtained using a two-fluid model. The whistler wave is propagating in a magnetized plasma with the magnetic field along the z-axis and wave vector, $\vec{k} = k_{0x}\hat{x} + k_{0z}\hat{z}$. Here, $B\,(= B_0 + \delta B)$ is the magnetic field, B_0 is the background magnetic field, and δB is the perturbation in the magnetic field.

To determine the whistler dynamics, the following basic equations are used:

Wave equation:

$$\nabla^2\vec{E} - \nabla(\nabla.\vec{E}) = \frac{4\pi}{c^2}\frac{\partial\vec{J}}{\partial t} + \frac{1}{c^2}\frac{\partial^2\vec{E}}{\partial t^2}. \tag{A1}$$

Equation of motion for electrons:

$$m_e\left[\frac{\partial\vec{v}_e}{\partial t} + (\vec{v}_e.\nabla\vec{v}_e)\right] = -e\vec{E} - \frac{e}{c}\left[\vec{v}_e \times (\vec{B}_0 + \delta B)\right]. \tag{A2}$$

Equation of motion for ions:

$$m_i\left[\frac{\partial\vec{v}_i}{\partial t} + (\vec{v}_i.\nabla\vec{v}_i)\right] = e\vec{E} + \frac{e}{c}\left[\vec{v}_i \times (\vec{B}_0 + \delta B)\right]. \tag{A3}$$

Using Eq. (A2), we obtain the velocity components of electrons as follows:

$$\begin{aligned}\frac{d\vec{v}_e}{dt} &= \frac{\partial\vec{v}_e}{\partial t} + \vec{v}_e.(\vec{\nabla}\vec{v}_e)\\ &= \frac{-e\vec{E}}{m_e} - \frac{e}{m_e c}(\vec{v}_e \times \vec{B}_0) - \frac{e}{m_e c}(\vec{v}_e \times \vec{B}_0)\frac{\delta B}{B_0},\end{aligned} \tag{A4}$$

where $\vec{v}_e = \vec{v}_{e0} + \vec{v}_{e1}$, $\vec{v}_{e0} = 0 =$ is no drift in equilibrium, and $\vec{B}$ and $\vec{E}$ represent the perturbed magnetic field and electric field, respectively.

From Eq. (A4), we obtain $-i\omega_0\,\vec{v}_{e1} = \frac{-e\vec{E}}{m_e} - \frac{e}{m_e c}(\vec{v}_{e1} \times \vec{B}_0) - \frac{e}{m_e c}(\vec{v}_{e1} \times \vec{B}_0)\frac{\delta B}{B_0}$.

Since $\vec{\omega}_{ce} = \frac{e\vec{B}_0}{m_e c}$, the above equation becomes

$$i\omega_0\,\vec{v}_{e1} = \frac{e\vec{E}}{m_e} + (\vec{v}_{e1} \times \omega_{ce}\hat{z}) + (\vec{v}_{e1} \times \omega_{ce}\hat{z})\frac{\delta B}{B_0}. \tag{A5}$$

The x-component of Eq. (A5) is

$$i\omega_0\,\vec{v}_{e1x} = \frac{e\vec{E}_x}{m_e} + \vec{v}_{e1y}\omega_{ce}\left(1 + \frac{\delta B}{B_0}\right). \tag{A6}$$

The y-component of Eq. (A5) is

$$i\omega_0\vec{v}_{e1y} = \frac{e\vec{E}_y}{m_e} - \vec{v}_{e1x}\omega_{ce}\left(1 + \frac{\delta B}{B_0}\right)$$

or

$$\vec{v}_{e1y} = \frac{e\vec{E}_y}{m_e i\omega_0} - \frac{\vec{v}_{e1x}\omega_{ce}}{i\omega_0}\left(1 + \frac{\delta B}{B_0}\right). \tag{A7}$$

We obtain Eq. (A8) by substituting Eq. (A7) into Eq. (A6),

$$v_{e1x} = \frac{-e(i\omega_0 E_x + \omega_{ce}E_y)}{m_e(\omega_0^2 - \omega_{ce}^2)} - \frac{e\omega_0 E_y}{m_e(\omega_0^2 - \omega_{ce}^2)}\frac{\delta B}{B_0}. \tag{A8}$$

Now substituting Eq. (A8) in Eq. (A7), we obtain

$$v_{e1y} = \frac{e(\omega_{ce}E_x - i\omega_0 E_y)}{m_e(\omega_0^2 - \omega_{ce}^2)} + \frac{e\omega_0 E_y}{m_e(\omega_0^2 - \omega_{ce}^2)}\frac{\delta B}{B_0}. \tag{A9}$$

The z-component of Eq. (A5) is

$$v_{e1z} = \frac{eE_z}{im_e\omega_0}. \tag{A10}$$

Similarly, the velocity component of ions can be expressed as

$$v_{ix} = \frac{e(i\omega_0 E_x - \omega_{ci}E_y)}{m_i(\omega_0^2 - \omega_{ci}^2)} - \frac{e\omega_0 E_y}{m_i(\omega_0^2 - \omega_{ci}^2)}\frac{\delta B_0}{B_0}, \tag{A11}$$

$$v_{iy} = \frac{e(\omega_{ci}E_x + i\omega_0 E_y)}{m_i(\omega_0^2 - \omega_{ci}^2)}E_x + \frac{e\omega_0 E_y}{m_i(\omega_0^2 - \omega_{ci}^2)}\frac{\delta B_0}{B_0}, \tag{A12}$$

$$v_{iz} = \frac{-eE_z}{im_i\omega_0}. \tag{A13}$$

Now the electron current density, $\vec{J}_{e1} = -n_e e\vec{v}_{e1}$, and x, y, and z-components of the current density due to electrons are

$$J_{e1x} = n_e\frac{e^2}{m_e}\left(\frac{i\omega_0 E_x + \omega_{ce}E_y}{\omega_0^2 - \omega_{ce}^2}\right) + \frac{n_e e^2\omega_0 E_y}{m_e(\omega_0^2 - \omega_{ce}^2)}\frac{\delta B}{B_0}, \tag{A14}$$

$$J_{e1y} = -n_e\frac{e^2}{m_e}\left(\frac{\omega_{ce}E_x - i\omega_0 E_y}{\omega_0^2 - \omega_{ce}^2}\right) - \frac{n_e e^2\omega_0 E_y}{m_e(\omega_0^2 - \omega_{ce}^2)}\frac{\delta B}{B_0}, \tag{A15}$$

$$J_{e1z} = -n_e\frac{e^2 E_z}{m_e i\omega_0}. \tag{A16}$$

Similarly, x, y, and z-components of the current density due to ions are

Published under an exclusive license by AIP Publishing

$$J_{i1x} = n_i \frac{e^2}{m_i}\left(\frac{i\omega_0 E_x - \omega_{ci}E_y}{\omega_0^2 - \omega_{ci}^2}\right) - \frac{n_i e^2 \omega_0 E_y}{m_i(\omega_0^2 - \omega_{ci}^2)}\frac{\delta B_0}{B_0}, \tag{A17}$$

$$J_{i1y} = -n_i \frac{e^2}{m_i}\left(\frac{-\omega_{ci}E_x - i\omega_0 E_y}{\omega_0^2 - \omega_{ci}^2}\right) - \frac{n_i e^2 \omega_0 E_y}{m_i(\omega_0^2 - \omega_{ci}^2)}\frac{\delta B_0}{B_0}, \tag{A18}$$

$$J_{i1z} = -n_i \frac{e^2 E_z}{m_i i\omega_0}. \tag{A19}$$

Now, from Eq. (A1), we obtain the x, y, and z components of the wave equation, the x-component of the wave equation, $\nabla^2\vec{E}_x - \frac{\partial}{\partial x}(\nabla.\vec{E}) = \frac{4\pi}{c^2}\frac{\partial \vec{J}_x}{\partial t} + \frac{1}{c^2}\frac{\partial^2 \vec{E}_x}{\partial t^2}$.

Using the equation of the current density, $J_x = n_i e v_{i1x} - n_e e v_{e1x}$ in the above equation, we obtain

$$\frac{\partial^2 \vec{E}_x}{\partial x^2} - \frac{\partial^2 E_z}{\partial x \partial z} = \frac{4\pi}{c^2}\frac{\partial(n_i e v_{i1x} - n_e e v_{e1x})}{\partial t} + \frac{1}{c^2}\frac{\partial^2 \vec{E}_x}{\partial t^2}.$$

Substituting the values v_{i1x} and v_{e1x} in Eqs. (A8) and (A11), we obtain

$$\frac{\partial^2 \vec{E}_x}{\partial x^2} - \frac{\partial^2 E_z}{\partial x \partial z} = \frac{4\pi}{c^2}\frac{\partial}{\partial t}\left[\begin{array}{l} \dfrac{1}{\left(\omega_{ci}^2 + \dfrac{\partial^2}{\partial t^2}\right)}\left(\dfrac{e}{m_i}\dfrac{\partial E_x}{\partial t} + \dfrac{e\omega_{ci}}{m_i}E_y\right) \\ + \dfrac{1}{\left(\omega_{ce}^2 + \dfrac{\partial^2}{\partial t^2}\right)}\left(\dfrac{e}{m_e}\dfrac{\partial E_x}{\partial t} + \dfrac{e\omega_{ce}}{m_e}E_y\right) \end{array}\right] + \frac{1}{c^2}\frac{\partial^2 \vec{E}_x}{\partial t^2}, \tag{A20}$$

or

$$\left(\begin{array}{l} \dfrac{\partial^2}{\partial z^2} - \dfrac{1}{\lambda_i^2\left(\omega_{ci}^2 + \dfrac{\partial^2}{\partial t^2}\right)}\dfrac{\partial^2}{\partial t^2} \\ - \dfrac{1}{\lambda_e^2\left(\omega_{ce}^2 + \dfrac{\partial^2}{\partial t^2}\right)}\dfrac{\partial^2}{\partial t^2} - \dfrac{1}{c^2}\dfrac{\partial^2}{\partial t^2} \end{array}\right) E_x - \left(\begin{array}{l} \dfrac{\omega_{ci}}{\lambda_i^2\left(\omega_{ci}^2 + \dfrac{\partial^2}{\partial t^2}\right)}\dfrac{\partial}{\partial t} + \dfrac{\omega_{ce}}{\lambda_e^2\left(\omega_{ce}^2 + \dfrac{\partial^2}{\partial t^2}\right)}\dfrac{\partial}{\partial t} \\ - \left(\dfrac{1}{\lambda_e^2\left(\omega_{ce}^2 + \dfrac{\partial^2}{\partial t^2}\right)}\dfrac{\partial}{\partial t} + \dfrac{1}{\lambda_i^2\left(\omega_{ci}^2 + \dfrac{\partial^2}{\partial t^2}\right)}\dfrac{\partial}{\partial t}\right)\dfrac{\delta B}{B_0} \end{array}\right) E_y - \frac{\partial^2 E_z}{\partial x \partial z} = 0. \tag{A21}$$

The y-component of the wave equation is

$$\left(\begin{array}{l} \dfrac{\omega_{ci}}{\lambda_i^2\left(\omega_{ci}^2 + \dfrac{\partial^2}{\partial t^2}\right)}\dfrac{\partial}{\partial t} - \\ \dfrac{\omega_{ce}}{\lambda_e^2\left(\omega_{ce}^2 + \dfrac{\partial^2}{\partial t^2}\right)}\dfrac{\partial}{\partial t} \end{array}\right) E_x + \left(\begin{array}{l} \dfrac{\partial^2}{\partial x^2} + \dfrac{\partial^2}{\partial z^2} - \dfrac{1}{\lambda_i^2\left(\omega_{ci}^2 + \dfrac{\partial^2}{\partial t^2}\right)}\dfrac{\partial^2}{\partial t^2} - \dfrac{1}{\lambda_e^2\left(\omega_{ce}^2 + \dfrac{\partial^2}{\partial t^2}\right)}\dfrac{\partial^2}{\partial t^2} \\ - \dfrac{1}{c^2}\dfrac{\partial^2}{\partial t^2} + \left(\dfrac{1}{\lambda_e^2\left(\omega_{ce}^2 + \dfrac{\partial^2}{\partial t^2}\right)}\dfrac{\partial}{\partial t} + \dfrac{1}{\lambda_i^2\left(\omega_{ci}^2 + \dfrac{\partial^2}{\partial t^2}\right)}\dfrac{\partial}{\partial t}\right)\dfrac{\delta B}{B_0} \end{array}\right) E_y = 0. \tag{A22}$$

The z-component of the wave equation is

$$\left(\frac{\partial^2}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \frac{1}{\lambda_e^2}\right) E_z - \frac{\partial^2}{\partial x \partial z} E_x = 0. \tag{A23}$$

From these linear relationships, we have

$$E_y = -\left(\begin{array}{l} \dfrac{\omega_{ci}}{\lambda_i^2\left(\omega_{ci}^2 + \dfrac{\partial^2}{\partial t^2}\right)}\dfrac{\partial}{\partial t} - \\ \dfrac{\omega_{ce}}{\lambda_e^2\left(\omega_{ce}^2 + \dfrac{\partial^2}{\partial t^2}\right)}\dfrac{\partial}{\partial t} \end{array}\right)\left(\begin{array}{l} \dfrac{\partial^2}{\partial x^2} + \dfrac{\partial^2}{\partial z^2} - \dfrac{1}{\lambda_i^2\left(\omega_{ci}^2 + \dfrac{\partial^2}{\partial t^2}\right)}\dfrac{\partial^2}{\partial t^2} - \dfrac{1}{\lambda_e^2\left(\omega_{ce}^2 + \dfrac{\partial^2}{\partial t^2}\right)}\dfrac{\partial^2}{\partial t^2} - \\ \dfrac{1}{c^2}\dfrac{\partial^2}{\partial t^2} + \left(\dfrac{1}{\lambda_e^2\left(\omega_{ce}^2 + \dfrac{\partial^2}{\partial t^2}\right)}\dfrac{\partial}{\partial t} + \dfrac{1}{\lambda_i^2\left(\omega_{ci}^2 + \dfrac{\partial^2}{\partial t^2}\right)}\dfrac{\partial}{\partial t}\right)\dfrac{\delta B}{B_0} \end{array}\right)^{-1} E_x,$$

$$E_z = \frac{\partial^2}{\partial x \partial z}\left(\frac{\partial^2}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \frac{1}{\lambda_e^2}\right)^{-1} E_x.$$

Published under an exclusive license by AIP Publishing

Combining Eqs. (A21)–(A23), by assuming $k\lambda_i \gg 1$ and neglecting the displacement current, we obtain

$$\left(\frac{\partial^2}{\partial z^2}-\frac{1}{\lambda_i^2}-\frac{1}{\lambda_e^2\left(\omega_{ce}^2+\frac{\partial^2}{\partial t^2}\right)}\frac{\partial^2}{\partial t^2}\right)\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial z^2}-\frac{1}{\lambda_e^2\left(\omega_{ce}^2+\frac{\partial^2}{\partial t^2}\right)}\frac{\partial^2}{\partial t^2}+\frac{1}{\lambda_e^2\left(\omega_{ce}^2+\frac{\partial^2}{\partial t^2}\right)}\frac{\partial}{\partial t}\frac{\delta B}{B_0}\right)\left(\frac{\partial^2}{\partial x^2}-\frac{1}{\lambda_e^2}\right)E_x$$
$$-\left(\frac{-\omega_{ci}}{\lambda_i^2}\frac{\partial}{\partial t}+\frac{\omega_{ce}}{\lambda_e^2\left(\omega_{ce}^2+\frac{\partial^2}{\partial t^2}\right)}\frac{\partial}{\partial t}+\frac{1}{\lambda_e^2\left(\omega_{ce}^2+\frac{\partial^2}{\partial t^2}\right)}\frac{\partial}{\partial t}\frac{\delta B}{B_0}\right)\left(\frac{\omega_{ci}}{\lambda_i^2}\frac{\partial}{\partial t}-\frac{\omega_{ce}}{\lambda_e^2\left(\omega_{ce}^2+\frac{\partial^2}{\partial t^2}\right)}\frac{\partial}{\partial t}\right)\left(\frac{\partial^2}{\partial x^2}-\frac{1}{\lambda_e^2}\right)E_x$$
$$+\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial z^2}-\frac{1}{\lambda_e^2\left(\omega_{ce}^2+\frac{\partial^2}{\partial t^2}\right)}\frac{\partial^2}{\partial t^2}+\frac{1}{\lambda_e^2\left(\omega_{ce}^2+\frac{\partial^2}{\partial t^2}\right)}\frac{\partial}{\partial t}\frac{\delta B}{B_0}\right)\frac{\partial^4 E_x}{\partial x^2\partial z^2}=0.$$

On solving the above equation, we obtain

$$\frac{-\lambda_e^2\lambda_i^2\omega_{ce}^2\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial z^2}\right)\left(1-\lambda_e^2\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial z^2}\right)\left(1+\frac{\delta B}{B_0}\right)\right)\frac{\partial^2 E_x}{\partial t^2}-\lambda_e^4\omega_{ce}^2\left(1-\lambda_e^2\frac{\partial^2}{\partial x^2}-\lambda_i^2\frac{\partial^2}{\partial z^2}\right)\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial z^2}\right)}{\lambda_e^6\lambda_i^2\left(\omega_{ce}^2+\frac{\partial^2}{\partial t^2}\right)^2}$$
$$+\frac{\lambda_i^2\left(1-\lambda_e^2\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial z^2}\right)\right)^2\frac{\partial^4 E_x}{\partial t^4}+\omega_{ce}^2\lambda_e^2\left(1-\lambda_e^2\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial z^2}\right)\right)\left(1-\lambda_e^2\frac{\partial^2}{\partial x^2}-\lambda_i^2\frac{\partial^2}{\partial z^2}\right)\frac{\partial^2 E_x}{\partial t^2}+\omega_{ce}^2\lambda_i^2\left(1-\lambda_e^2\frac{\partial^2}{\partial x^2}\right)\frac{\partial^2 E_x}{\partial t^2}}{\lambda_e^6\lambda_i^2\left(\omega_{ce}^2+\frac{\partial^2}{\partial t^2}\right)^2}=0$$

or

$$-\lambda_e^2\omega_{ce}^2\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial z^2}\right)\left(1-\lambda_e^2\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial z^2}\right)\left(1+\frac{\delta B}{B_0}\right)\right)\frac{\partial^2 E_x}{\partial t^2}-\lambda_e^2\omega_{ce}^3\omega_{ci}\left(1-\lambda_e^2\frac{\partial^2}{\partial x^2}-\lambda_i^2\frac{\partial^2}{\partial z^2}\right)\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial z^2}\right)$$
$$+\lambda_i^2\left(1-\lambda_e^2\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial z^2}\right)\right)^2\frac{\partial^4 E_x}{\partial t^4}+\omega_{ce}\omega_{ci}\left(1-\lambda_e^2\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial z^2}\right)\right)\left(1-\lambda_e^2\frac{\partial^2}{\partial x^2}-\lambda_i^2\frac{\partial^2}{\partial z^2}\right)\frac{\partial^2 E_x}{\partial t^2}$$
$$+\omega_{ce}^2\left(1-\lambda_e^2\frac{\partial^2}{\partial x^2}\right)\frac{\partial^2 E_x}{\partial t^2}=0$$

or

$$\left(-\frac{\partial^2}{\partial t^2}-\omega_{ce}^2\right)\begin{bmatrix}-\left(1-\lambda_e^2\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial z^2}\right)\right)^2\frac{\partial^2 E_x}{\partial t^2}-\lambda_i^2 v_A^2\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial z^2}\right)\frac{\partial^2 E_x}{\partial z^2}\\ \left(1+\frac{\delta B}{B_0}\right)+\omega_{ce}\omega_{ci}\lambda_e^2\left(1-\lambda_e^2\frac{\partial^2}{\partial x^2}\right)\left(\frac{\partial^2 E_x}{\partial x^2}+\frac{\partial^2 E_x}{\partial z^2}\right)\end{bmatrix}=0. \tag{A24}$$

Using the conditions $-\frac{\partial^2}{\partial t^2} > \omega_{ce}\omega_{ci}$, we obtain

$$\left(-\frac{\partial^2}{\partial t^2}-\omega_{ce}^2\right)\left[-\left(1-\lambda_e^2\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial z^2}\right)\right)^2\frac{\partial^2 E_x}{\partial t^2}-\lambda_i^2 v_A^2\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial z^2}\right)\left(1+\frac{\delta B}{B_0}\right)\frac{\partial^2 E_x}{\partial z^2}\right]=0. \tag{A25}$$

From Eq. (A25), the whistler wave's dynamical equation is expressed as

$$\frac{-\partial^2 E_x}{\partial t^2}-\lambda_e^4\frac{\partial^6 E_x}{\partial x^4\partial t^2}-\lambda_e^4\frac{\partial^6 E_x}{\partial z^4\partial t^2}-2\lambda_e^4\frac{\partial^6 E_x}{\partial x^2\partial z^2\partial t^2}+2\lambda_e^2\frac{\partial^4 E_x}{\partial x^2\partial t^2}+2\lambda_e^2\frac{\partial^4 E_x}{\partial z^2\partial t^2}=\lambda_i^2 v_A^2\left(\frac{\partial^4 E_x}{\partial z^2\partial x^2}+\frac{\partial^4 E_x}{\partial z^4}\right)\left(1+\frac{\delta B}{B_0}\right). \tag{A26}$$

Published under an exclusive license by AIP Publishing

REFERENCES

[1]G. Vetoulis and J. F. Drake, J. Geophys. Res. **104**, 6919, https://doi.org/10.1029/1998JA900112 (1999).
[2]G. M. Mason, M. I. Desai, R. A. Mewaldt, and C. M. S. Cohen, AIP Conf. Proc. **1516**, 117 (2013).
[3]A. A. Schekochihin, S. C. Cowley, W. Dorland, G. W. Hammett, G. G. Howes, E. Quataert, and T. Tatsuno, Astrophys. J., Suppl. Ser. **182**, 310 (2009).
[4]W. H. Matthaeus and M. L. Goldstein, J. Geophys. Res. **87**, 10347, https://doi.org/10.1029/JA087iA12p10347 (1982).
[5]J. P. Eastwood, T. D. Phan, S. D. Bale, and A. Tjulin, Phys. Rev. Lett. **102**, 035001 (2009).
[6]D. A. Roberts, W. Klein, M. L. Goldstein, and W. H. Matthaeus, J. Geophys. Res. **92**, 11021, https://doi.org/10.1029/JA092iA10p11021 (1987).
[7]R. B. Horne, R. M. Thorne, S. A. Glauert, J. M. Albert, N. P. Meredith, and R. R. Anderson, J. Geophys. Res. **110**, A03225, https://doi.org/10.1029/2004JA010844 (2005).
[8]S. Servidio, P. Dmitruk, A. Greco, M. Wan, S. Donato, P. A. Cassak, M. A. Shay, V. Carbone, and W. H. Matthaeus, Nonlinear Process Geophys. **18**, 675 (2011).
[9]W. H. Matthaeus and S. L. Lamkin, Phys. Fluids **28**, 303 (1985).
[10]G. Kowal, D. A. Falceta-Gonçalves, A. Lazarian, and E. T. Vishniac, Astrophys. J. **838**, 91 (2017).
[11]H. Karimabadi, V. Roytershteyn, W. Daughton, and Y.-H. Liu, Space Sci. Rev. **178**, 303 (2013).
[12]S. Saito, S. P. Gary, H. Li, and Y. Narita, Phys. Plasmas **15**, 102305 (2008).
[13]S. P. Gary, S. Saito, and H. Li, Geophys. Res. Lett. **35**, L02104, https://doi.org/10.1029/2007GL032327 (2008).
[14]O. Chang, S. P. Gary, and J. Wang, J. Geophys. Res. **118**, 2824, https://doi.org/10.1002/jgra.50365 (2013).
[15]D. Shaikh, Mon. Not. R. Astron. Soc. **405**, 2521 (2010).
[16]G. Ganguli, C. Crabtree, M. Mithaiwala, L. Rudakov, and W. Scales, Phys. Plasmas **22**, 112904 (2015).
[17]V. A. Svidzinski, H. Li, H. A. Rose, B. J. Albright, and K. J. Bowers, Phys. Plasmas **16**, 122310 (2009).
[18]R. L. Stenzel, J. Geophys. Res. **104**, 14379, https://doi.org/10.1029/1998JA900120 (1999).
[19]Y. Zhang, H. Matsumoto, and H. Kojima, J. Geophys. Res. **103**, 11841, https://doi.org/10.1029/98JA00760 (1998).
[20]W. Masood, S. J. Schwartz, M. Maksimovic, and A. N. Fazakerley, Ann. Geophys. **24**, 1725 (2006).
[21]L. Y. Li, J. Yu, J. B. Cao, J. Y. Yang, X. Li, D. N. Baker, G. D. Reeves, and H. Spence, J. Geophys. Res. **122**, 5431, https://doi.org/10.1002/2016JA023634 (2017).
[22]S. Y. Huang, Z. G. Yuan, F. Sahraoui, H. S. Fu, Y. Pang, M. Zhou, K. Fujimoto, X. H. Deng, A. Retinò, D. D. Wang, X. D. Yu, and H. M. Li, J. Geophys. Res. **122**, 7188–7196, https://doi.org/10.1002/2016JA023670 (2017).
[23]R. B. Horne, G. V. Wheeler, and H. Alleyne, J. Geophys. Res. **105**, 27597, https://doi.org/10.1029/2000JA000018 (2000).
[24]R. B. Horne, R. M. Thorne, S. A. Glauert, N. P. Meredith, D. Pokhotelov, and O. Santolík, Geophys. Res. Lett. **34**, L17107, https://doi.org/10.1029/2007GL030267 (2007).
[25]B. Lembege, S. T. Ratliff, J. M. Dawson, and Y. Ohsawa, Phys. Rev. Lett. **51**, 264 (1983).
[26]H. Zhang, Z. Liu, F. Wang, B. Liu, L. Wei, and W. Duan, Astron. Space Sci. **367**, 88 (2022).
[27]X. Gao, J. Sun, Q. Lu, L. Chen, and S. Wang, Geophys. Res. Lett. **45**, 8029, https://doi.org/10.1029/2018GL079090 (2018).
[28]S. Saito, Y. Nariyuki, and T. Umeda, Phys. Plasmas **25**, 072118 (2018).
[29]S. Q. Zhao, C. J. Xiao, T. Z. Liu, H. Chen, H. Zhang, M. J. Shi, S. Teng, H. S. Zhang, X. G. Wang, Z. Y. Pu, and M. Z. Liu, J. Geophys. Res. **126**, e2020JA028525, https://doi.org/10.1029/2020JA028525 (2021).
[30]C. Krafft and A. S. Volokitin, Plasma Phys. Control Fusion **62**, 024007 (2020).
[31]G. D. Reeves, G. L. Delzanno, P. A. Fernandes, K. Yakymenko, B. E. Carlsten, J. W. Lewellen, M. A. Holloway, D. C. Nguyen, R. F. Pfaff, W. M. Farrell, D. E. Rowland, M. Samara, E. R. Sanchez, E. Spanswick, E. F. Donovan, and V. Roytershteyn, Front. Astron. Space Sci. **7**, 23 (2020).
[32]Y. Narita, R. Nakamura, W. Baumjohann, K.-H. Glassmeier, U. Motschmann, B. Giles, W. Magnes, D. Fischer, R. B. Torbert, C. T. Russell, R. J. Strangeway, J. L. Burch, Y. Nariyuki, S. Saito, and S. P. Gary, Astrophys. J. Lett. **827**, L8 (2016).
[33]D. A. Gurnett, R. R. Anderson, B. T. Tsurutani, E. J. Smith, G. Paschmann, G. Haerendel, S. J. Bame, and C. T. Russell, J. Geophys. Res. **84**, 7043, https://doi.org/10.1029/JA084iA12p07043 (1979).
[34]V. I. Karpman, "Whistler solitons, their radiation and the self-focusing of whistler wave beams," in *Nonlinear MHD Waves and Turbulence*, Lecture Notes in Physics Vol. 536 (Springer, Berlin, 2007), p. 311.
[35]A. Das, R. Singh, P. Kaw, and S. Champeaux, Phys. Plasmas **9**, 2602 (2002).
[36]D. Biskamp, E. Schwarz, and J. F. Drake, Phys. Rev. Lett. **76**, 1264 (1996).
[37]A. Das and P. H. Diamond, Phys. Plasmas **7**, 170 (2000).
[38]S. Galtier and A. Bhattacharjee, Phys. Plasmas **10**, 3065 (2003).
[39]M.-M. Shen and D. R. Nicholson, Phys. Fluids **30**, 1096 (1987).
[40]G. D. Doolen, D. F. Dubois, and H. A. Rose, Phys. Rev. Lett. **54**, 804 (1985).
[41]R. P. Sharma, P. Stubbe, and A. D. Verga, Phys. Plasmas **101**, 022311 (1996).
[42]Jyoti, S. C. Sharma, N. Pathak, and R. P. Sharma, Phys. Plasmas **29**, 092104 (2022).
[43]J. Gazdag, J. Comput. Phys. **20**, 196 (1976).

Published under an exclusive license by AIP Publishing

Proceedings of the International Conference on Automation, Computing and Renewable Systems (ICACRS 2022)
IEEE Xplore Part Number: CFP22CB5-ART: ISBN: 978-1-6654-6084-2

2022 International Conference on Automation, Computing and Renewable Systems (ICACRS) | 978-1-6654-6084-2/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICACRS55517.2022.10029285

Low Power, Highly Stable and Enhanced Read Speed 7T SRAM

Deberjeet Usham, Malti Bansal*
Department of Electronics & Communication Engineering,
Delhi Technological University (DTU), Delhi-110042, India
*maltibansal@gmail.com

***Abstract*— Cache memory is one of the most important functional part of any system ranging from System on Chips (SoC) to personal computers. The cache memories are composed of millions of SRAM cells. Thus, for the cache memory to be reliable, it is important that SRAM cells have some specific features such as: high stability of the data retained, low static and dynamic power consumption, low read delay. As an attempt to accomplish the incorporation these features, a novel SRAM cell consisting of 7 transistors is proposed in this paper. The cell has separate bitlines for the read and write operation thereby isolating the two operations. Also, the storage node formed by the back-to-back inverters are not disturbed by read operation, this leads to enhanced robustness of data stored. The cell also makes use of dual Vt(threshold voltage) transistors in order to minimize leakage power consumption. Transistors in the cell is configured in such a way that there are minimum number of transistors in the read critical delay path, thereby reducing the read delay. Performance of the proposed model is compared with that of the conventional 6T SRAM cell to showcase the enhancements obtained.**

Keywords—6T SRAM (6 Transistor Static Random Access Memory) , 7T SRAM (7 Transistor SRAM), Dual Threshold voltage transistors, SNM (Static Noise Margin), Butterfly curve

I. INTRODUCTION

Today's technology market has become more memory depend than ever. Every device such as mobile phones and computers performs memory intensive tasks on a day-to-day basis. It is not an exaggeration to state that memory is one of the most important components of any system ranging from microcontrollers and System on Chips (SoCs) to personal computers and workstations. And in the memory hierarchy, cache memory is very crucial for quick processing of data by the CPU (Central Processing Unit). They work as an augmented memory to the CPU by temporarily storing data frequently requested by the CPU. This cache memory is mainly composed of a class of random access memory known as Static Random Access Memory (SRAM). Thus, it is of utmost importance to ensure that the SRAMs used are of highest quality and reliability.

Certain features that can be highlighted to guarantee this are: high stability of the data retained (also termed as robustness), low static and dynamic power consumption, low read delay. But the aforementioned features are threatened because of various reasons. Firstly, in an attempt to reduce the size of devices and transistors, voltages have been scaled down drastically such that it degrades the stability of data stored in the cell [1, 5]. In addition to this, since SRAM cells are a memory element, a significant amount of power consumptions particularly leakage power consumption takes place. The standard 6T SRAM cell and current model SRAM cells are ineffective in addressing these issues [6, 7].

So, in this paper a novel SRAM cell is proposed that intends to solve the above-mentioned problems. The paper is organized into five sections to help in better comprehension of our proposition. Section II gives a brief overview of the SRAM cells which includes the working mechanics of the traditional 6T SRAM cell and also the complications and the shortcomings of the model. Then, in Section III, implementation of the proposed model will be discussed. Section IV will be composed of simulation results and analysis of the same. The paper will be concluded with Section V discussing the conclusion and future scopes.

II. OVERVIEW OF SRAM CELLS

A. Conventional 6T SRAM cell

An SRAM cell is a piece of circuitry that can hold one bit of data, either 0 or 1 as long as there is power supplied. It should also be able to have its data read or written into as long as power supply is not cut off. The SRAM cell that will be used as a point of reference for comparison with the proposed model is the conventional model consisting of 6 transistors. It uses a weak back to back CMOS inverter to hold the data and a pair of access transistors on either side of the back to back inverter to make reading and writing onto the cell fairly easy. The circuitry is shown in Fig. 1.

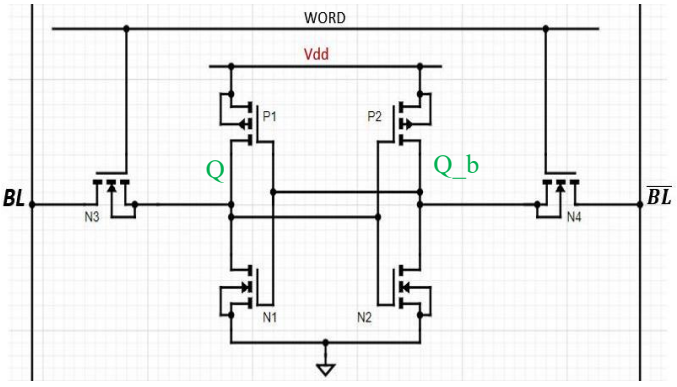


Fig. 1. Conventional 6T SRAM (Reproduced from [2])

To write a data onto the cell, the bit line(BL) and the bitline bar ($\overline{BL}$) are first set equal to the data bit and its complement respectively, then the WORD line is set high to turn on the access transistors (N3 and N4), so as to force the data on the two nodes of the weak cross coupled inverter. To read data from the cell, the bit line(BL) and the bit line bar ($\overline{BL}$) are pre-charged high with the help of a precharge circuit, then the WORD line is set high to turn on the access transistors N3 and N4, so that one among the BL and $\overline{BL}$ is discharged to ground according to the data stored on the nodes.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on February 13,2023 at 06:39:52 UTC from IEEE Xplore. Restrictions apply.

There are two constraints that are to be followed to ensure correct reading and writing of data onto the cell. First is the *read stability* constraint [2]. This constraint states that "the driver transistors N1 and N2 should be stronger than the access transistors N3 and N4" so that during the read operation when one among the BL and $\overline{BL}$ is to be discharged to ground by one of the nodes via the access transistors, the particular node which is 0 is not raised high enough by the current flowing through the access transistor. So, the driver transistors N1 and N2 should be sized bigger than that of the access transistors, N3 and N4. Second is the *writability* constraint [2]. This constraint states that "the access transistors N3 and N4 should be stronger than the pull up transistors P1 and P2". To explain this, suppose Q = 0 and Q_b =1 in the cell. Now a 1 is to be written, i.e, Q has to be forced to 1 by the bitline BL. But due to the read stability constraint as explained above, N1 is stronger than N3 and thus Q can't be forced 1 by the weaker access transistor. As an alternative to accomplish the write operation, Q_b has to be forced to 0 by $\overline{BL}$, this will only be possible if the pull up transistor P2 is weaker than the access transistor N4. To summarize, according to the *writability* constraint, the pull up transistors P1 and P2 must be sized smaller than the access transistors N3 and N4.

The evaluation of resilience to noise in various operation modes is done through the use of a quantity called Static Noise Margins (SNMs), measured in various operation modes, namely read, write and hold [8-10]. Static Noise Margin is defined as the amount of noise that can be applied to the two nodes of the back-to-back inverter before a data stored in the node is flipped.

B. Complications in 6T SRAM

As compact and simple a conventional 6T SRAM is, it has its own shortcomings. First, comes the robustness. The term "robustness" in an SRAM cell quantifies how well a data held (stored) in the nodes is not disturbed during any of the access operations. Configuration of transistors in 6T SRAM is in such a way that the nodes Q and Q_b are directly accessed by the bitlines BL and $\overline{BL}$ via the access transistors, N3 and N4. Careful observation reveals that a voltage division happens in the storage nodes. This voltage division occurs between the transistors in the back-to-back inverters and the access transistor. So, the storage nodes are already at a raised voltage level and thus are vulnerable and can be easily disturbed during a read operation. This is termed as *destructive read* in paper [3]. Second matter of concern is the leakage or static power consumption. As staying static is an inevitable act required in all memories, an SRAM cell suffers from power dissipation during the static state. And since SRAMs mainly make up the cache memories, it is highly desirable that the power consumption is reduced through innovations [11-13].

III. Proposed 7T SRAM Cell

A. Architecture of the proposed model

In an attempt to solve the complications in 6T SRAM cells as discussed above, a novel 7T SRAM cell is proposed, which makes use of two types of MOS (Metal Oxide Semiconductor) transistors, one type with low threshold voltage and the other type with high threshold voltage. The circuit schematic of the proposed 7T SRAM is as shown in Fig. 2.

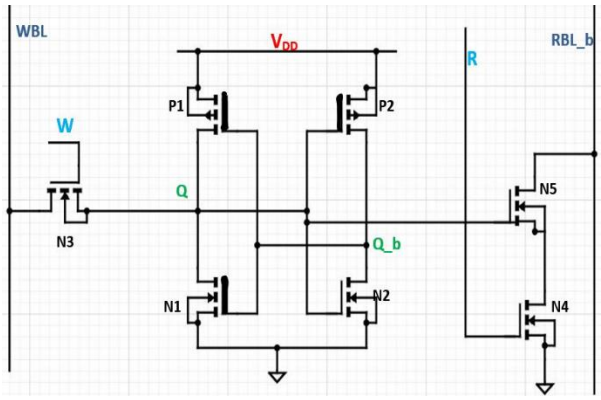


Fig. 2. Proposed 7T SRAM schematic

First noticeable change is the MOS transistors illustrated in thick lines, they are the MOSFETs with high V_T (Threshold Voltage). The main motive behind using high-V_T transistors for 3 of the 4 transistors of the cross-coupled inverter is to reduce static(mainly leakage) power consumption. This can be explain using the mathematical equation that governs the current in off state.

$$I_{OFF} = I_S . e^{\frac{q(V_{GS} - V_T - V_{offset})}{nKT}} (1 - e^{\frac{-qV_{DS}}{KT}}) \quad (1)$$

Where, V_T = Threshold Voltage, and

$$P_{static} = I_{OFF} . V_{DD} \quad (2)$$

As it can be seen from the above equation, as V_T increase, the off state current (subthreshold current) decreases exponentially. So, this in turn decreases the static power consumption.

The next thing noticeable is the use of separate read and write lines (indicated by R and W in the Fig. 2) and also separate write bitline (WBL) and read bitline bar (RBL_b) for writing to and reading the stored data respectively. This is done so as to isolate the read and write operations completely. The write bitline (WBL) alone is used to write a new data onto node Q via the access transistor N3 controlled by the write signal. When node Q is forced with a new data, Q_b is also automatically changed because of the cross coupling of the inverters. The read bitline bar (RBL_bar) alone is used to read the complement of the data from the storage node via the stacked transistors N4 and N5, where the transistor stack is controlled both by the read signal R and also Q_b of the storage node. The purpose of such an architecture can be better understood after understanding the read and write operation.

B. Read and Write operation

To perform a read operation, first the read bitline bar (RBL_b) is precharged to V_{DD}, then the read (R) signal is set high, while write (W) signal is set low so that N4 is ON. Now since N5 is directly connected to Q, if a '1' was stored on Q (which means the cell), N5 would become ON and the RBL_b will be discharged to GND via the series stacked transistors N5 and N4. On the other hand, if a '0' was stored on Q, the RBL_b will stay charged (high). So the RBL_b reads a complement of the bit stored on the cell, and hence

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on February 13,2023 at 06:39:52 UTC from IEEE Xplore. Restrictions apply.

the name. Because of the architecture the storage nodes Q and Q_b are completely isolated from the read bitline bar, since the nodes are not directly accessed by the bitline as was the case in 6T SRAM cell. This enhances the read stability of the proposed SRAM cell. Also, the read bitline bar is conditionally discharged through the transistor stack with the data on storage node acting as one of the control signals. So, the pull up and the pull-down transistors of the back-to-back inverter does not take part in the critical read delay path and hence their sizing does not affect the read delay.

To write onto the cell, the Write Bitline (WBL) is first pre-charged to V_{DD}, then the Write (W) signal is set high, while Read (R) signal is set low. Now if the input data to be written is a '1', the WBL remains charged to V_{DD} and thus the bit '1' is forced to the node Q via the low V_T access transistor N3 (for easy write process). Alternatively, if the input data to be written is a '0', the WBL is discharged to GND and thus the bit '0' is force to node Q via N3. It should be noted here that WBL is required to be precharged to V_{DD} every time regardless of whether the input data is 0 or 1 because of the fact that the write driver that are used in every memory circuitry can only do the job of pulling down the WBL to GND via the control of Write (W) signal and input signal D. The write driver circuit is shown in Fig. 3. for better understanding.

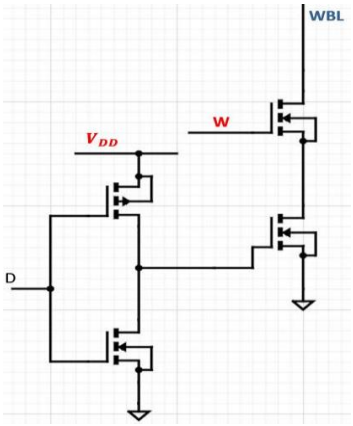


Fig. 3. Write driver used for 7T SRAM cell.

Like the constraints related to read stability and writability in 6T SRAM cell, there are design constraints to be followed here as well. To accomplish a successful write of bit '0' onto the cell, the pull up transistor (P1) must be weaker than the access transistor (N3). Correspondingly, to accomplish a successful write of bit '1' onto the cell, the pull-down transistor N1 must be weaker than the access transistor N3. This is automatically accomplished by the use of high V_T transistors N1 and P1. So, the high V_T transistors serve dual purpose here.

Also, there is a problem of transfer of weak (degraded) '1' through the access nMOS transistor N3 from the WBL due the threshold drop in N3. This is solved by making the inverter formed by transistors P2 and N2 have lower switching threshold so that the weak '1' will not be misinterpreted as a '0'. To make this possible low V_T transistor N3 is used, as evident form the equation defining switching threshold.

$$V_M = \frac{V_{DD} - |V_{Tp}| + V_{Tn}\sqrt{\frac{\beta_n}{\beta_p}}}{1 + \sqrt{\frac{\beta_n}{\beta_p}}} \quad (2)$$

Where, $|V_{Tp}|$= Threshold Voltage of pMOS
V_{Tn}= Threshold Voltage of nMOS.

Thus, it is clear that if the pMOS is of high V_T and the nMOS of low V_T ,the inverter will have a decreased switching threshold, V_M . Therefore, the use of dual V_T is beneficial at many levels while decreasing leakage power consumption remains of the main reasons.

IV. Simulation Results and Analysis

Simulations were performed on the Cadence Virtuoso tool by constructing schematics for the suggested 7T SRAM cell as well as for the conventional 6T SRAM cell for the purpose of comparison. Sizes of transistors used in the 6T SRAM is shown in Table I.

TABLE I. TRANSISTOR SIZES OF 6T SRAM CELL

	Width (W)	Length (L)
N1,N2, (Pull Down Transistors)	400n	100n
P1,P2 (Pull-Up Transistors)	120n	100n
N3, N4 (Access Transistors)	200n	100n

So, the Cell Ratio (β)= ratio of pull-down transistor to the access transistor, is set to 1. And the sizing is in accordance to the read stability and writability constraints as discussed in above section. Similarly, sizes of transistors used for 7T SRAM cell is as shown in Table II.

TABLE II. TRANSISTOR SIZES OF PROPOSED 7T SRAM CELL

	Width (W)	Length (L)
N1,N2 (Pull Down Transistors)	120n	100n
P1,P2 (Pull-Up Transistors)	120n	100n
N3, (Write Access Transistors)	120n	100n
N4, N5 (Read Stacked Transistors)	300n	100n

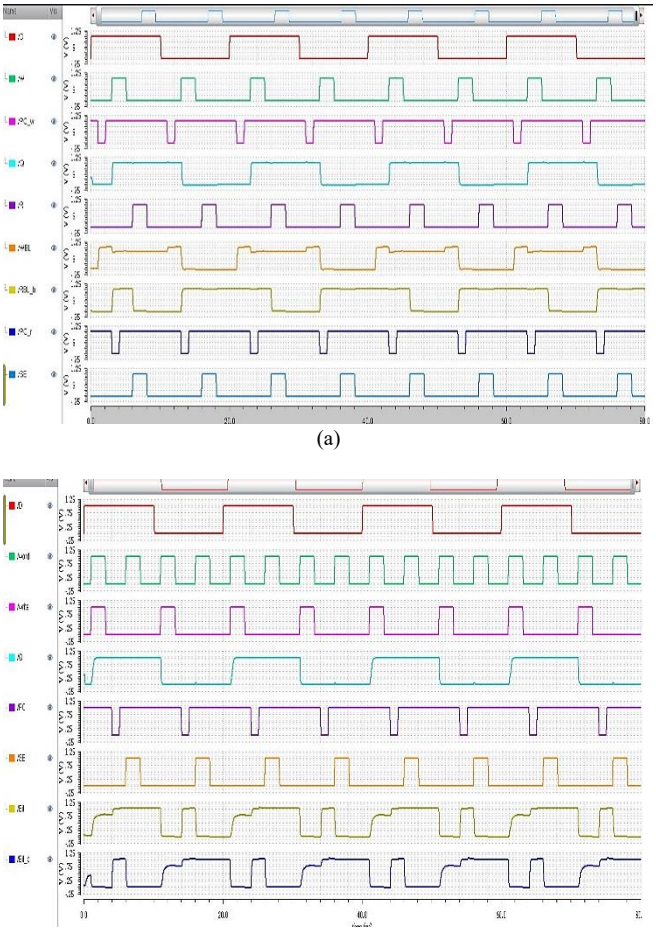
Since the stack transistors are connected in series, it results in increased resistance, so to get a less equivalent resistance overall they needed to be sized double that of the other transistors. All the other transistors have the minimum possible size, thus making the overall area occupied by the circuit much less compared to that occupied by the conventional 6T SRAM despite the fact that the proposed model has 1 extra transistor [14-15].

A. Transient Analysis

Both the conventional and the proposed model were put under transient analysis to verify the correct functioning of the read, write and the hold operations. Both the models were simulated by using the same input data signal, similar write and read signals, and precharge signals for read (RBL) and

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on February 13,2023 at 06:39:52 UTC from IEEE Xplore. Restrictions apply.

write bitline (WBL) (BL and $\overline{BL}$ in case of 6T SRAM). Similar precharge circuits, sense amplifiers and write drivers were used for both the models so that the performance comparison results remain reliable. The simulation outputs are illustrated in Fig. 4.



(a)

(b)

Fig. 4. (a)Transient analysis of proposed 7T SRAM, (b)Transient analysis of 6T SRAM.

As evident from the output of transient analysis, read, write and hold operation is performed satisfactorily well in both the models. For the 7T SRAM, separate precharge signals are used prior to each read and write operation. PC_r signal is used to precharge RBL just before a read operation. Likewise, PC_w is used to precharge WBL just before a write operation. While, in 6T SRAM, one precharge signal, PC is used to precharge BL and $\overline{BL}$ just before a read operation.

B. Data Stability

Stability of the data in read, write and hold operation in an SRAM cell is quantified by Static Noise Margin (SNM) and is measured for the three different operation modes. SNM is defined as the maximum amount of noise that can be applied to the storage nodes before that state stored is flipped and this applies to all three operating modes. SNM is measured by the use of butterfly curves and the detail is explained in paper [3].

The read SNM (RSNM) for both the models are shown in Fig. 5. The read margin is hugely improved by about 3.5 times. This because of the fact than the storage nodes Q and Q_b are completely isolated from the bitlines. To understand this better, suppose, node Q of 7T SRAM is storing a '1' (V_{DD}), then node Q_b will store a proper '0' (GND) and will remain unchanged always. But this is not the case in node Q and Q_b in 6T SRAM cell. Suppose node Q stores a'1' (V_{DD}), node Q_b will not be proper 0 but a slightly higher value because of voltage division happening during the read operation because of the structure formed by the driver transistor (N2) and the access transistor (N4). As a result of the already elevated voltage level at Q_b, the bit stored is most prone to noise while a read operation is taking place.

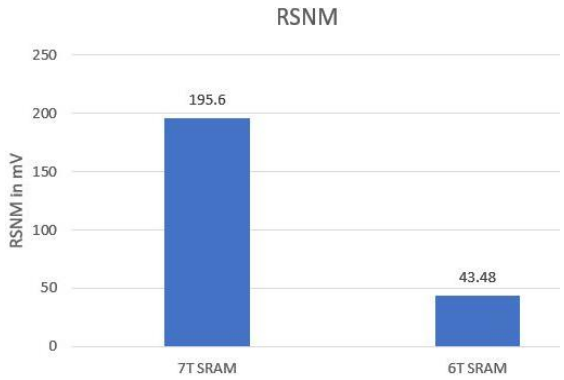


Fig. 5. RSNM comparison between the proposed model and 6T SRAM

The write SNM (WSNM) for both models are illustrated in Fig. 6. The proposed novel 7T SRAM has two write margins, one for writing '0' and the other for writing '1'. This is because of the unsymmetrical nature of the 7T SRAM cell.

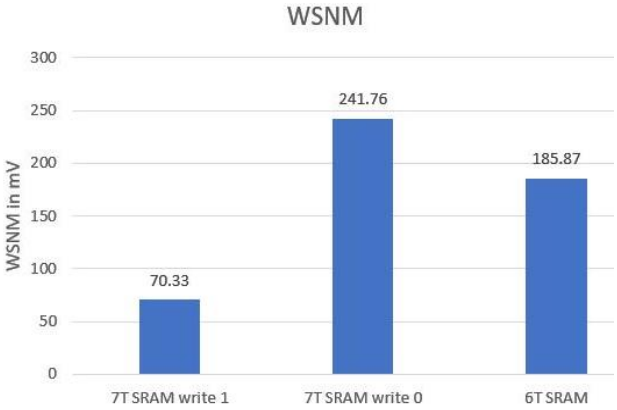


Fig. 6. WSNM comparison between 6T SRAM and proposed model

The WSNM for the 7T SRAM is observed to be small during the write '1' mode because of the use of single bitline for writing purpose. But the margin is till fairly manageable and is compensated by the large margins in other modes. The Hold SNM (HSNM) for both models are shown in Fig 7.

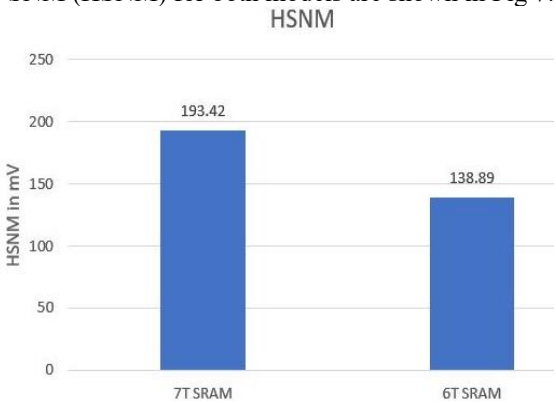


Fig. 7. HSNM comparison between 6T SRAM and proposed model

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on February 13,2023 at 06:39:52 UTC from IEEE Xplore. Restrictions apply.

Hold stability was not a problem for the conventional 6T SRAM, but still it is improved by a small margin in the proposed model.

C. Power consumption

Comparison between the two models in terms of power consumed is illustrated graphically in figure 8.

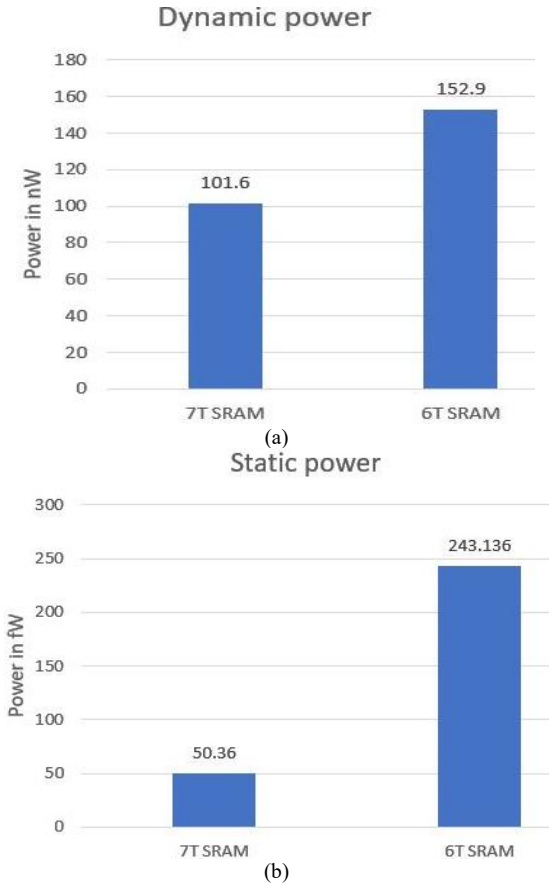


Fig. 8. Comparison of (a)Dynamic power consumption, (b)Static power consumption

The problem of power consumption is one among the major issues faced in the conventional model because of the off-state leakage currents during steady state. But this has been greatly reduced thanks to the usage of high threshold transistors in the back to back inverter. Also the stacked transistors N4 and N5 also help to further reduce leakage power consumption. To understand this better, suppose node Q is storing '0' and is in steady state, then N5 and N4 are both turned off. So, the leakage current through the stack N4 and N5 is significantly reduced due to the stacking effect explained in detail in paper [4]. On the other hand, if a '1' is stored in Q in steady state, then N4 is ON but N5 is OFF, so it is N5 that will have an off-state current. However, the N5 off-state current is also significantly reduced because the drain-to-source voltage of N5 will not be V_{DD} due to voltage drop in N4. This can be understood better with the equation (1). So, as V_{DS} decreases, I_{OFF} decreases exponentially. Also, the dynamic power consumption is also observed to be reduced owing to the use of minimum sized transistors.

D. Read and write delay.

In addition to the improvements mentioned above, the proposed model also has reduced read and write access delay, meaning they can be read or written to faster than the conventional model. The comparison chart is shown in Fig. 9.

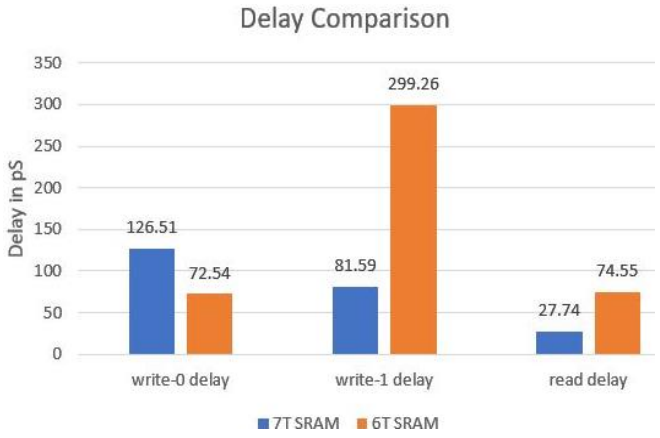


Fig. 9. Delay comparison between 6T SRAM and proposed model

The decrease in read delay can be mainly accredited to the configuration of transistors in the proposed model such that critical read delay path is minimized. The critical read delay path in 6T SRAM comprises of the access transistors (N4 or N5) and the pull down transistors (N1 or N2) in series connection. Whereas in the proposed 7T SRAM, read happens via the stack transistors N4 and N5 that are sized double that of other transistors, thereby reducing the equivalent series resistance and hence the critical read delay path. One may be confused that the usage of high V_T transistors must decrease the current flowing through it and hence increase the delay. But this does not happen because of the fact that the high V_T transistors in the cross coupled inverters do not form any part in the critical read delay path.

One may expect the write delay in 7T SRAM to be more than that of 6T SRAM because of the use of only one bit line to force a data onto the storage nodes. It is indeed a little bit more while writing a '0'. But the delay for writing '1' and the average write delay (i.e. average of writing '1' and '0') is significantly lesser as compared to that of 6T SRAM. This can be accredited to the precharge writing mechanism used in conjunction with the write driver as shown in Fig. 3 and explained in section III(*B*). Because the bit lines are prechared prior to a writing operation, writing 1 is much easier now, even if it is only with one bit line.

V. Conclusion and Future Scope

From all the simulation results and examination of the same, the proposed novel SRAM model is clearly seen to be more enhanced in almost every aspect. First of all, stability of the data stored, which is also stylistically termed as robustness of the cell is significantly better in the proposed SRAM cell. As discussed, this is mainly attributed to the fact the storage nodes are completely secluded from the read bit lines. Secondly, the power consumption is also greatly reduced both in terms of dynamic and static or leakage power consumption. This is primarily attributed to the fact that dual V_T transistors are used to make the cross-coupled inverter and also to the use of minimum sized transistors. Thirdly, the read

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on February 13,2023 at 06:39:52 UTC from IEEE Xplore. Restrictions apply.

and write access speed is also observed to be fairly improved in the proposed 7T SRAM. This is mainly attributed to two things: I) reduction of critical read delay path, and II) the precharge mechanism of the write bit line. Last but not the least, although a comparison of area occupied is not simulated, a fair deduction can be made that the proposed model is more compact despite the use of one extra transistor. This is because of the use of minimum sized transistors as illustrated in Table I and Table II, which is carefully achieved by the use of dual threshold voltage transistors and the configuration of transistors in the cell. As a part of future scope, some more innovations can be added on the proposed 7T SRAM such as using two bit lines to enhance the writing performance.

REFERENCES

[1] P. N. V. Kiran and N. Saxena, "Design and analysis of different types SRAM cell topologies," 2015 2nd International Conference on Electronics and Communication Systems (ICECS), 2015, pp. 1060-1065, doi: 10.1109/ECS.2015.7124742.

[2] Neil Weste and David Harris. 2010. CMOS VLSI Design: A Circuits and Systems Perspective (4th. ed.). Addison-Wesley Publishing Company, USA.

[3] Jose, A.A., Balan, N.C. (2016). Static Noise Margin Analysis of 6T SRAM Cell. In: Dash, S., Bhaskar, M., Panigrahi, B., Das, S. (eds) Artificial Intelligence and Evolutionary Computations in Engineering Systems. Advances in Intelligent Systems and Computing, vol 394. Springer, New Delhi. https://doi.org/10.1007/978-81-322-2656-7_22

[4] V. Kursun and E. G. Friedman, *Multi-Voltage CMOS Circuit Design*, 242 pp., John Wiley & Sons Ltd., 2006, ISBN # 0-470-01023-1.

[5] N. Raghav and M. Bansal, "Analysis of Power Efficient 6-T SRAM Cell with Performance Measurements," 2017 International Conference on Innovations in Control, Communication and Information Systems (ICICCI), 2017, pp. 1-4, doi: 10.1109/ICICCIS.2017.8660819.

[6] Malti Bansal, Vishal Bharti and Vibhor Chander, "Comparison between Conventional Fast Multipliers and Improved Fast Multipliers using PTL Logic", 2021 IOP Conf. Ser.: Mater. Sci. Eng. 1126 012041

[7] Z. Liu and V. Kursun, "High Read Stability and Low Leakage Cache Memory Cell," 2007 IEEE International Symposium on Circuits and Systems, 2007, pp. 2774-2777, doi: 10.1109/ISCAS.2007.378628.

[8] Kiran, P & Saxena, Nikhil. (2015). Design and analysis of different types SRAM cell topologiesDesign and analysis of different types SRAM cell topologies. 167-173. 10.1109/ECS.2015.7124870.

[9] Rath, Subhashree & Panda, Siba. (2017). Analysis of 6T SRAM Cell in Different Technologies. Circulation in Computer Science. MCSP2017. 7-10. 10.22632/ccs-2017-mcsp026.

[10] Shikha Saun and Hemant Kumar," Design and performance analysis of 6T SRAM cell on different CMOS technologies with stability characterization", 2019 *IOP Conf. Ser.: Mater. Sci. Eng.* **561** 012093, DOI 10.1088/1757-899X/561/1/012093

[11] M. Bansal and J. Singh, "Qualitative Analysis of 2-bit CMOS Magnitude Comparator and GDI Magnitude Comparator using FinFET Technology (18nm)," 2020 International Conference on Smart Electronics and Communication (ICOSEC), 2020, pp. 1323-1327, doi: 10.1109/ICOSEC49089.2020.9215251.

[12] A. Bhaskar, "Design and analysis of low power SRAM cells," 2017 Innovations in Power and Advanced Computing Technologies (i-PACT), 2017, pp. 1-5, doi: 10.1109/IPACT.2017.8244888.

[13] R. M. Gangadari and D. S. Yadav, "Extraction and Comparative Inspection of several parameters of 6T, 8T, 10T SRAM," 2021 First International Conference on Advances in Computing and Future Communication Technologies (ICACFCT), 2021, pp. 119-124, doi: 10.1109/ICACFCT53978.2021.9837366.

[14] Bansal, M., Maiya, R.R. (2023). A Taxonomical Review of MOS Power Transistor for Electronic Circuits and Devices. In: Namrata, K., Priyadarshi, N., Bansal, R.C., Kumar, J. (eds) Smart Energy and Advancement in Power Technologies. Lecture Notes in Electrical Engineering, vol 927. Springer, Singapore. https://doi.org/10.1007/978-981-19-4975-3_49

[15] M. Bansal, R. Arora and R. Bharti, "VLSI Layout: Concept to Realization," 2022 International Conference on Applied Artificial Intelligence and Computing (ICAAIC), 2022, pp. 1590-1596, doi: 10.1109/ICAAIC53929.2022.9792835.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on February 13,2023 at 06:39:52 UTC from IEEE Xplore. Restrictions apply.

2022 5th International Conference on Multimedia, Signal Processing and Communication Technologies (IMPACT) | 978-1-6654-7647-8/22/$31.00 ©2022 IEEE | DOI: 10.1109/IMPACT55510.2022.10029271

Machine Learning Based Obfuscated Malware Detection in the Cloud Environment with Nature-Inspired Feature Selection

Mohd. Rehan Ghazi
Dept. of Electronics and Comm. Engineering
Delhi Technological University, Delhi, India
er.rehan.aras@gmail.com

N. S. Raghava
Dept. of Electronics and Comm. Engineering
Delhi Technological University, Delhi, India
nsraghava@dtu.ac.in

Abstract: **One of the most significant and widely used IT breakthroughs nowadays is cloud computing. Today, the majority of enterprises use private or public cloud computing services for their computing infrastructure. Cyber-attackers regularly target Cloud resources by inserting malicious code or obfuscated malware onto the server. These malware programmes that are obfuscated are so clever that they often manage to evade the detection technology that is in place. Unfortunately, they are discovered long after they have done significant harm to the server. Machine Learning (ML) techniques have shown to be effective at finding malware in a wide range of fields. To address feature selection (FS) challenges, this study uses the wrapper-based Binary Bat Algorithm (BBA), Cuckoo Search Algorithm (CSA), Mayfly Algorithm (MA), and Particle Swarm Optimization (PSO), and then k-Nearest Neighbor (kNN), Random Forest (RF), and Support Vector Machine (SVM) are used to classify the benign and malicious records to measure the performance in terms of various metrics. CIC-MalMem-2022, the most recent malware memory dataset, is used to evaluate and test the proposed approach and it is found that the proposed system is an acceptable solution to detect malware.**

Keywords: Cloud Security, Malware detection, Machine learning, Random Forest, SVM, kNN, Cloud Computing.

I. INTRODUCTION

With the usage of a pay-per-use model, cloud computing offers a very flexible and scalable platform for on-demand access to computer system resources, data storage, and computational power [1]. The usage of cloud resources for a variety of always-on services in the private, public, and commercial spheres is starting to spread. In the face of problems like cybersecurity threats, component failures, and erroneous configurations, they need to be secure and resilient [2]. Due to the frequent migration of businesses to third-party cloud providers, there is a considerable danger of new malware being injected into user systems, which could result in data breaches, VM hijacking, and other security issues. The providers also experience issues like service interruptions and reputational harm at the same time. Because there are so many varied apps using shared resources, enforcing security and resilience in a cloud platform is a crucial but difficult topic. The cloud infrastructure must contain a security analysis system that can identify risks or malware.

Malicious software that targets a computer system's confidentiality, integrity, and availability of data is referred to as malware in general [1]. Nearly 1363.92 million malware reports have been made in the last five years, according to AV-Test [3]. Because there is so much unstructured data, malware analysis is quite challenging regularly. Since the 1980s, when malware first surfaced, it has occupied a prominent position in the cybersecurity sector [4]. The user's privacy may be jeopardized, classified information may be accessed, applications and files may be added or removed, and it may cause harm in a variety of other ways. There are several types of malware, including viruses, rootkits, botnets, spyware, bots, ransomware, Trojan horses, worms, and more.

ML algorithms do not work as expected when a dataset includes a large number of features; more specifically, when the extracted feature set has a significant amount of redundant and/or unimportant data. To overcome this problem, researchers apply FS techniques, which reduce the dimension of the feature set by choosing the optimal subset from all the features. Since FS is an NP-hard problem, the most effective solutions are found using methods that are inspired by nature.

The two main categories of malware detection strategies are 1) Signature-based detection and 2) Anomaly-based detection [5]. By matching a piece of malware's pattern to a previously saved signature of that malware in the database, signature-based detection approaches can identify malware. It quickly and reliably detects existing malware, but it misses new threats. Malware is discovered via anomaly-based detection, which looks for deviations in malware execution behavior from expected behavior. It has a high rate of false alarms yet can detect invisible malware [6]. Due to the change in malware behavior, methods of malware detection based on signatures are no longer effective [7], [8]. Because of this, attacks on virtualized systems are a particularly delicate subject that should be handled with extreme attention when offering cloud services.

There are numerous techniques for identifying malware that has been masked that are based on memory analysis. The majority of the works, however, are too complicated to be applied in practical settings. This is the motivation for the proposed efficient and easy-to-implement approach for obfuscated malware detection that takes advantage of the most beneficial features identified through memory analysis. The following is a list of this paper's key contributions:

a. A new optimization method is proposed to address the problem of obfuscated malware detection in cloud systems.
b. Methodology for malware analysis that makes use of ML techniques to differentiate between benign and obfuscated malware samples.
c. The proposed approach relies on nature-inspired optimization techniques for FS that applies the key operators of the BBA [9], CSA [10], MA [11], and PSO [12]Optimizers.
d. The newest dataset, CIC-Mal-Mem 2022 [13], used in obfuscated memory malware detection is used to benchmark the proposed approach.
e. The findings of the proposed method demonstrated its capacity to address the issues with obfuscated memory malware identification.

The organization of the research paper is outlined as: Section II presents a review of the relevant literature on malware, FS approaches, and classification approaches. In

The first author would like to thank the Department of Science & Technology (DST), Ministry of Science and Technology, Government of India, for DST-INSPIRE Fellowship.

978-1-6654-7647-8/22/$31.00 ©2022 IEEE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on February 06,2023 at 05:23:39 UTC from IEEE Xplore. Restrictions apply.

section III, the dataset employed is elaborated. Section IV is comprised of the proposed methodology of FS and classification approach. The performance evaluation metrics with results are discussed in Section V and the conclusion is in Section VI.

II. RELATED WORK

Since its beginnings, malware has gotten a lot of interest in the cybersecurity field due to its various delivery methods and classifications. Even though there are numerous detection methods, each has its drawbacks. This section highlights the pertinent works on malware detection via memory analysis as well as the continuing issues and challenges in this field of research.

Numerous machine-learning algorithms have been investigated in the field of malware detection during the past ten years [1], [5]. Although these models perform quite well, scalability is a problem for them. To combat malware obfuscation techniques, the detection system must be built with obfuscation in mind. This can be accomplished by utilizing a dataset designed specifically to evaluate how well a detection system handles malware that has been disseminated. By using a sizable dataset made up of intricate payload systems, positive and negative memory snapshots, and malware obfuscation, the designers of [14] challenged detection methods.

Malware detection on a cloud is significantly more difficult with more liabilities. As cloud operations expand, using the minHash technique can increase effectiveness and dependability, as several studies like those in [15] have shown. The results of this experiment show the impact of the classifiers on the different kinds of malware that are fed into the system because there are significant differences in how the malware was detected by each classifier. The Long Short-Term Memory-Based (LSTM) approach of Vectorising assembly source code is advised by [16] due to the time and accuracy requirements of malware classification analysis. Compared to other methods, the accuracy improvements while employing word2vec with the LSTM system are around 0.5 percent higher.

For FS, PSO has been used extensively. The location values were transformed into probabilities for the inclusion of features in the discrete or binary PSO that was first proposed for FS in [17] using a sigmoid function. Discrete and continuous PSO have occasionally been coupled to enhance the feature parameters and SVM, respectively [18]. Eesa *et al.* [19] created a hybrid model for the identification of network breaches by combining the Cuttlefish Optimization Method (CFA) and a Decision Tree (DT) classifier. The detection rate and accuracy are greatly increased when the number of features is less than 20 when utilizing the KDDCup99 dataset. The Artificial Bee Colony (ABC) technique for IDS FS was developed by researchers in [20]. The suggested approach makes use of a hybrid ABC-PSO methodology, a cutting-edge classification system.

Virtual domains in the cloud are protected from sneaky and obfuscated malware attacks by employing a technology known as VMShield [21]. VMShield can recognize over 97 percent of attacks using various introspection techniques, even camouflaged and obfuscated ones. Finally, VMshield obtains the model's output and delivers a status report that administrators can inspect. Overall, it is clear that there are a variety of techniques for memory-based malware detection that is obfuscated. This study provides a paradigm for memory analysis of obfuscated malware that focuses on an immediate and practical fix.

III. DATASET ENUMERATION

The most recent dataset for this investigation had 58,596 records, including 29,298 benign and 29,298 malicious records. It is called CIC-MalMem-2022. This dataset included three different types of malware—Spyware, Trojan-Horse, and Ransomware—along with 57 features, as given in Table I below.

TABLE I Enumeration of the employed dataset

No.	Name	No.	Name	No.	Name
1	“Category”	20	“handles.nsection”	39	“psxview.not_in_pslist_false_avg”
2	“pslist.nproc”	21	“handles.nmutant”	40	“psxview.not_in_eprocess_pool_false_avg”
3	‘pslist.nppid”	22	“ldrmodules.not_in_load”	41	“psxview.not_in_ethread_pool_false_avg”
4	“pslist.avg_threads”	23	“ldrmodules.not_in_init”	42	“psxview.not_in_pspcid_list_false_avg”
5	“pslist.nprocs64bit”	24	“ldrmodules.not_in_mem’	43	“psxview.not_in_csrss_handles_false_avg”
6	“pslist.avg_handlers	25	“ldrmodules.not_in_load_avg”	44	“psxview.not_in_session_false_avg”
7	“dlllist.ndlls”	26	“ldrmodules.not_in_init_avg”	45	“psxview.not_in_deskthrd_false_avg”
8	“dlllist.avg_dlls_per_proc”	27	“ldrmodules.not_in_mem_avg”	46	‘modules.nmodules”
9	“handles.nhandles”	28	“malfind.ninjections”	47	“svcscan.nservices”
10	“handles.avg_handles_per_proc”	29	“malfind.commitCharge”	48	“svcscan.kernel_drivers”
11	“handles.nport”	30	“malfind.protection”	49	“svcscan.fs_drivers”
12	“handles.nfile”	31	“malfind.uniqueInjections”	50	“svcscan.process_services”
13	“handles.nevent”	32	“psxview.not_in_pslist”	51	“svcscan.shared_process_services”
14	‘handles.ndesktop”	33	“psxview.not_in_eprocess_pool”	52	“svcscan.interactive_process_services”
15	‘handles.nkey’	34	“psxview.not_in_ethread_pool”	53	“svcscan.nactive”
16	“handles.nthread”	35	“psxview.not_in_pspcid_list”	54	“callbacks.ncallbacks”
17	“handles.ndirectory”	36	“psxview.not_in_csrss_handles”	55	“callbacks.nanonymous”
18	“handles.nsemaphore”	37	“psxview.not_in_session”	56	“callbacks.ngeneric”
19	“handles.ntimer”	38	“psxview.not_in_deskthrd”	57	“Class”

IV. PROPOSED METHODOLOGY

We outline our suggested methods for malware detection for a specific sample of records in this section. The research aims to identify malware. It shows the effectiveness of ML techniques in processing samples to differentiate between normal data and malware. The preprocessed, compact dataset is then subjected to wrapper-based, nature-inspired algorithms, including the BBA, CSA, MA, and PSO, algorithms for FS, as shown in Fig. 1; classification is carried out utilizing kNN, RF, and SVM for outcomes.

A. Data Pre-processing

To give a more acceptable and pertinent feature for the Python-oriented FS and classification algorithms, the dataset underwent several pre-processing steps. The actions to take are as follows:

1. *Features removal*: Two of the 57 features in the aforementioned dataset—Category and Class—are the class labels. It's essential to get rid of it as a result. Since the

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on February 06,2023 at 05:23:39 UTC from IEEE Xplore. Restrictions apply.

project's primary objective is to reduce the number of features, it is imperative to get rid of it.

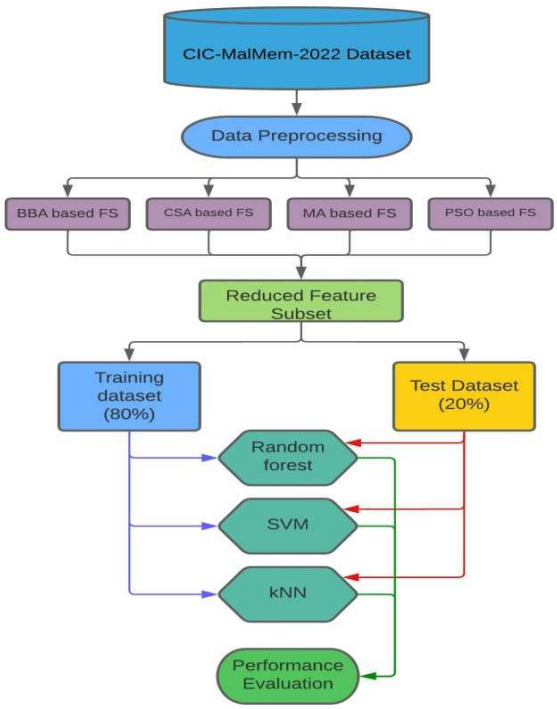


Fig. 1. The Proposed Malware Detection Model's Operational Flow

2. *Dataset Split:* After the FS process, there are 46876 instances set aside for training and 11720 instances set aside for testing, with the training section of the dataset accounting for 80% of the total and the testing section for 20%.

B. Feature Selection

The complete feature set of the dataset won't be useful for malware detection. The time-consuming procedure of adding undesired features for detection affects the system's accuracy. The BBA, CSA, MA, and PSO were used to develop the FS section. Every nature-inspired algorithm has 20 agents (population size) and 20 iterations. As the FS approach is wrapper based, the weight for classification accuracy is set to 0.75.

1. BBA

We must discretize the bat algorithm to handle the FS. The binary bat algorithm was proposed by Nakamura *et al.*[22] for FS and image processing. According to their study, the entire search space can be represented as a d-dimensional Boolean lattice in which bats traverse a hypercube's corners. The positions of the bats in BBA are constrained by the sigmoid function in accordance with (1) and (2).

$$S(v_i^j) = \frac{1}{1+e^{-v_i^j}} \tag{1}$$

$$x_i^j = \begin{cases} 1 & if \quad S(v_i^j) > \sigma \\ 0 & otherwise \end{cases} \tag{2}$$

The velocity v_i^j is the j-dimensional of the bat i^{th} and $\sigma \sim$ U (0, 1) is a uniformly distributed random value. To enhance the malware detection performance, we apply BBA in this paper to choose the optimal feature subset of features.

2. CSA

This method is based on the behavior of some cuckoo species, which are obligatory brood parasites. We now refer to our new Cuckoo Search[10] using the following three idealized principles for simplicity's sake: 1) Each cuckoo lays one egg at a time, which it deposits in a nest that is selected at random; 2) The best nests produce the highest-quality eggs, which are passed down to the following generations; 3) There are a fixed number of host nests accessible, and the host bird has a probability of $p_a \in [0, 1]$ of finding a cuckoo egg. In this situation, the host bird has two options: either toss the egg away or quit the nest and establish a brand-new one. For ease of calculation, the fraction p_a of the n nests that are replaced by new nests can be used to approximate the last premise (with new random solutions).

3. MA

A novel approach to resolving FS issues has been devised by Zervoudakis and Tsafarakis [23]and is known as MA. It is a hybrid approach that combines the benefits of traditional optimization techniques like PSO and Genetic algorithms [24]. It was demonstrated in the study of [25] that PSO required adjustments since it is prone to be stuck in a local optimum, particularly for problems with a high dimension. The algorithm's performance is improved across both small and big-scale feature sets thanks to the improvements made by the MA.

4. PSO

A meta-heuristic global optimization strategy that is a member of the swarm intelligence family of algorithms, particle swarm optimization was first developed by Kennedy and Eberhart [12]. The PSO idea involves updating each particle's velocity at each instant of time concerning its best position and the best position that its neighborhood has discovered [26]. "G_{best}" is the "global best," or the overall best solution for the swarm, and "P_{best}" is the particle's "personal best," or the coordinates of the best solution that particular particle has so far accomplished [27]. The stages of the PSO algorithm are as follows:

a. Set the location and velocity of each particle to random.
b. Evaluate each particle's fitness.
c. Update each particle's P_{best} and G_{best} values.
d. Update the velocity and position of each particle using (3) and (4) respectively.

$$V_p^{(t+1)} = w.V_p^{(t)} + c_1 * r_1 * \left(P_{best} - x_p^{(t)}\right) + c_2 * r_2 * \left(G_{best} - x_p^{(t)}\right) \tag{3}$$

$$x_p^{(t+1)} = x_p^{(t)} + V_p^{(t+1)} \tag{4}$$

e. Continue until the condition is satisfied.

C. Classification

An effective ML algorithm is essential when developing intelligent automation systems. The ML classifiers receive their input from the feature set produced by the nature-inspired algorithms for FS of the proposed models. The RF, kNN, and SVM classifiers are investigated in the current study.

1. RF

L. Breiman introduced an RF classifier [28]. It uses trees as the basis for an ensemble learning classifier [28]. The forecasts of several trees, each of which had been trained independently, were combined to produce it. Several earlier studies [29], [30] have found that RF classifier has a significant positive impact on IDS accuracy. The potential for the training set to overfit is corrected by RF. the RF algorithm is the supervised multi-classification technique.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on February 06,2023 at 05:23:39 UTC from IEEE Xplore. Restrictions apply.

2. SVM

Each data point is plotted in n-dimensional space using the supervised ML algorithm (SVM), where n is the total number of features. Each feature value corresponds to a specific coordinate's value. The classification process is then carried out by locating a hyper-plane that separates the data of one class from another [31]. To identify the hyper-plane, support vector training examples and the margins they establish are used. SVM falls within the topic of binary classifiers. It's a well-liked approach to classifying two groups. In this experiment, the "Radial Basis Function" (RBF) kernel is used. This kernel function is a good option because it works well in nonlinear forecasting and has fewer configurable parameters.

3. k-NN

k-NN is an example of a lazy ML algorithm that saves the training data before waiting for the test data [32]. When the test data instance is obtained, prediction is performed by searching the training data for the *k* examples or neighbors that are the most comparable. These students are in favor of gradual learning. These, however, are expensive to compute. To identify the training instance that is the closest to the specified test instance, we have chosen $k = 5$ and used normalized Euclidean distance.

V. PERFORMANCE MEASURE AND RESULTS

We use a confusion matrix to describe all potential outcomes of a prediction and to measure the classification performance. The below-mentioned metrics in Table II are widely used to assess models. The following are the performance metrics: True Positive (TP) represents a malware specimen that is accurately detected as malware. True Negative (TN) represents a normal specimen that is accurately classified as normal. False Positive (FP) means a normal specimen is misidentified as malware. False Negative (FN) represents a malware specimen that is wrongly labeled as normal.

TABLE II Metrics generated for performance evaluation

Metrics	Formula
Accuracy	(TP + TN) / (TP +FP+FN+TN)
Precision	TP / (TP + FP)
Recall	TP / (TP + FN)

TABLE IV Confusion matrices of kNN, RF and SVM

kNN			
BBA		**CSA**	
5857	3	5859	1
9	5851	11	5849
PSO		**MA**	
5857	3	5858	2
7	5853	10	5850

RF			
BBA		**CSA**	
5858	2	5859	1
0	5860	2	5858
PSO		**MA**	
5859	1	5859	1
1	5859	0	5860

SVM			
BBA		**CSA**	
5742	118	5720	140
30	5830	17	5843
PSO		**MA**	
5807	53	5719	141
52	5808	32	5828

TABLE III List of selected relevant features

Method	Selected Relevant features	Total
BBA	3,7,13,17,18,21,30,31,33,40,42,47,49,50,52,54	16
CSA	3,6,10,12,14,15,16,23,27,29,35,39,40,49,51,55	16
MA	1,4,6,8,10,12,16,18,19,21,28,33,36,37,42,46,52,53	18
PSO	9,17,19,21,24,29,35,43,51,53,54	11

The proposed methodology is utilized to choose the pertinent features, as indicated in Table III above. By limiting the number of features and choosing the pertinent ones, the proposed methodology met the study's goal. 16 pertinent features from BBA and CSA were selected. 18 and 11 pertinent features were chosen from MA and PSO, respectively. The confusion matrix for RF, kNN, and SVM outcomes is shown in Tables IV above, after applying the appropriate features produced by the BBA, CSA, MA, and PSO, techniques**.**

This chapter will detail the experimental findings from the proposed approach. The environment in which this research is conducted comprised Windows 10 pro-64-bit OS with Intel® Xeon® Gold 6130 CPU of 2.10 GHz and 64 GB of RAM. The python-based programming model is deployed on jupyter.

TABLE V Results of RF for malware detection

	Accuracy	Precision	Recall
BBA	0.9998293515358362	0.999829413	0.999829452
CSA	0.9997440273037542	0.999744042	0.999744027
MA	**0.999914675767918**	**0.99991469**	**0.999914676**
PSO	0.9998293515358362	0.999829352	0.999829352

TABLE VI Results of kNN for malware detection

	Accuracy	Precision	Recall
BBA	0.9989761092150171	0.998976632	0.998976109
CSA	0.9989761092150171	0.998977562	0.998976109
MA	0.9989761092150171	0.998977039	0.998976109
PSO	**0.9991467576791809**	**0.999914676**	**0.999146758**

TABLE VII Results of SVM for malware detection

	Accuracy	Precision	Recall
BBA	0.9873720136518771	0.987481947	0.987372014
CSA	0.9866040955631399	0.986818573	0.986604096
MA	0.9852389078498294	0.985406852	0.985238908
PSO	**0.9910409556313993**	**0.99104097**	**0.991040956**

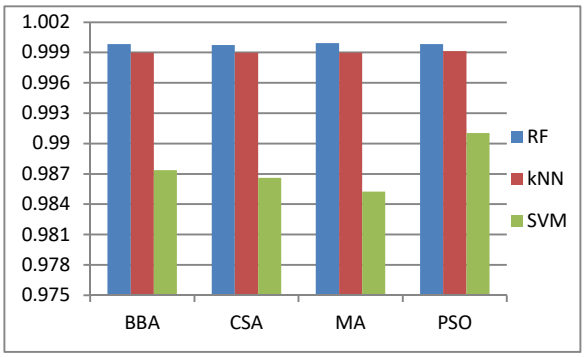


Fig. 2. Accuracy Comparison

As depicted in Tables V, VI, VII, and Fig. 2, RF-based classifiers with features selected by the MA approach outperformed other approaches for malware detection. The maximum accuracy attained is almost 1. In the case of kNN and SVM, features selected with PSO-based nature-based algorithm outperformed other approaches in terms of accuracy, precision and recall. It is clear from the results data that the RF algorithm with features selected from the MA nature-inspired approach outperforms and efficiently detects malware. It is found that there is a little difference between the outcomes but as the dataset size increases, this minor

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on February 06,2023 at 05:23:39 UTC from IEEE Xplore. Restrictions apply.

difference may have a huge impact on the outcome for detecting malware in the big network like cloud networks.

VI. CONCLUSION

Nature-inspired algorithms make it easier to locate the best feature sets. Nature-inspired algorithms can improve classification accuracy while decreasing the number of features used. To detect malware in a cloud environment, this improvement would consequently affect the degree of security. In cloud networks, the proposed paradigm is more effective in identifying malware threats. There were two sections to the experiment. The first stage is selecting features using techniques inspired by nature, then evaluating suggested models using RF, SVM, and kNN classifiers. The MA selected features used in the random forest technique yield good results and attains accuracy of 0.999914675767918. In the future, ML algorithms may incorporate more nature-inspired algorithms.

REFERENCES

[1] M. R. Watson, N. U. H. Shirazi, A. K. Marnerides, A. Mauthe, and D. Hutchison, "Malware Detection in Cloud Computing Infrastructures," *IEEE Trans. Dependable Secur. Comput.*, vol. 13, no. 2, pp. 192–205, 2016.

[2] A. Marnerides, C. James, A. Schaeffer-Filho, S. Sait, A. Mauthe, and H. Murthy, "Multi-level network resilience: Traffic analysis, anomaly detection and simulation," *ICTACT J. Commun. Technol. Spec. Issue Next Gener. Wirel. Networks Appl.*, vol. 2, no. 2, 2011.

[3] A.-T. Institute, "Malware," 2022. [Online]. Available: https://www.av-test.org/en/statistics/malware/. [Accessed: 25-May-2022].

[4] T. Carrier, P. Victor, A. Tekeoglu, and A. Lashkari, "Detecting Obfuscated Malware using Memory Feature Engineering," no. Icissp, pp. 177–188, 2022.

[5] K. Sethi, R. Kumar, L. Sethi, P. Bera, and P. K. Patra, "A novel machine learning based malware detection and classification framework," in *2019 International Conference on Cyber Security and Protection of Digital Services (Cyber Security)*, 2019, pp. 1–4.

[6] R. Kumar, K. Sethi, N. Prajapati, R. R. Rout, and P. Bera, "Machine learning based malware detection in cloud environment using clustering approach," in *2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, 2020, pp. 1–7.

[7] N. Scaife, H. Carter, P. Traynor, and K. R. B. Butler, "Cryptolock (and drop it): stopping ransomware attacks on user data," in *2016 IEEE 36th international conference on distributed computing systems (ICDCS)*, 2016, pp. 303–312.

[8] J. M. Vidal, A. L. S. Orozco, and L. J. G. Villalba, "Alert correlation framework for malware detection by anomaly-based packet payload analysis," *J. Netw. Comput. Appl.*, vol. 97, pp. 11–22, 2017.

[9] S. Mirjalili, S. M. Mirjalili, and X.-S. Yang, "Binary bat algorithm," *Neural Comput. Appl.*, vol. 25, no. 3, pp. 663–681, 2014.

[10] X.-S. Yang and S. Deb, "Cuckoo search via Lévy flights," in *2009 World congress on nature & biologically inspired computing (NaBIC)*, 2009, pp. 210–214.

[11] T. Bhattacharyya, B. Chatterjee, P. K. Singh, J. H. Yoon, Z. W. Geem, and R. Sarkar, "Mayfly in harmony: A new hybrid meta-heuristic feature selection algorithm," *IEEE Access*, vol. 8, pp. 195929–195945, 2020.

[12] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of ICNN'95 - International Conference on Neural Networks*, 1995, vol. 4, pp. 1942–1948 vol.4.

[13] T. Carrier, P. Victor, A. Tekeoglu, and A. H. Lashkari, "Detecting Obfuscated Malware using Memory Feature Engineering.," in *ICISSP*, 2022, pp. 177–188.

[14] I. Sadek, P. Chong, S. U. Rehman, Y. Elovici, and A. Binder, "Memory snapshot dataset of a compromised host with malware using obfuscation evasion techniques," *Data Br.*, vol. 26, p. 104437, Oct. 2019.

[15] N. Nissim, O. Lahav, A. Cohen, Y. Elovici, and L. Rokach, "Volatile memory analysis using the MinHash method for efficient and secured detection of malware in private cloud," *Comput. Secur.*, vol. 87, p. 101590, Nov. 2019.

[16] J. Kang, S. Jang, S. Li, Y. S. Jeong, and Y. Sung, "Long short-term memory-based Malware classification method for information security," *Comput. Electr. Eng.*, vol. 77, pp. 366–375, Jul. 2019.

[17] J. Kennedy and R. C. Eberhart, "A discrete binary version of the particle swarm algorithm," in *1997 IEEE International conference on systems, man, and cybernetics. Computational cybernetics and simulation*, 1997, vol. 5, pp. 4104–4108.

[18] C.-L. Huang and J.-F. Dun, "A distributed PSO–SVM hybrid system with feature selection and parameter optimization," *Appl. Soft Comput.*, vol. 8, no. 4, pp. 1381–1391, 2008.

[19] A. S. Eesa, Z. Orman, and A. M. A. Brifcani, "A novel feature-selection approach based on the cuttlefish optimization algorithm for intrusion detection systems," *Expert Syst. Appl.*, vol. 42, no. 5, pp. 2670–2679, 2015.

[20] W. A. H. M. Ghanem and A. Jantan, "Novel multi-objective artificial bee colony optimization for wrapper based feature selection in intruction detectoin," *Int. J. Adv. Soft Comput. its Appl.*, vol. 8, no. 1, pp. 70–81, 2016.

[21] P. Mishra, P. Aggarwal, A. Vidyarthi, P. Singh, B. Khan, H. H. Alhelou, and P. Siano, "VMShield: Memory introspection-based malware detection to secure cloud-based services against stealthy attacks," *IEEE Trans. Ind. Informatics*, vol. 17, no. 10, pp. 6754–6764, 2021.

[22] R. Y. M. Nakamura, L. A. M. Pereira, K. A. Costa, D. Rodrigues, J. P. Papa, and X.-S. Yang, "BBA: a binary bat algorithm for feature selection," in *2012 25th SIBGRAPI conference on graphics, patterns and images*, 2012, pp. 291–297.

[23] K. Zervoudakis and S. Tsafarakis, "A mayfly optimization algorithm," *Comput. Ind. Eng.*, vol. 145, p. 106559, 2020.

[24] J. J. Grefenstette, "Genetic algorithms and machine learning," in *Proceedings of the sixth annual conference on Computational learning theory*, 1993, pp. 3–4.

[25] M. Li, H. Chen, X. Shi, S. Liu, M. Zhang, and S. Lu, "A multi-information fusion 'triple variables with iteration' inertia weight PSO algorithm and its application," *Appl. Soft Comput.*, vol. 84, p. 105677, 2019.

[26] I. Sousa-Ferreira and D. Sousa, "A review of velocity-type PSO variants," *J. Algorithm. Comput. Technol.*, vol. 11, no. 1, pp. 23–30, 2017.

[27] F. Marini and B. Walczak, "Particle swarm optimization (PSO). A tutorial," *Chemom. Intell. Lab. Syst.*, vol. 149, pp. 153–165, 2015.

[28] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.

[29] Y. T. and R. M. P. Negandhi, "Intrusion detection system using random forest on the NSL-KDD dataset,"," in *Emerging Research in Computing, Information, Communication and Applications*, 2019, pp. 519–531.

[30] N. Farnaaz and M. A. Jabbar, "Random Forest Modeling for Network Intrusion Detection System," *Procedia Comput. Sci.*, vol. 89, pp. 213–217, 2016.

[31] H. Xue, Q. Yang, and S. Chen, "SVM: Support vector machines," in *The top ten algorithms in data mining*, Chapman and Hall/CRC, 2009, pp. 51–74.

[32] G. Guo, H. Wang, D. Bell, Y. Bi, and K. Greer, "KNN model-based approach in classification," in *OTM Confederated International Conferences" On the Move to Meaningful Internet Systems"*, 2003, pp. 986–996.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on February 06,2023 at 05:23:39 UTC from IEEE Xplore. Restrictions apply.

Environmental Science and Pollution Research
https://doi.org/10.1007/s11356-023-25896-8

Measuring circularity of a manufacturing organization by using sustainable balanced scorecard

Abhishek Sahu[1] · **Saurabh Agrawal**[2] · **Chandra Prakash Garg**[3]

Received: 30 September 2022 / Accepted: 8 February 2023
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2023

Abstract

In recent years, circular economy has become a matter of great importance because of its ability to contribute toward economic, environmental, and social aspects of the sustainability. The circular economy approaches help in resource conservation by reducing, reusing, and recycling products/parts/components/materials. On the other hand, Industry 4.0 is coupled with emerging technologies, which support the firms in efficient resource utilization. These innovative technologies can transform the present manufacturing organizations by reducing resource extraction, CO2 emissions, environmental damage, and power consumption and improve it into a more sustainable manufacturing organization. Industry 4.0 along with circular economy concepts greatly improves the circularity performance. However, there is no framework found for measuring the circularity performance of the firm. Therefore, the current study aims to develop a framework for measuring performance in terms of circularity percentage. In this work, graph theory and matrix approach are employed for measuring the performance based on a sustainable balanced scorecard such as internal process, learning and growth, customer and financial with environmental and social perspectives. A case of an Indian barrel manufacturing organization is discussed for the illustration of proposed methodology. Based on "circularity index" of the organization and the maximum possible circularity index, the circularity was found to be 5.10%. It indicates that there is a huge potential for the improvement in the circularity of the organization. An in-depth sensitivity analysis and comparison are also performed to validate the findings. There are very few studies on measuring the circularity. The study developed the approach for measuring circularity, which may be utilized by industrialists and practitioners for improving the circularity.

Keywords Balanced scorecard · Circular economy · Graph theory and matrix approach · Industry 4.0 · Sustainability

Introduction

In the present scenario, manufacturing organizations are struggling due to low availability of raw materials, excessive waste generation, high CO_2 emission, and lack of involvement of digital technology-based initiatives. Further, managing circularity in the supply chain has become more challenging with an increase in the number of returned items (Kumar et al. 2021). Due to sustainable development (economic, environmental, and social development), it is necessary to transform the linear business model into a circular formation, which can be represented as a circular economy (CE) (Fathollahi-Fard et al. 2022). CE is a modern approach, pointing to the sustainable usage of natural resources and eliminating waste. It emphasizes on maximizing the circularity of energy and resources within manufacturing processes. It works on the principle of a circular supply chain that emphasizes restorative and regenerative aspects (Saidani et al. 2019). CE-based

Responsible Editor: Philippe Garrigues

✉ Abhishek Sahu
er.abhisheksahu2012@gmail.com

Saurabh Agrawal
agrawals.iit@gmail.com

Chandra Prakash Garg
cpgarg86@gmail.com

[1] Department of Mechanical, Production and Industrial Engineering, Delhi Technological University, Delhi 110042, India

[2] Delhi School of Management, Delhi Technological University, Delhi 110042, India

[3] Department of Operations Management and Quantitative Techniques, Indian Institute of Management Rohtak, Rohtak 124010, India

Published online: 18 February 2023

approaches are intended to reduce dependency on fossil fuels, eliminate the use of toxic chemicals, carbon emissions, and water consumption, and suggest stakeholders in the planning of sustainable development. Kirchherr et al. (2017) suggested a 3R framework of CE, i.e., reduce, recycle, and reuse in the process of circularity. CE implementation proposes an industrial eco-system at meso-level, eco-design and green design (micro-level), and social-economical metabolism at macro-level (Ellen MacArthur Foundation 2013). All these levels improve the circularity of resources and natural systems that promote restorative design, continuous improvement, and zero waste culture.

On the other hand, the evolution of information and communication technology and data management underpins the principle of Industry 4.0 (I4.0), which creates a digital environment within an organization. I4.0 is coupled with emerging technologies, such as Internet of Things (IoT), Big Data, Cyber-Physical-Systems, Cloud Computing, and Cybersecurity (Sivarajah et al. 2017). These technologies have the ability to transform circularity levels in the various stages of supply chain. The main role of I4.0 implementation is to achieve a higher level of functional effectiveness and productivity. It may also provide the information regarding failure and errors, unusual wastage, and product usage patterns through real-time monitoring. Further, I4.0-based technologies increase operational efficiency, stock optimization, network monitoring, and reduction in the waste generation that may support the longer life of products and services.

Several research studies have already presented the benefits of implementing I4.0 from the CE perspective and measured circularity index. Vimal et al. (2021) estimated the circularity index by utilizing a multi-grade fuzzy technique in the manufacturing organization. Moreover, De Oliveira et al. (2021) also reviewed CE indicators for assisting the decision-making in circularity evaluations. Saidani et al. (2019) analyzed the fifty-five sets of circularity indicators to improve the CE performance. Furthermore, Bocken et al. (2017) took circularity to the highest level by introducing the product and material circularity, which encourages a circular business model. Also, several number of studies (de Sousa Jabbour et al. 2018; Kristoffersen et al. 2020; Bag et al. 2021) have utilized various multi criteria decision-making techniques, such as Analytic Network Process (ANP), Analytic Hierarchy Process (AHP), Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS), Elimination and Choice Translating Reality (ELECTRE), Decision Making Trial and Evaluation Laboratory (DEMATEL), and Preference Ranking Organization Method for Enrichment Evaluation (PROMETHEE). AHP and TOPSIS are applied when attributes are relative and independent, while ANP does not develop a hierarchical structure among attributes (Rao 2007). However, such limitations are not present in the graph theory and matrix approach (GTMA). Therefore, GTMA is adopted for the proposed research because no one has applied this approach for measuring the circularity. It is also found that most of the studies provide the comparative value of circularity. The proposed study measures the circularity in terms of percentage, by using the value of circularity index and maximum value of the circularity index. Further, in-depth sensitivity analysis and comparison are also performed to validate the findings. The study utilizes a sustainable balanced scorecard (SBSC) for measuring the circularity index using permanent function value. A case of ABC barrel manufacturing organization is considered to demonstrate the current research methodology. The research objectives of the study are as follows:

- To identify and analyze attributes and sub-attributes for the implementation of CE and I4.0.
- To develop GTMA-based approach for evaluating the circularity in terms of percentage.

"The main purpose of this article is not only to improve the circularity and triple bottom line perspective of sustainability but also to integrate I4.0 technologies with SBSC in order to have a comprehensive understanding of circularity decisions. This integration may also improve circularity level in the various stages of supply chain and reduce the total power/energy consumption, greenhouse gas emissions, and wasteful actions, and also prevents environmental degradation caused by resource extraction. Further, adoption and implementation of I4.0 and CE in the context of sustainable development is the primary stage of development, and very few studies are available on those considered GTMA using SBSC. To reduce material wastage, energy consumption, and environmental degradation, the circularity of the resources is an important step toward sustainable development, which provides low capital investment, less uncertainty, more focus on core competency, and better utilization of modern technology. This significant contribution makes the study unique, which is relevant from the viewpoint of a manufacturing organization and practitioners who are trying to produce high-end products at the lowest possible cost while simultaneously monitoring the product's end-of-life in real-time."

The remaining part of the article is structured as follows. Section 2 consists of a literature review related to CE and I4.0. Section 3 comprises the research methodology of GTMA, including step-by-step procedures discussed for calculating the "Circularity index." A case elucidation of the Indian barrel manufacturing organization is presented in Sect. 4. Finally, Sect. 5 considers results and discussion along with research implications. Finally, Sect. 6

summarizes conclusion along with the limitations and future research directions of the study.

Literature review

To accomplish the research objectives, the study performed a comprehensive literature review to investigate the principles of CE and I4.0. The detailed literature review is explained in the following sub-section.

Principles of circular economy

In the late 1990s, China implemented CE to supersede the end-of-life criteria with restoration. CE works on the principles of energy conservation, circularity of resources, and waste management (De Oliveira et al. 2021). It seeks to maintain materials, components, and products at their highest utility and increases its value under all circumstances (Bag et al. 2021). The CE approaches include reuse, recycle, reduce, refurbish, remanufacturing, recover, repair, and refuse (Kirchherr et al. 2017; Wang et al. 2021). These approaches may improve the circularity performance of the circular supply chain. The key objective of CE is to reduce the design of waste, eliminate the use of toxic substances, and work on optimum design life criteria. Circularity is the key strategy for CE implementation. Circularity is represented in the form of products, waste materials, recycled components, and customer dissatisfaction. Ellen MacArthur Foundation (2013) proposed the hierarchy structure of the circularity strategy using a butterfly figure and suggested a circular flow of materials, waste reduction, and resource optimization along with ecological sustainability. These strategies may have the potential to improve CE performance through a circular business model. Soleimani et al. (2022) have utilized the lagrangian relaxation, reformulations, and heuristics, which represent the redesigning strategies of the CE to achieve a sustainable closed-loop supply chain by minimization of waste, energy consumption, and maintenance of ecological balance.

Principles of Industry 4.0

The German Industry Science Research Alliance first proposed the principle of Fourth Industrial Revolution, i.e., I4.0 in 2011. I4.0 integrates embedded systems, manufacturing technologies, and smart production processes to pave the way to a modern technological era, which may profoundly transform organizations and value chain partners along with business models (Kristoffersen et al. 2020; Virmani et al. 2021). Further, it improves the data exchange, real-time monitoring, optimization capability, and failure detection, which reduces the breakdowns of machines and minimizes equipment downtime. This improves system performance besides revealing new product, service, and business model possibilities. I4.0 encompasses numerous technological notions, such as interoperability among varied systems, semantics between machines, decentralization, cyber-physical systems, virtualization, computerization, and others (Bag et al. 2021; Sun et al. 2021). The IoT technology enables machines to interconnect with several other web-enabled machines, enabling collection, processing, and exchange of huge amounts of information on a real-time basis (Chong et al. 2022; Singh et al. 2022). Furthermore, Big Data technology stores and processes a large amount of data and provides real-time data support for comprehensive data analysis. The cloud computing technologies share information based on internet-connected systems and sub-systems (Ali et al. 2022). This provides cost-effecting solutions for sharing information and improves the performance of the system. Cybersecurity permits safety and privacy from the unauthorized servers and cyber-attacks (Sivarajah et al. 2017; Kamble et al. 2018).

Besides this, the implementation of I4.0 in CE is the primary stage of development. The major role of I4.0 implementation is to enhance the circularity of products through real-time status, which enhances visibility and risk mitigation throughout the product life cycle. Several studies are working toward the CE implementation using I4.0-based technologies. Bag et al. (2021) found 10 redesign strategies of the CE in the implementation of I4.0. Dantas et al. (2021) integrated CE and I4.0 and identified the 17 most significant factors in achieving sustainable development goals. Rajput and Singh (2019) utilized the DEMATEL approach for the principal component analysis of I4.0 and CE analyzing key challenging and enabling factors. The current status of I4.0 and CE is summarized in the following Table 1. It is found from the Table 1 that most of the studies considered the qualitative as well as quantitative analysis and conceptual framework in the integration of I4.0 and CE. However, a very limited study is considered for the integration of I4.0 technologies from the CE perspective in a specific sector. Therefore, the article provides a case study of a manufacturing organization focusing on I4.0 and CE integration. Furthermore, most of the work is carried out in developed economies, while very few works are performed in developing economies like India. Last but not least, the combination of GTMA using SBSC is found in none of the studies from the perspective of I4.0 and CE.

Sustainable balanced scorecard

In the present scenario, developing environmental concerns, corporate social responsibility, and laws in several emerging economies have forced organizations to look beyond the

Table 1 Current status of I4.0 and CE

Authors	Country	Approach	Sector	Findings
Tang et al. (2022)	Hong Kong	PLS-SEM	Manufacturing	Identified the role of blockchain technology on organization performance
Chau et al. (2021)	Vietnam	Qualitative	Remanufacturing industry	Discussed the impact of IoT in remanufacturing and identified key challenges and barriers of the I4.0, remanufacturing and CE
Laskurain-Iturbe et al. (2021)	Spain	Case study	Aerospace	Identified the major technologies of I4.0 in the CE implementation
Abdul-Hamid et al. (2020)	Taiwan	Fuzzy-Delphi	Palm oil	Identified the critical challenges of I4.0 and CE in the palm oil industry
Kristoffersen et al. (2020)	Norway	Qualitative	General	Established a digital-enabled circular strategies framework for I4.0 and CE
Yadav et al. (2020)	India	BWM-ELECTRE	Automotive	Found 28 challenges and 22 solution measures of I4.0 and CE implementation
Rajput and Singh et al. (2019)	India	DEMATEL	General	Measured 26 enabling factors and 15 barriers to study the cause-effect relationship between I4.0 and CE
Lopes de Sousa Jabbour et al. (2018)	France	ReSOLVE model	General	Developed a roadmap to increase the application of I4.0 from the CE perspective

traditional methods. Epstein and Wisner (2001) utilized an SBSC for developing strategies for sustainability and identified the various factors that improved the performance of sustainability. Agrawal et al. (2016a) also proposed an SBSC framework to select the outsourcing index of reverse logistics. Social and environmental responsibility is seen as a center of the organization's strategy that produces competitive advantage rather than being considered a strategy to facilitate operational efficiency. The present study employed SBSC based on experts' opinions and existing literature review. SBSC approach encompasses the internal process, learning and growth, customer, financial as well as environmental, and social perspectives. The proposed SBSC attributes and sub-attributes are presented in Table 2.

Research methodology

GTMA is a rigorous and rational decision-making technique, i.e., valuable in investigating directional graphs. Further, it works most accurately when nodes are larger in number. In many areas of science and technology, graph/digraph model illustrations have been beneficial for modeling and analyzing many types of systems and challenges (Rao 2007). When a graph or digraph is complicated, visual analysis becomes challenging. There are several computer programs or languages available to solve and analyze matrices. Existing studies used advanced graph theory for analyzing various systems and sub-systems for evaluating real-life challenges. Rao (2007) established the GTMA approach for the selection of material and evaluating material suitability index. Agrawal et al. (2016b) identified disposition index of reverse logistics and prioritized five alternatives using GTMA. Additionally, Virmani et al. (2021) utilized GTMA to identify major barriers and a comprehensive I4.0 index in the automobile industry. Consequently, it is found that most of the studies implemented GTMA for disposition decisions, contractor ranking, material selection, and I4.0 ranking.

In addition, nodes and edges are two essential elements of GTMA that express inter-relationship among various nodes (also various attributes) through a directed graph (also termed as digraph) and a matrix. The attributes are represented by these nodes. A digraph depicting these attributes and their interaction is developed with the selection of attributes. The number of nodes in the digraph corresponds to the number of attributes examined, and with the help of a directed edge or arrow linking the two nodes that indicate their relative importance. The relevance of individual attributes to the organization or inheritance is shown by diagonal elements. The digraph or matrix may transform when the number of attributes changes. The common version of matrix function is termed a permanent function, which is estimated instead of the matrix's determinant. Because of the existence of minus signs in determinant function of a matrix, several critical parameters could be lost. Therefore, researchers selected to use a matrix's permanent function to deliver comprehensive information without compromise. [A] represents a square matrix of $M \times M$ size for M number of attributes taking all the attributes inheritance (C_i, $i = j$) along with their relative importance (a_{ij}, $i \neq j$).

Table 2 The attributes and sub-attributes for the considered study

Attributes/Sub-attributes	Explanation	References
Internal Process Perspective (IPP)		
Recycle (RCY)	Recycling is a method of treating resources to generate a product of the same or lower quality	Bocken et al. (2017); Saidani et al. (2019)
Reuse (RUS)	Reuse the things after repairing or recovering the material	Kirchherr et al. (2017); Slaveykova et al. (2019)
Refurbished (RFB)	The term "refurbish" mentions to the process of updating an older product	Chau et al. (2021); De Oliveira et al. (2021)
Reduce (RDC)	Reducing means the lesser amount of natural resources and raw materials used in the manufacturing process	Rajput and Singh (2019)
Learning and Growth Perspective (LGP)		
Technological Roadmap (TR)	I4.0-based technologies provide real-time status of the product and its usage pattern. It also provided information on product quality, quantity, and remaining life of the product	Lopes de Sousa Jabbour et al. (2018); Garg et al. (2021); Tang et al. (2022)
Big Data Analytics (BDA)	It stores, processes, and converts the unstructured data into the structured form, which may be helpful in circular supply chain implementation	Sivarajah et al. (2017); Chau et al. (2021); Singh et al. (2021)
Cybersecurity (CBS)	It is a technique of protecting critical and sensitive data from unwanted threats known as cybersecurity	Virmani et al. (2021); Bag et al. (2021); Fallahpour et al. (2021)
Seamless Integration (SI)	It is essential for developing industrial system integration to exchange information and data sharing to confirm the minimum utilization of available resources and costs for achieving organizational sustainability	Rajput and Singh (2019); Kristoffersen et al. (2020)
Customer Perspective (CP)		
Return Product Quality (RPQ)	Consumer behavior or dissatisfaction with the quality of returned and recycled products is the major issue in the success of I4.0 and CE	Epstein and Wisner (2001); de Sousa Jabbour et al. (2018)
Management Support (MS)	Management support and coordination with supply chain partners play a significant role in implementing I4.0 and CE	Agrawal et al. (2016a); Chauhan et al. (2022); Virmani et al. (2021); Bangwal et al. (2022)
Strategic Planning (SP)	Smart technologies enable the circular strategy for long-term sustainability goals	Kristoffersen et al. (2020)
Decision-Making Strategy (DMS)	It provides strategic planning and decisions for future demands and material availability	Chau et al. (2021); Abdul-Hamid et al. (2020)
Financial Perspective (FP)		
Infrastructure Cost (IC)	It includes the cost of software and hardware, including sensors technology, storing devices, integrated database server technology, and recycling equipment	Prakash et al. (2015); Garg et al. (2021) Laskurain-Iturbe et al. (2021)
Return on Investment (ROI)	It includes profit sharing and revenue generation	Chauhan et al. (2022)
Redesigning Cost (RC)	High redesigning or recycling is the major challenge in employing I4.0 and CE	Saidani et al. (2019)
High Implementation Cost (HIC)	High software and hardware costs of developing I4.0 infrastructure and recycling/remanufacturing/inspection techniques are the major challenges	Kristoffersen et al (2020); Kashav et al. (2021)

Table 2 (continued)

Attributes/Sub-attributes	Explanation	References
Environmental Perspective (EP)		
Waste Reduction (WR)	It impacts on ecological performance and reduces the adverse effect on environment by resource extraction	de Sousa Jabbour et al. (2018); Chau et al. (2021)
Reduce CO_2 Emission (RCE)	Reducing CO_2 emissions, carbon footprints, and heat emissions is the major attribute of CE implementation	Chauhan et al. (2022); Awan and Sroufe (2022)
Reduce Energy Consumption (REC)	The organization can reduce power consumption, water consumption, CO2 emissions, and energy requirement by implementing solar energy	Jabbour et al. (2018); Bag et al. (2021)
Zero Wastage (ZW)	Reduce solid waste, energy consumption, and raw material and available resources	Bocken et al. (2017); Chau et al. (2021)
Social Perspective (SP)		
Employee Satisfaction (ES)	It provides social responsibility, workload distribution, health standards, and working environment conditions	Epstein and Wisner (2001); Tian et al. (2022)
Corporate Social Responsibility (CSR)	Establish CSR in the organization for employee welfare, health, education, and supporting social projects	Awan and Sroufe (2022); Agrawal et al. (2016b)
Employment Training Programs (ETP)	Promote employment skill development programs, carrier opportunities, and legal and ethical	Gómez et al. (2018); Vimal et al. (2021)
Increase Worker Safety (IWS)	Provide safety equipment, hand gloves, and lost time injury rate for workers' health and safety	

$$[A] = \begin{bmatrix} C_1 & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} & \cdots & \cdots & a_{1M} \\ a_{21} & C_2 & a_{23} & a_{24} & a_{25} & a_{26} & \cdots & \cdots & a_{210} \\ a_{31} & a_{32} & C_3 & a_{34} & a_{35} & a_{36} & \cdots & \cdots & a_{310} \\ a_{41} & a_{42} & a_{43} & C_4 & a_{45} & a_{46} & \cdots & \cdots & a_{410} \\ a_{51} & a_{52} & a_{53} & a_{54} & C_5 & a_{56} & \cdots & \cdots & a_{510} \\ a_{61} & a_{62} & a_{63} & a_{64} & a_{65} & C_6 & \cdots & \cdots & a_{610} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{M1} & a_{M2} & a_{M3} & a_{M4} & a_{M5} & a_{M6} & \cdots & \cdots & C_{MM} \end{bmatrix}$$

The permanent function of matrix [A] is as follows:

$\prod_{i=1}^{M} C_i \sum_{i=1}^{M-1} \sum_{j=i+1}^{M} \cdots \cdots \sum_{M=t+1}^{M} (a_{ij}a_{ji}$) $C_k C_l C_m C_n C_o C_t C_M$ . . . M $\neq$ pus + $\sum_{i=1}^{M-2} \sum_{j=i+1}^{M-1} \sum_{k=j+1}^{M} \cdots \cdot \sum_{M=t+1}^{M} (a_{ij}a_{jk}a_{ki} + a_{ik}a_{kj}a_{ji})$ $C_1C_mC_nC_o...C_tC_M$ K...M $\neq$ pus.

$\prod_{i=1}^{M} C_i + \sum_{i=1}^{M-1} \sum_{j=i+1}^{M} \cdots \cdots \sum_{M=t+1}^{M} (a_{ij}a_{ji}$) $C_kC_lC_mC_nC_o...C_tC_mK...M \neq pus +$

$\sum_{i=1}^{M-2} \sum_{j=i+1}^{M-1} \sum_{k=j+1}^{M} \cdots \cdots \sum_{M=t+1}^{M} (a_{ij}a_{ji} + a_{ik}a_{kj}a_{ji})$ $C_1C_mC_nC_o...C_tC_mK, l...M \neq pus +$

$\sum_{i=1}^{M-3} \sum_{j=i+1}^{M} \sum_{k=i+2}^{M-1} \sum_{l=i+2}^{M} \cdots \cdots \sum_{M=t+1}^{M} (a_{ij}a_{ji})\ (a_{kl}a_{lk})$ $C_mC_nC_o...C_tC_M$ k,l,m...M $\neq$ pus +

$\sum_{i=1}^{M-3} \sum_{j=i+1}^{M+1} \sum_{k=i+1}^{M} \sum_{l=i+1}^{M} \cdots \cdots \sum_{M=t+1}^{M} (a_{ij}$ $a_{jk}a_{kl}a_{li} + a_{il}a_{lk}a_{kj}a_{ji}) \times C_mC_nC_o...C_tC_M$ k,l,m...M $\neq$ pus + $\sum_{i=1}^{M-2} \sum_{j=i+1}^{M-1} \sum_{k=j+1}^{M} \sum_{l=1}^{M-1} \sum_{m=l+1}^{M} \cdots \cdots \sum_{M=t+1}^{M} (a_{ij}$ $a_{jk}a_{ki} + a_{ik}a_{kj}a_{ji})\ (a_{lm}a_{ml}) \times C_nC_o...C_tC_M$ k,l,m...M $\neq$ pus + $\sum_{i=1}^{M-4} \sum_{j=i+1}^{M-1} \sum_{k=i+1}^{M} \sum_{l=i+1}^{M} \sum_{m=j+1}^{M} \cdots \cdots \sum_{M=t+1}^{M} (a_{ij}$ $a_{jk}a_{kl}a_{lm}a_{mi} + a_{im}a_{ml}a_{lk}a_{kj}a_{ji}) \times C_nC_o...C_tC_M$ k,l,m...M $\neq$ pus + $\sum_{i=1}^{M-5} \sum_{j=i+1}^{M-1} \sum_{k=j+1}^{M} \sum_{l=1}^{M-2} \sum_{m=l+1}^{M-1} \sum_{n=m+1}^{M} \cdots \cdots \sum_{M=t+1}^{M} (a_{ij}$ $a_{jk}a_{ki} + a_{ik}a_{kj}a_{ji}) \times (a_{lm}a_{mn}a_{ln} + a_{lm}a_{nm}a_{lm})\ C_o...C_tC_M$ k,l,m...M $\neq$ pus (1).

Generally, the value of a permanent function of a matrix [A] along with attributes a_{ij} is represented by:

$$per(A) = \sum_P \prod_{i=1}^{N} a_i, P(i). \quad (2).$$

The value of permanent functions for each alternative will be evaluated by using a C + + program. The logical flow diagram of this program and detailed GTMA procedure are provided by Agrawal et al. (2016b). A research framework for the proposed study has been established using the GTMA, which is presented in Fig. 1. The stepwise procedure of GTMA is as follows:

Step 1: To identify and analyze several factors of circularity alternatives.

Step 2: Select the attributes and sub-attributes of CE and I4.0 using SBSC.

Step 3: Draw an attributes digraph for the considered strategy and transform the digraph into a matrix. Figure 2 shows the digraph of inter-relationship among the various circularity attributes.

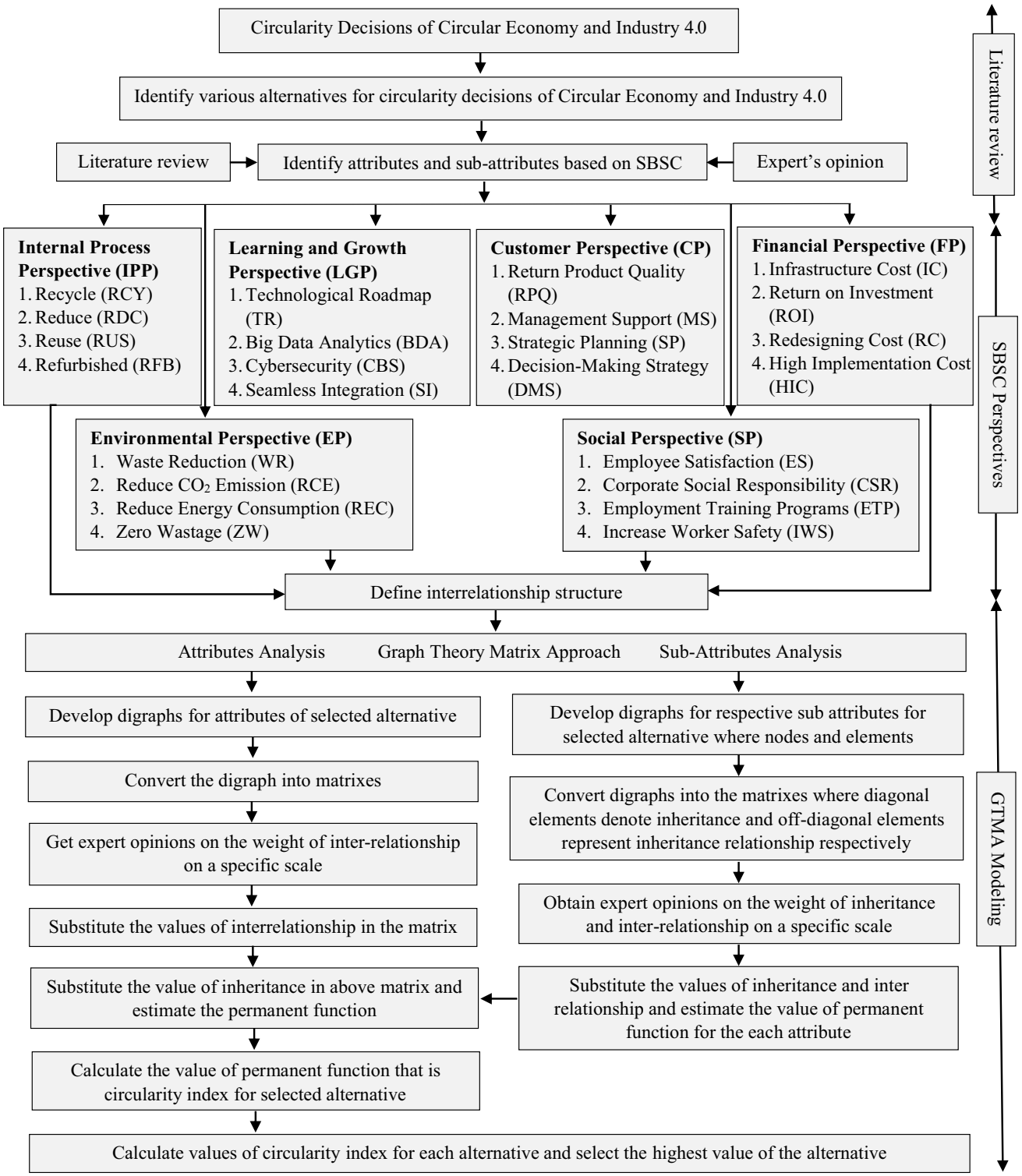


Fig. 1 Framework for circularity decisions of CE and I4.0

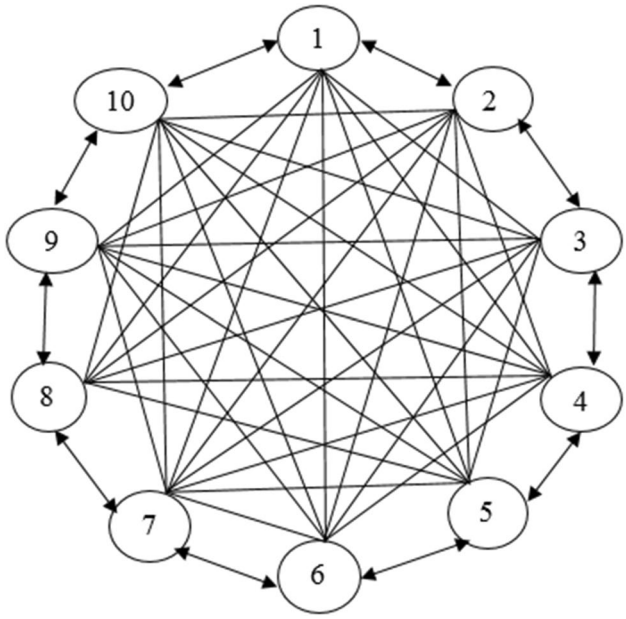


Fig. 2 Digraph of inter-relationship among circularity attributes

Table 3 Values of relative importance

Class description	Relative importance	
	a_{ij}	$a_{ji} = 1\text{-}a_{ij}$
Two attributes are equally significant	0.5	0.5
One attribute is slightly more significant than the other	0.6	0.4
One attribute is strongly more significant than the other	0.7	0.3
One attribute is very strongly significant than the other	0.8	0.2
One attribute is extremely significant than the other	0.9	0.1
One attribute is exceptionally more significant than the other	1	0

Step 4: Using a suitable scale, take opinions from experts to evaluate the weights of inter-relationship and relative importance between sub-attributes. Table 3 shows the scale for the proposed research. It is used to estimate the relative importance of several attributes.

Step 5: The values of diagonal elements of the matrix can be derived as follows.

i. Select each of the attributes individually. The proposed research contains six attributes, each of which has four sub-attributes. Choose the first attribute and its corresponding sub-attribute.
ii. Draw a sub-attributes digraph and convert it to a matrix with off-diagonal elements representing the inter-relationship between them.

Table 4 Inheritance values of attributes for each circularity alternative

Subjective measure of the attributes	Assigned value of the attributes (C_{ij})
Extremely low	0.1
Very low	0.2
Low	0.3
Below average	0.4
Average	0.5
Above average	0.6
High	0.7
Very high	0.8
Extremely high	0.9
Exceptionally high	1.0

iii. Obtain expert opinions on the weights of inheritance and inter-relationships on a preferred scale. The inter-relationship and inheritance scales are presented in Tables 3 and 4, respectively, employed in the proposed work.
iv. After substituting the values of interrelationship and inheritance into the matrix, use function per(A) to estimate the permanent function for the first attribute. In the proposed work, a C++ program is used to evaluate the first attribute's permanent function.
v. Repeat the procedure from step 1 to step 4 to evaluate the value of the permanent function for another five attributes.

Step 6: Substitute inheritance values into the matrix and use a C++ program or function per(A) to estimate the permanent function of the considered strategy.

Step 7: The permanent function is computed to determine circularity value for strategy.

Step 8: Select the optimum strategy of I4.0 and CE using the values of circularity index.

Case illustration

Indian barrel manufacturing organization ABC is taken for the case illustration. The annual turnover of the organization is around 161 billion INR with a 2.3 billion annual profit. The organization is the market leader in barrel manufacturing with the low-cost market segment. Further, organization has developed lean culture and always worked on sustainable development and emphasizes circularity of the resources. It has recently started selling steel barrels in Dubai also and is planning to implement CE and I4.0 in their manufacturing systems; therefore, circularity of resources within the organization will be maximized. To apply the proposed

methodology, suitable circularity alternatives and sub-alternatives linked with MS barrel for ABC organization are identified.

These alternatives are based on six standard SBSC. To provide diagonal and off-diagonal elements values, researchers enlisted using four experts. The experts from the organization, manufacturing industry, and third-party service providers are selected for analysis. These experts have substantial experience in industrial, production, and marketing fields.

Results and discussion

The implementation of the proposed methodology for barrel manufacturing organization is described in this section. The expert's opinion is selected to identify attributes, sub-attributes, and their inter-relationships as mentioned in steps 1 and 2 of the research methodology. As per literature reviews and discussions with experts, the strategy-based decisions were chosen using SBSC for the study. The digraph for CE and I4.0 is based on the GTMA, which includes six nodes pertaining to six attributes from the SBSC perspectives that can produce a direct impact on the circularity decision. Figure 3 shows the digraphs for all six attributes. The directed edges represent inter-relationships among attributes. Further, these digraphs are then converted to matrices. Matrix [CS] is used to illustrate the matrix. The diagonal element values and inheritance are calculated using the permanent function of the sub-attributes matrices. In the subsequent step, diagonal element values are evaluated for the matrix [CS]. The sub-attributes digraph will have four nodes, and each of the sub-attributes has a corresponding attribute. Figure 4 illustrates the digraphs for corresponding sub-attributes for each of the six SBSC perspectives. The directed edges represent inter-relationships among sub-attributes. These digraphs are then transformed into the matrices. The values of off-diagonal elements are estimated based on the expert's opinion and derived from Table 3. Table 5 contains the values of the diagonal elements for strategy. The values in Table 5 are according to the expert opinions chosen from Table 4. The next step is to calculate the values of permanent functions for strategy. Permanent function value for strategy can be calculated (0.6, 0.8, 0.7, 0.9) by putting the values of the diagonal elements in the Table 5. Permanent function value is calculated by putting these values in the matrix [IPP] and then using the matrix in a C++program. C11 in the matrix [CS] is used to represent this value. For strategy, the values of inheritance or diagonal elements have been determined.

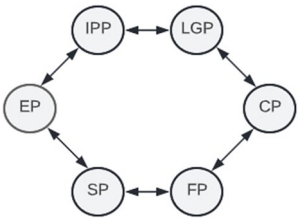


Fig. 3 Digraph for the circularity decisions of CE and I4.0

Circularity matrix for strategy [CS].

$$= \begin{bmatrix} - & \text{IPP} & \text{LGP} & \text{CP} & \text{FP} & \text{EP} & \text{SP} \\ \text{IPP} & C_{11} & 0.7 & 0.4 & 0.6 & 0.4 & 0.5 \\ \text{LGP} & 0.3 & C_{12} & 0.4 & 0.6 & 0.5 & 0.6 \\ \text{CP} & 0.6 & 0.6 & C_{13} & 0.5 & 0.6 & 0.4 \\ \text{FP} & 0.4 & 0.4 & 0.5 & C_{14} & 0.7 & 0.5 \\ \text{EP} & 0.6 & 0.5 & 0.4 & 0.3 & C_{15} & 0.6 \\ \text{SP} & 0.5 & 0.4 & 0.6 & 0.5 & 0.4 & C_{16} \end{bmatrix}$$

(i) Matrix from an internal process perspective (IPP) and value found for strategy is C11 = 2.5056.

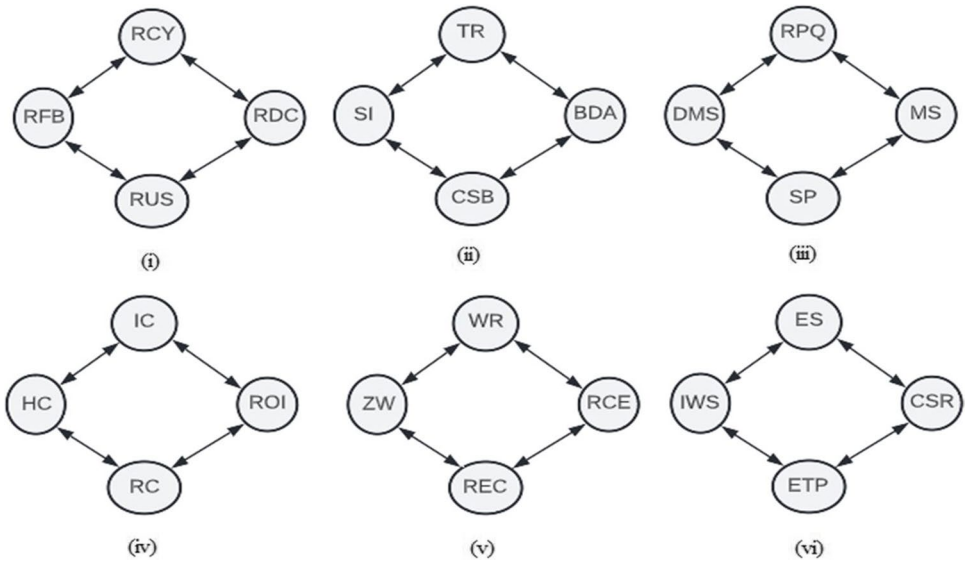


Fig. 4 Sub-attributes digraph from SBSC perspectives, i.e., (i) IPP, (ii) LGP, (iii) CP, (iv) FP, (v) EP, (vi) SP

Table 5 The diagonal elements values for the corresponding sub-attributes

Attributes/Sub-attributes	Strategy	Attributes/Sub-attributes	Strategy
Internal Process Perspective (IPP)		Financial Perspective (FP)	
Recycle (RCY)	0.6	Infrastructure Cost (IC)	0.8
Reuse (RUS)	0.8	Return on Investment (ROI)	0.7
Refurbished (RFB)	0.7	Redesigning Cost (RC)	0.6
Reduce (RDC)	0.9	High Implementation (HIC)	0.8
Learning and Growth Perspective (LGP)		Environmental Perspective (EP)	
Technological Roadmap (TR)	0.9	Waste Reduction (WR)	0.8
Big Data Analytics (BDA)	0.8	Reduce CO_2 Emission (RCE)	0.6
Cybersecurity (CBS)	0.8	Reduce Energy Consumption (REC)	0.7
Seamless Integration (SI)	0.6	Zero Wastage (ZW)	0.8
Customer Perspective (CP)		Social Perspective (SP)	
Return Product Quality (RPQ)	0.6	Employee Satisfaction (ES)	0.5
Management Support (MS)	0.7	Corporate Social Responsibility (CSR)	0.7
Strategic Planning (SP)	0.5	Employment Training Programs (ETP)	0.4
Decision-Making Strategy (DMS)	0.4	Increase Worker Safety (IWS)	0.6

$$= \begin{bmatrix} - & \text{RCY} & \text{RDC} & \text{RUS} & \text{RFB} \\ \text{RCY} & \ldots & 0.4 & 0.7 & 0.4 \\ \text{RDC} & 0.6 & \ldots & 0.4 & 0.4 \\ \text{RUS} & 0.3 & 0.6 & \ldots & 0.7 \\ \text{RFB} & 0.6 & 0.6 & 0.3 & \ldots \end{bmatrix}$$

(ii) Matrix from a learning and growth perspective (LGP) and value found for strategy is C12 = 2.3779.

$$= \begin{bmatrix} - & \text{TR} & \text{BDA} & \text{CBS} & \text{SI} \\ \text{TR} & \ldots & 0.5 & 0.6 & 0.4 \\ \text{BDA} & 0.5 & \ldots & 0.7 & 0.6 \\ \text{CBS} & 0.4 & 0.3 & \ldots & 0.3 \\ \text{SI} & 0.6 & 0.4 & 0.7 & \ldots \end{bmatrix}$$

(iii) Matrix from a customer perspective (CP) and value found for strategy is C13 = 1.5272.

$$= \begin{bmatrix} - & \text{RPQ} & \text{MS} & \text{SP} & \text{DMS} \\ \text{RPQ} & \ldots & 0.6 & 0.5 & 0.5 \\ \text{MS} & 0.4 & \ldots & 0.6 & 0.7 \\ \text{SP} & 0.5 & 0.4 & \ldots & 0.8 \\ \text{DMS} & 0.5 & 0.3 & 0.2 & \ldots \end{bmatrix}$$

(iv) Matrix from a financial perspective (FP) and value found for strategy is C14 = 1.7162.

$$= \begin{bmatrix} - & \text{IC} & \text{ROI} & \text{RC} & \text{HIC} \\ \text{IC} & \ldots & 0.8 & 0.7 & 0.8 \\ \text{ROI} & 0.2 & \ldots & 0.7 & 0.8 \\ \text{RC} & 0.3 & 0.3 & \ldots & 0.7 \\ \text{HIC} & 0.2 & 0.2 & 0.3 & \ldots \end{bmatrix}$$

(v) Matrix from an environment perspective is (EP) and value found for strategy is C15 = 2.1813.

$$= \begin{bmatrix} - & \text{WR} & \text{RCE} & \text{REC} & \text{ZW} \\ \text{WR} & \ldots & 0.5 & 0.7 & 0.8 \\ \text{RCE} & 0.5 & \ldots & 0.6 & 0.3 \\ \text{REC} & 0.3 & 0.4 & \ldots & 0.4 \\ \text{ZW} & 0.2 & 0.7 & 0.6 & \ldots \end{bmatrix}$$

(vi) Matrix from a social perspective (SP) and value found for strategy is C16 = 1.3115.

$$= \begin{bmatrix} - & \text{ES} & \text{CSR} & \text{ETP} & \text{IWS} \\ \text{ES} & \ldots & 0.8 & 0.7 & 0.8 \\ \text{CSR} & 0.2 & \ldots & 0.6 & 0.5 \\ \text{ETP} & 0.3 & 0.4 & \ldots & 0.4 \\ \text{IWS} & 0.2 & 0.5 & 0.6 & \ldots \end{bmatrix}$$

For evaluating the circularity index, substitute the values of inheritance or diagonal elements into the matrix [CS]; matrix [CS] is represented as follows:

$$[CS] = \begin{bmatrix} - & IPP & LGP & CP & FP & EP & SP \\ IPP & 2.5056 & 0.7 & 0.4 & 0.6 & 0.4 & 0.5 \\ LGP & 0.3 & 2.3779 & 0.4 & 0.6 & 0.5 & 0.6 \\ CP & 0.6 & 0.3 & 1.5272 & 0.5 & 0.6 & 0.4 \\ FP & 0.8 & 0.5 & 0.6 & 1.7162 & 0.7 & 0.5 \\ EP & 0.5 & 0.6 & 0.2 & 0.4 & 2.1813 & 0.6 \\ SP & 0.5 & 0.4 & 0.5 & 0.6 & 0.4 & 1.3115 \end{bmatrix}$$

Further, estimate the value of circularity index for strategy by estimating the permanent function for [CS]. The permanent function value has been evaluated using either the function per(A) or the C + + program. The result for strategy is presented as [CS] index = 184.811, and the maximum value of [CS] is estimated by substituting the maximum value of attributes, i.e., 1 (from Table 4) in the diagonal elements. The permanent function value for each of the SBSC is obtained in the same manner. For, per(IPP) = 4.0792, per(LGP) = 3.8051, per(CP) = 3.8459, per(FP) = 3.158, per(EP) = 3.7951, and per(SP) = 3.4661. Further, the maximum value of circularity index is $[CS]_{max} = 3619.58$, and percentage of circularity is 5.10%. Therefore, the highest circularity attribute of SBSC is found to be internal process perspectives (IPP). The permanent function value of internal process perspective, i.e., per(IPP) is 2.5056, and maximum value is 4.0792. It includes reduce, reuse, recycle, and refurbished sub-attributes perspective of the internal process perspective. The findings are supported with a study performed by Kirchherr et al. (2017), which recommends that CE approaches (reduce, reuse, and recycle) are helpful in the economic prosperity, environmental quality, and social equity for the future generations. Further, Chau et al. (2021) also observed that recycling industries in developing economies like India are encouraged by the opportunity for profit.

The second significant perspective is the customer perspective (CP), which consists of return product quality, management support, strategic planning, and decision-making strategy. Customer or stakeholder engagement mainly focuses on meeting the requirements and expectations of the customer (Agrawal et al. 2016a). In most cases, quality of the returned product is not up to the mark. Therefore, the customer does not take much interest in buying such products. To overcome these issues, management needs to develop a proper strategy and plan to maintain the quality of the secondary product. The I4.0-based technologies have the potential to overcome these situations by providing real-time analytics, data availability for strategic development, minimizing the rate of failure, and improving services to end customers.

The permanent function value of learning and growth perspective (LGP) is the third highest, i.e., 3.8051. It consists of the technological roadmap, big data analytics, cybersecurity, and seamless integration. The technological roadmap provides the information on product quality, quantity, and remaining life of the product using I4.0-based technology. de Sousa Jabbour et al. (2018) also found that the adoption of I4.0 technologies from the CE viewpoint enhances the circularity of products through real-time status, which enhances visibility and risk mitigation entire the product life cycle. Further, big data analytics may be useful for storing, processing, and converting unstructured data into a structured form, which may help in dynamic reverse flow of the products or CE implementations. Cybersecurity may protect critical and sensitive data from unwanted threats, and finally, seamless integration may exchange the useful information and data sharing to confirm the minimum utilization of available resources and costs for achieving organizational sustainability (Rajput and Singh 2019). These technologies aim to reduce power and water consumption, CO_2 emission, environmental damage, and low-cost circular supply chain with feedback control.

The next significant perspectives are related to sustainability, i.e., environmental, social, and financial perspectives that manufacturing organizations should initially implement. The environmental perspective is concerned with waste reduction, zero waste, and the reduction of energy consumption (Awan and Sroufe 2022). Similarly, the findings of de Sousa Jabbour et al. (2018), Rajput and Singh (2019), and Kristoffersen et al. (2020) suggested that the sustainable performance of manufacturing organizations could be measured through I4.0 standards. These perspectives maximize the life cycle of end-of-life products through sustainable production and consumption. Similarly, social perspective deals with employee satisfaction, employment training programs, corporate social responsibility, and increased worker safety. The social perspective establishes the social responsibilities within the organization for developing working environment conditions for employees, welfare schemes, employment skill development programs, education, health standards, and supporting social projects. Finally, financial perspective is concerned with infrastructure cost, return on investment, redesigning cost, and high implementation cost. The implementation of I4.0 technologies needs dedicated infrastructure, including sensors technology, storing device, and integrated database servers, which requires high investment. Further, developing recycling, remanufacturing, and inspection facilities of CE requires a huge investment within the organization. Consequently, manufacturers and industrialists wait for a long time to get a return on investment, revenue generation, and profit sharing. Considering all these perspectives and taking proactive measures to address them may encourage organizations to adopt these perspectives and implement I4.0 and CE with relative ease.

Sensitivity analysis

Sensitivity analysis can be utilized to examine whether an optimal solution for any system or framework is robust

enough to be used for decision-making (Pannell, 1997). This analysis can be applied in two different ways: first, by adjusting the weights given to each challenge, and second, by changing the weights given to each expert. Xia et al. (2015) used sensitivity analysis by altering the weights given to each expert to test the viability of the study of the cause-and-effect relationship. Further, sensitivity analysis is implemented to improve the reliability of results that ensure minimum uncertainty in the final results. This technique aids in determining whether a certain expert is biased, which may have an impact on the study's findings. The authors of this research adopted this technique by changing the weight assigned to each expert. Primarily, expert 1 is calculated by permanent function values (0.6, 0.8, 0.7, 0.9) by putting these values in the matrix [IPP]. This situation is represented by Case 1. Further, the weight of experts are varied for the other cases, such as Cases 2, 3, and 4. Table 6 shows the changing weight of the experts for the sensitivity analysis. The results of the sensitivity analysis graph are shown in Fig. 5. The results show that Case 1, Case 3, and Case 4 have the highest value for Internal Process Perspective and Learning and Growth Perspective. Similarly, In Case 2 Learning and Growth Perspective has the highest value followed by Internal Process Perspective. Only Case 2 exhibits variances among the other cases. The ranking order of the most important aspects remained similar even in these cases, with only minor variations.

Table 6 Changing the weight of the experts for the sensitivity analysis

Attributes/Sub-attributes	Case 1	Case 2	Case 3	Case 4
Internal Process Perspective (IPP)				
Recycle (RCY)	0.6	0.7	0.6	0.8
Reuse (RUS)	0.8	0.8	0.9	0.9
Refurbished (RFB)	0.7	0.6	0.7	0.5
Reduce (RDC)	0.9	0.8	0.8	0.7
Learning and Growth Perspective (LGP)				
Technological Roadmap (TR)	0.9	0.7	0.8	0.9
Big Data Analytics (BDA)	0.8	0.9	0.6	0.7
Cybersecurity (CBS)	0.8	0.8	0.9	0.6
Seamless Integration (SI)	0.6	0.7	0.8	0.5
Customer Perspective (CP)				
Return Product Quality (RPQ)	0.6	0.5	0.7	0.6
Management Support (MS)	0.7	0.7	0.6	0.8
Strategic Planning (SP)	0.5	0.6	0.5	0.7
Decision-Making Strategy (DMS)	0.4	0.6	0.4	0.5
Financial Perspective (FP)				
Infrastructure Cost (IC)	0.8	0.7	0.8	0.9
Return on Investment (ROI)	0.7	0.9	0.6	0.8
Redesigning Cost (RC)	0.6	0.8	0.6	0.5
High Implementation (HIC)	0.8	0.8	0.7	0.7
Environmental Perspective (EP)				
Waste Reduction (WR)	0.8	0.9	0.6	0.7
Reduce CO_2 Emission (RCE)	0.6	0.5	0.7	0.5
Reduce Energy Consumption (REC)	0.7	0.6	0.5	0.5
Zero Wastage (ZW)	0.8	0.5	0.6	0.7
Social Perspective (SP)				
Employee Satisfaction (ES)	0.5	0.4	0.6	0.3
Corporate Social Responsibility (CSR)	0.7	0.6	0.7	0.8
Employment Training Programs (ETP)	0.4	0.6	0.4	0.5
Increase Worker Safety (IWS)	0.6	0.4	0.8	0.5

Managerial and strategic implications

This section discusses the managerial and strategic implications of the study. The study may provide several research implications for industrialists, practitioners, policymakers, and supply chain partners. The presence of SBSC attributes and sub-attributes may improve the operational and financial efficiency of manufacturing organizations and may assist industrialists in contributing to the organization's sustainability initiatives. In the proposed framework, GTMA has been used as a decision-making technique. A combination of GTMA along with SBSC provides a feasible and comprehensive understanding of the attribute for circularity decisions. The circularity alternatives in this case study are strategy based, which provides direction to the industrialists for alternative selection in accordance with their organizational needs. The barrel manufacturing organization has contributed significantly toward economic push to the developing economies, although it is restricted because of low return on investment, high implementation cost, and irregular cash flows. Further, organization has focused more toward the redesigning strategies such as reduce, reuse, recycle, and lean manufacturing to achieve sustainable development. Moreover, the internal process perspective (found maximum circularity value in the proposed study) is the most optimum alternative among the other considered alternative strategies for the organization.

The findings reveal that 'Customer Perspective (CP)' and 'Learning and Growth Perspective (LGP)' attributes have ranked in second and third position, respectively. The other attributes with the highest priority are 'Environmental Perspective (EP)', Social Perspective (SP), and 'Financial Perspective (FP)', which manufacturing organizations should initially implement. These results may also help to take better decisions in the context of economic, environmental, and social perspectives. The results of the study may be utilized by the industrialists and practitioners to improve the circularity in supply chain using I4.0-based technologies. Additionally, the implementation of I4.0 from CE perspective may provide advanced process control, minimize equipment

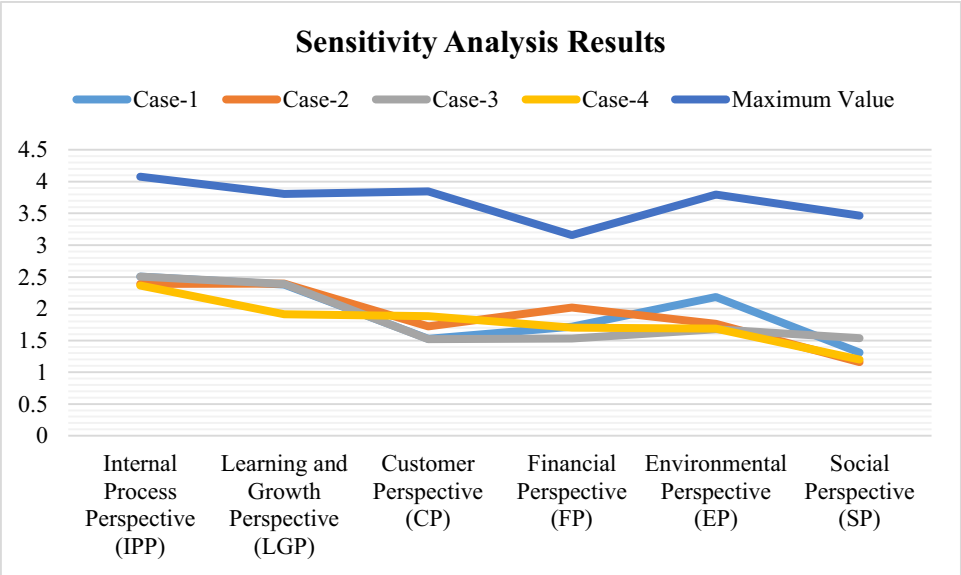


Fig. 5 Sensitivity analysis graph of all SBSC perspectives from all four cases

downtime, and make organizations more flexible, efficient, and productive. The integration may also improve circularity as well as triple bottom line performance, which reduces the greenhouse gas emissions, energy consumption, wasteful actions, and also prevents environmental degradation caused by resource extraction. It also helps to reduce global warming and ultimately results in the development of sustainable competitive advantage by utilizing the CE and I4.0-based technologies (Fathollahi-Fard et al. 2021). Therefore, the findings may be very beneficial for those industrialists and practitioners willing to transform their manufacturing operations from traditional to smart industries.

Conclusions

The present study considers the economic, environmental, and social characteristics of the sustainability and integrates them with a newly constructed SBSC to deliver a comprehensive summary of circularity decisions. Further, by integrating CE with I4.0 technologies, the overall performance of the organization increases. The GTMA is used in the proposed research to determine the optimum circularity alternative in accordance with several circularity attributes. The attributes are chosen by virtue of a detailed literature review and from the opinions of experts. All corresponding attributes, as well as alternatives, are considered from the sustainable development and technologies implementation perspective of the CE and I4.0. The successful implementation of CE in I4.0 can provide cost-effective solutions, reduction of breakdowns, optimize production flow, and maximize the circularity of resources in the considered barrel manufacturing organization. Also, GTMA is utilized since it can handle a large number of nodes and enables the interaction among various attributes. The case illustration provides a quantitative analysis to understand circularity decisions of CE and I4.0 to achieve sustainability.

The attributes, sub-attributes, and alternative strategies in the proposed research are all focused on a single organization, which is one of its limitations. Conversely, methodologies may easily adapt to various strategies and can incorporate a variety of qualitative and quantitative aspects into account depending on the needs of the organization. Further limitation is the study's use of a group of four experts. In the future, a larger number of experts may be employed, and therefore more case studies might be established for the purposes of generalizing findings and recommendations. To avoid any ambiguity or subjective bias, a scale with a fuzzy base may also be utilized. Further, the identified circularity alternatives can be used to develop their causal and dependencies. Additionally, researchers may also work to establish a survey model by using structural equation modeling or *t* test in future research. Furthermore, the proposed framework may be highly valuable for opting the appropriate circularity decisions.

Authors' contributions All authors contributed to the study conception and design. Document preparation, data collection, and analysis were performed by AS, SA, and CPG. The first draft of the manuscript was written by AS, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Data availability Not applicable.

Declarations

Ethics approval and consent to participate Not applicable.

Consent for publication Not applicable.

Competing interests The authors declare no competing interests.

References

Abdul-Hamid AQ, Ali MH, Tseng ML, Lan S, Kumar M (2020) Impeding challenges on industry 40 in circular economy: palm oil industry in Malaysia. Comput Operat Res 123:105052

Agrawal S, Singh RK, Murtaza Q (2016) Outsourcing decisions in reverse logistics: sustainable balanced scorecard and graph theoretic approach. Resour Conserv Recycl 108:41–53

Agrawal S, Singh RK, Murtaza Q (2016b) Disposition decisions in reverse logistics: graph theory and matrix approach. J Clean Prod 137:93–104

Ali Q, Parveen S, Yaacob H, Zaini Z (2022) The management of Industry 4.0 technologies and environmental assets for optimal performance of industrial firms in Malaysia. Environ Sci Pollut Res 1–20.

Awan U, Sroufe R (2022) Sustainability in the circular economy: insights and dynamics of designing circular business models. Appl Sci 12(3):1521

Bag S, Gupta S, Kumar S (2021) Industry 4.0 adoption and 10R advance manufacturing capabilities for sustainable development. Intern J Prod Econ 231:107844.

Bangwal D, Bahuguna PC, Kumar R, Damodaran A, Walker J (2022) Transforming HR through BI and information technology. Bus Int Hum Resour Manage 49–61.

Bocken NM, Olivetti EA, Cullen JM, Potting J, Lifset R (2017) Taking the circularity to the next level: a special issue on the circular economy. J Ind Ecol 21(3):476–482. https://doi.org/10.1111/jiec.12606

Chau MQ, Nguyen XP, Huynh TT, Chu VD, Le TH, Nguyen TP, Nguyen DT (2021) Prospects of application of IoT-based advanced technologies in remanufacturing process towards sustainable development and energy-efficient use. Energy Sources A: Recovery Util Environ Eff 1–25.

Chauhan C, Parida V, Dhir A (2022) Linking circular economy and digitalisation technologies: a systematic literature review of past achievements and future promises. Technol Forecast Soc Chang 177:121508.

Chong P L, Singh AK, Kyong FY (2022) Renewable energy from living plants to power IoT sensor for remote sensing. ADBU J Eng Technol 11(1).

Dantas TE, De-Souza ED, Destro IR, Hammes G, Rodriguez CMT, Soares SR (2021) How the combination of Circular Economy and Industry 4.0 can contribute towards achieving the Sustainable Development Goals. Sustain Prod Consum 26:213–227. https://doi.org/10.1016/j.spc.2020.10.005

De Oliveira CT, Dantas TET, Soares SR (2021) Nano and micro level circular economy indicators: assisting decision-makers in circularity assessments. Sustain Prod Consum 26:455–468. https://doi.org/10.1016/j.spc.2020.11.024

de Sousa Jabbour ABL, Jabbour CJC, Foropon C, Godinho Filho M (2018) When titans meet—can industry 4.0 revolutionise the environmentally-sustainable manufacturing wave? The role of critical success factors. Technol Forecast Soc Chang 132:18–25. https://doi.org/10.1016/j.techfore.2018.01.017

Ellen MacArthur Foundation (2013) Towards a circular economy: opportunities for the consumer goods sector. 2. https://ellenmacarthurfoundation.org/business/report/ce2013

Epstein MJ, Wisner PS (2001) Using a balanced scorecard to implement sustainability. Environ Qual Manag 11(2):1–10. https://doi.org/10.1002/tqem.1300

Fallahpour A, Wong KY, Rajoo S, Fathollahi-Fard AM, Antucheviciene J, Nayeri S (2021) An integrated approach for a sustainable supplier selection based on Industry 4.0 concept. Environ Sci Pollut Res 1–19.

Fathollahi-Fard AM, Ahmadi A, Karimi B (2022) Sustainable and robust home healthcare logistics: a response to the covid-19 pandemic. Symmetry 14(2):193. https://doi.org/10.3390/sym14020193

Fathollahi-Fard AM, Woodward L, Akhrif O (2021) Sustainable distributed permutation flow-shop scheduling model based on a triple bottom line concept. J Ind Inf Integr 24:100233

Garg CP, Kashav V (2021) Modeling the supply chain finance (SCF) barriers of Indian SMEs using BWM framework. J Bus Ind Mark 37(1):128–145. https://doi.org/10.1108/JBIM-05-2020-0248

Gómez AMM, González FA, Bárcena MM (2018) Smart eco-industrial parks: a circular economy implementation based on industrial metabolism. Resour Conserv Recycl 135:58–69. https://doi.org/10.1016/j.resconrec.2017.08.007

Kamble SS, Gunasekaran A, Gawankar SA (2018) Sustainable Industry 4.0 framework: a systematic literature review identifying the current trends and future perspectives. Process Saf Environ Prot 117:408–425. https://doi.org/10.1016/j.psep.2018.05.009

Kashav V, Garg CP, Kumar R (2021) Ranking the strategies to overcome the barriers of the maritime supply chain (MSC) of containerized freight under fuzzy environment. Ann Oper Res 1–46

Kirchherr J, Reike D, Hekkert M (2017) Conceptualizing the circular economy: an analysis of 114 definitions. Resour Conserv Recycl 127:221–232. https://doi.org/10.1016/j.resconrec.2017.09.005

Kristoffersen E, Blomsma F, Mikalef P, Li J (2020) The smart circular economy: a digital-enabled circular strategies framework for manufacturing companies. J Bus Res 120:241–261. https://doi.org/10.1016/j.jbusres.2020.07.044

Kumar R, Tiwari S, Kansara S (2021) Barriers prioritization of the indian steel industry supply chain: applying AHP and fuzzy AHP method. Vision, 09722629211065687.

Laskurain-Iturbe I, Arana-Landín G, Landeta-Manzano B, Uriarte-Gallastegi N (2021) Exploring the influence of industry 4.0 technologies on the circular economy. J Clean Prod 321:128944.

Prakash C, Barua MK (2015) Integration of AHP-TOPSIS method for prioritizing the solutions of reverse logistics adoption to overcome its barriers under fuzzy environment. J Manuf Syst 37:599–615. https://doi.org/10.1016/j.jmsy.2015.03.001

Rajput S, Singh SP (2019) Connecting circular economy and industry 4.0. Int J Inf Manage 49:98–113. https://doi.org/10.1016/j.ijinfomgt.2019.03.002

Rao RV (2007) Decision making in the manufacturing environment: using graph theory and fuzzy multiple attribute decision making methods. London: Springer 2:294.

Saidani M, Yannou B, Leroy Y, Cluzel F, Kendall A (2019) A taxonomy of circular economy indicators. J Clean Prod 207:542–559. https://doi.org/10.1016/j.jclepro.2018.10.014

Singh RK, Agrawal S, Sahu A, Kazancoglu Y (2021) Strategic issues of big data analytics applications for managing health-care sector: a systematic literature review and future research agenda. TQM J. https://doi.org/10.1108/TQM-02-2021-0051

Singh P, Elmi Z, Meriga VK, Pasha J, Dulebenets MA (2022) Internet of Things for sustainable railway transportation: past, present, and future. Clean Log Supp Chain 4:100065

Sivarajah U, Kamal MM, Irani Z, Weerakkody V (2017) Critical analysis of Big Data challenges and analytical methods. J Bus Res 70:263–286. https://doi.org/10.1016/j.jbusres.2016.08.001

Slaveykova VI, Couture P, Duquesne S, D'hugues P, Sánchez W, (2019) Recycling, reuse, and circular economy: a challenge for ecotoxicological research. Environ Sci Pollut Res 26(21):22097–22100

Soleimani H., Chhetri P, Fathollahi-Fard AM, Mirzapour Al-e-Hashem SMJ, Shahparvari S (2022) Sustainable closed-loop supply chain with energy efficiency: Lagrangian relaxation, reformulations and heuristics. Ann Oper Res 1–26.

Lopes de Sousa Jabbour AB, Jabbour CJC, Godinho Filho M, Roubaud D (2018) Industry 4.0 and the circular economy: a proposed research agenda and original roadmap for sustainable operations. Ann Oper Res 270(1): 273–286.

Sun X, Yu H, Solvang WD, Wang Y, Wang K (2021) The application of Industry 4.0 technologies in sustainable logistics: a systematic literature review (2012–2020) to explore future research opportunities. Environ Sci Pollut Res 1–32

Tang YM, Chau KY, Fatima A, Waqas M (2022) Industry 4.0 technology and circular economy practices: business management strategies for environmental sustainability. Environ Sci Pollut Res 1–18.

Tian G, Zhang C, Fathollahi-Fard AM, Li Z, Zhang C, Jiang Z (2022) An enhanced social engineering optimizer for solving an energy-efficient disassembly line balancing problem based on bucket brigades and cloud theory. IEEE Trans Industr Inform 1–11.

Vimal KEK, Kulatunga AK, Ravichandran M, Kandasamy J (2021) Application of multi grade fuzzy approach to compute the circularity index of manufacturing organizations. Procedia CIRP 98:476–481. https://doi.org/10.1016/j.procir.2021.01.137

Virmani N, Salve UR, Kumar A, Luthra S (2021) Analyzing roadblocks of Industry 4.0 adoption using graph theory and matrix approach. IEEE Trans Eng Manag.

Wang W, Tian G, Zhang H, Xu K, Miao Z (2021) Modeling and scheduling for remanufacturing systems with disassembly, reprocessing, and reassembly considering total energy consumption. Environ Sci Pollut 1–17.

Yadav G, Luthra S, Jakhar SK, Mangla SK, Rai DP (2020) A framework to overcome sustainable supply chain challenges through solution measures of industry 4.0 and circular economy: an automotive case. J Clean Prod 254: 120112

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.

JMEPEG
https://doi.org/10.1007/s11665-023-07925-2

©ASM International
1059-9495/$19.00

TECHNICAL ARTICLE

Mechanical and Erosion Characterization of Untreated and Solution-Treated Nitrogen-Alloyed (23-8N) Austenitic Stainless Steel

Mohit Vishnoi, Qasim Murtaza, and Paras Kumar

Submitted: 31 May 2022 / Revised: 15 January 2023 / Accepted: 20 January 2023

The comparative study of as-rolled untreated and solution-treated 23-8N nitronic steel was studied. The results obtained from mechanical characterization revealed that the solution-treated nitronic steel specimens have shown higher toughness and ductility with marginal reduction in hardness as compared to as-rolled untreated nitronic steel specimens. The erosion behavior with surface characterizations of nitronic steel specimen has been evaluated by air jet and slurry jet erosion test rig at four different impingement angles (15°, 45°, 60°, and 90°) and (30°, 45°, 60°, and 90°), respectively. It is observed that the air jet and slurry jet erosion rate for untreated and solution-treated specimens, decreases as the impingement angle increases. The potentiodynamic polarization test in 3.5 wt.% NaCl aqueous solution has been conducted to evaluate the corrosion rate of specimens. The as-rolled solution-treated steel exhibits better corrosion resistance as compared to untreated steel. The surface characterization of eroded specimens was analyzed using scanning electron microscope and possible erosion mechanisms have been studied on untreated and solution-treated specimens.

Keywords air jet erosion, nitronic steel, potentiodynamic polarization, slurry jet erosion

1. Introduction

In the ever-increasing search for materials with high corrosion resistance with substantial tensile strength coupled with hardness, new materials are extensively being tested and experimented with. Austenitic steel is known for its good ductility and superior immunity to a wide range of varying temperatures, from cryogenic temperatures to high temperatures along with its non-magnetic nature (Ref 1). Above all, austenitic stainless steel is extensively utilized in applications subjected to high corrosion primarily, hydro-turbine blades where it is subjected to slurry erosion, and internal combustion engine, valves by virtue of their high corrosion resistance (Ref 2, 3). Though austenitic steel exhibits high corrosion resistance, its mechanical properties such as impact and tensile strength can be increased for better efficiency in hydro and internal combustion engines.

The nitrogen is added to improve the mechanical properties of austenitic steel. The incorporation of nitrogen in austenitic steel exhibited augmentation in wear resistance and strength, without significantly affect in ductility and toughness (Ref 4). This enhancement in mechanical properties, erosion, and corrosion resistance was attributed to nitrogen being an austenite stabilizer due to its superior solubility in austenite (Ref 5). In addition, alloying of nitrogen further decreases the need for incorporation of nickel in the alloy as nitrogen replaces the nickel in alloy formation. This reduction in the quantity of nickel mitigates nickel's harmful effects such as severe allergic reactions to the human system, frequently observed in biomedical applications. Nitrogen-alloyed steel is broadly categorized as nitronic steel, having nitrogen wt % less than 0.4%, and high nitrogen steel (HNS) having nitrogen wt.% more than 0.4% (Ref 6, 7). Among various grades of nitronic steel, 23-8N austenitic steel having $\sim$ 0.29 wt.% nitrogen has immerged as the focus of research in present times. 23-8N exhibited high corrosion and erosion resistance along with high mechanical tensile strength (Ref 8, 9).

To explore the effect of hardening on the mechanical characteristic of steel, various hardening processes namely solid solution hardening, phase transformation hardening, and precipitation hardening were researched (Ref 10-14). Among them, solution annealing (solution hardening) surfaced as a promising method for post heat treatment of 23-8N steel, prominent enhancement in mechanical properties was observed which was later explained based on the effective dissolution of carbides in austenite alloy (Ref 15). Solution treatment also enhanced work hardening (Ref 16). It is well established that work hardening plays a crucial role in determining the erosion resistance of a material. It was observed that as the work hardening coefficient increased in conjugation with ductility, there was a prominent enhancement in erosion resistance of nitronic steel (Ref 17). Similarly, treatment of 23-8N austenitic steel at higher temperature ranges $\sim$ 1200 °C showcased enhancement in mechanical properties pertaining to crack propagation and erosion.

Mohit Vishnoi, Qasim Murtaza, and **Paras Kumar**, Department of Mechanical Engineering, Delhi Technological University, Delhi 110042, India. Contact e-mails: vishnoi.mohit06@gmail.com, qasimmurtaza@dce.ac.in, and paraskumar@dtu.ac.in.

Published online: 14 February 2023

These properties were mechanical flow stress, stacking fault energy, ferrite–austenite mode, and carbide percentage to name a few (Ref 18). This enhancement was attributed to the more effective dissolution of hard carbides into the austenite alloy which consequently formed equiaxed grains (Ref 9). The relation between erosion and carbide elucidates that as the carbide particulate increase, more brittle failure occurs which further promotes erosion. In addition to carbide, ferrite formation and ferrite–austenite interface significantly affected crack propagation. This further determined the impact strength, tensile strength, and ductility (Ref 11, 19-21). Lai et al. (Ref 22) explored the effect of ultrasonic surface rolling on 23-8N austenitic steel for application in engine valves. It was observed that by subjecting 23-8N to ultrasonic rolling, wear resistance increased along with a decrease in coefficient of friction.

In the present study, an attempt has been made to elucidate a comparison between as-rolled untreated and solution-treated 23-8N nitronic steel. Mechanical characterization such as tensile, impact, and hardness test has been carried out on specimens. Moreover, erosion analysis encompassing slurry erosion test and air jet erosion have been performed. The corrosion analysis has also been carried out on untreated and solution-treated specimens. In addition, surface morphology has been examined using scanning electron microscope (SEM).

2. Experimental Details

Plates of 23-8N austenitic stainless steel measuring 80 mm × 80 mm × 40 mm were received from M/S Star Wire (India) Ltd., Ballabhgarh (Haryana). The chemical composition of the material is given in Table 1.

The Pilot study is done at temperatures 1050, 1100, and 1150 °C for 15, 30, 45, and 60 min. The best combinations of properties are obtained at 1100 °C/30 min and the same has been delivered for the experiment.

Kumar et al. (Ref 9) carried out the analysis in the temperature range of 1180–1240 °C on cast nitronic steel and claimed the equiaxed grains and carbide dissolution at an optimum cycle of heat treatment. In the case of as-rolled nitronic steel, the optimum temperature and time were less for grain growth as compared to cast nitronic steel.

In the current work, the optimum cycle of solution treatment at 1100 °C/30 min followed by water quenching has been performed on as-rolled nitronic steel (23-8N). For the characterization of untreated and solution-treated nitronic steel; the samples as per different standards according to the test were cut using computer-controlled wire EDM machine.

2.1 Hardness Test

The Vickers microhardness tester (Model: Matrix DVS 1AT8, Maximum hardness: 1600HV and Maximum load: 1 kg) is used to know the hardness value of the specimens. As per the ASTM-A270 standard, the Vickers hardness (HV30) test is performed on the rolled untreated and solution-treated nitronic steel specimens.

2.2 Impact Test

The 55*10*10 mm^3 (2 mm notch and hammer angle 45°) dimension specimens are prepared by considering the ASTM-E23 standard and the Charpy test (striking energy of test equipment: 300 Joule) is performed on untreated and solution-treated specimens.

2.3 Tensile Test

The Universal testing machine (50KN) with heating chamber software controlled (make-Walter Bai) is used to know the tensile strength value of the specimens. The specimens are prepared as per the ASTM-E8 standard, and the tensile test is performed on untreated and solution-treated specimens.

2.4 Air Jet Erosion Test Rig

The abrasive jet erosion testing is performed according to ASTM G76-13 standard using an air jet erosion test rig (Model: TR-471, Make: DUCOM- Bangalore, India) as shown in Fig. 1 on as-rolled and solution-treated nitronic steel.

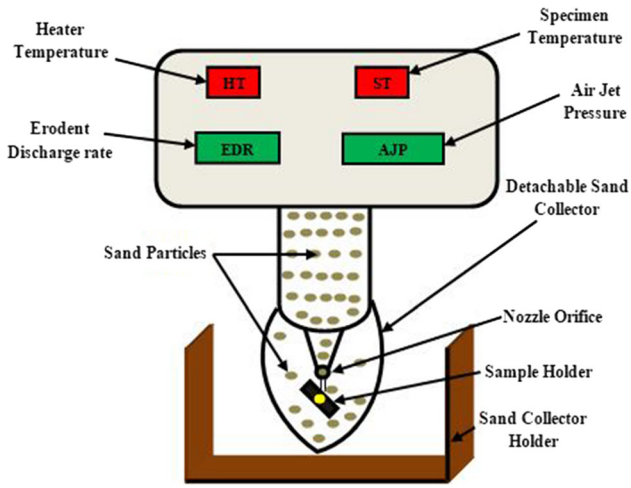


Fig. 1 A scheme of the air jet erosion tester

Table 1 Chemical composition of as-rolled nitronic steel

Element	C	Si	Mn	P	S	Cr	Mo	Ni	Al	Cu
Wt.%	0.32	0.69	1.94	0.03	0.003	22.58	0.12	7.52	0.002	0.13
Element	**Sn**	**Ti**	**V**	**W**	**B**	**Co**	**Nb**	**Pb**	**N**	
Wt.%	0.005	0.004	0.09	0.02	0.0003	0.12	0.06	0.001	0.33	

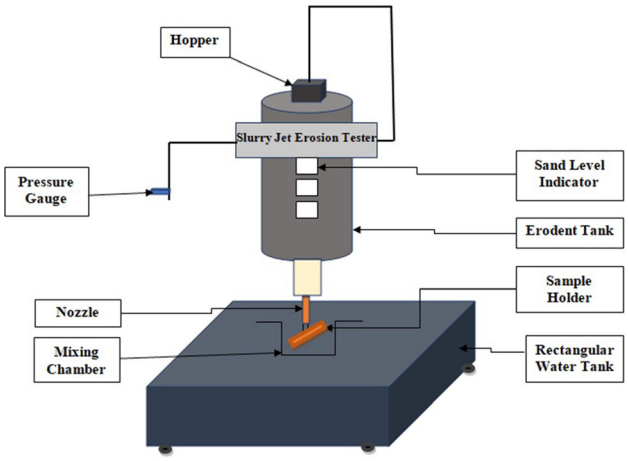


Fig. 2 A scheme of the water jet erosion tester

The samples having dimensions 25 × 25 × 5 mm^3 were used for the air jet erosion test. The naturally available silica having irregular shape sand is used as an erodent.

2.5 Slurry Jet Erosion Test Rig

The slurry erosion testing is performed according to ASTM G-136 standard using a slurry jet erosion test rig (Model: TR-41, Make: DUCOM- Bangalore, India) as shown in Fig. 2 on rolled and solution-treated nitronic steel.

The samples having dimensions 25 × 25 × 5 mm^3 were used for the slurry jet erosion test. The naturally available silica having irregular shape sand is used as an erodent which is mixed with water to form slurry.

2.6 Potentiodynamic Polarization (PDP) Test

The corrosion tests are carried out on a computer-controlled potentiostat PGSTAT30 with GPES and FRA2 software. In this experiment, the sample was connected to the working electrode (WE 1), a platinum electrode was connected to the auxiliary electrode (AE) and the calomel electrode was connected to the reference electrode (RE).

2.7 Scanning Electron Microscope

Surface characterizations of specimens at 2000X are carried out using a scanning electron microscope (SEM) (Model: EVO—Scanning Electron Microscope MA15/18, CARL ZEISS MICROSCOPY LTD.

3. Results and Discussion

3.1 Mechanical Characterization

3.1.1 Hardness Test. There are three readings considered during the experiment and the average value with a deviation (approximate ± 5%) is shown in Fig. 3.

The hardness values of the untreated specimen were more as compared to the solution-treated specimen. Nitronic steel is in the category of austenitic steel and after solution treatment, the hardness of the steel reduced marginally. The hardness of as-rolled steel is probably due to the presence of carbides and after solution treatment; the hardness has reduced due to the dissolution of carbides at 1100 °C (Ref 15). The percentage decrease in hardness value is about 13% as compared to the untreated specimen.

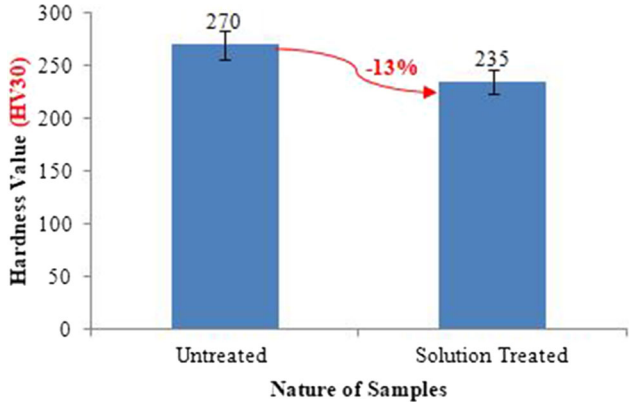


Fig. 3 Vickers hardness values of 23-8N untreated and solution-treated specimen

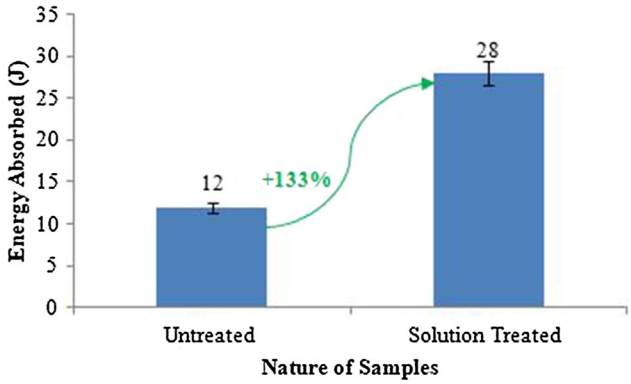


Fig. 4 Impact energy values of 23-8N untreated and solution-treated specimen

3.1.2 Impact Test. There are three samples considered during the experiment and the average value with a deviation (approximate ± 5%) is shown in Fig. 4.

The impact energy absorbed by the solution-treated specimen was more as compared to the untreated as-rolled steel. The solution-treated specimen showed high toughness as compared to the untreated specimen. The percentage increase in impact energy is about 133% as compared to the untreated specimen because after the solution treatment the ductility of the specimen has increased which increases the tendency to absorb more energy.

3.1.3 Fractography of Impact Test Specimens. The fractured specimens of untreated and solution-treated steel are shown in Fig. 5(a) and (c), respectively. The microscopic images of the fractured surfaces are taken at 150X for untreated and solution-treated nitronic steel as shown in Fig. 5(b) and (d).

The fractography analysis showed that the surface of the untreated specimen was smooth and brighter (leads to brittle fracture) but solution-treated specimen surface showed extensive plastic deformation (leads to ductile fracture).

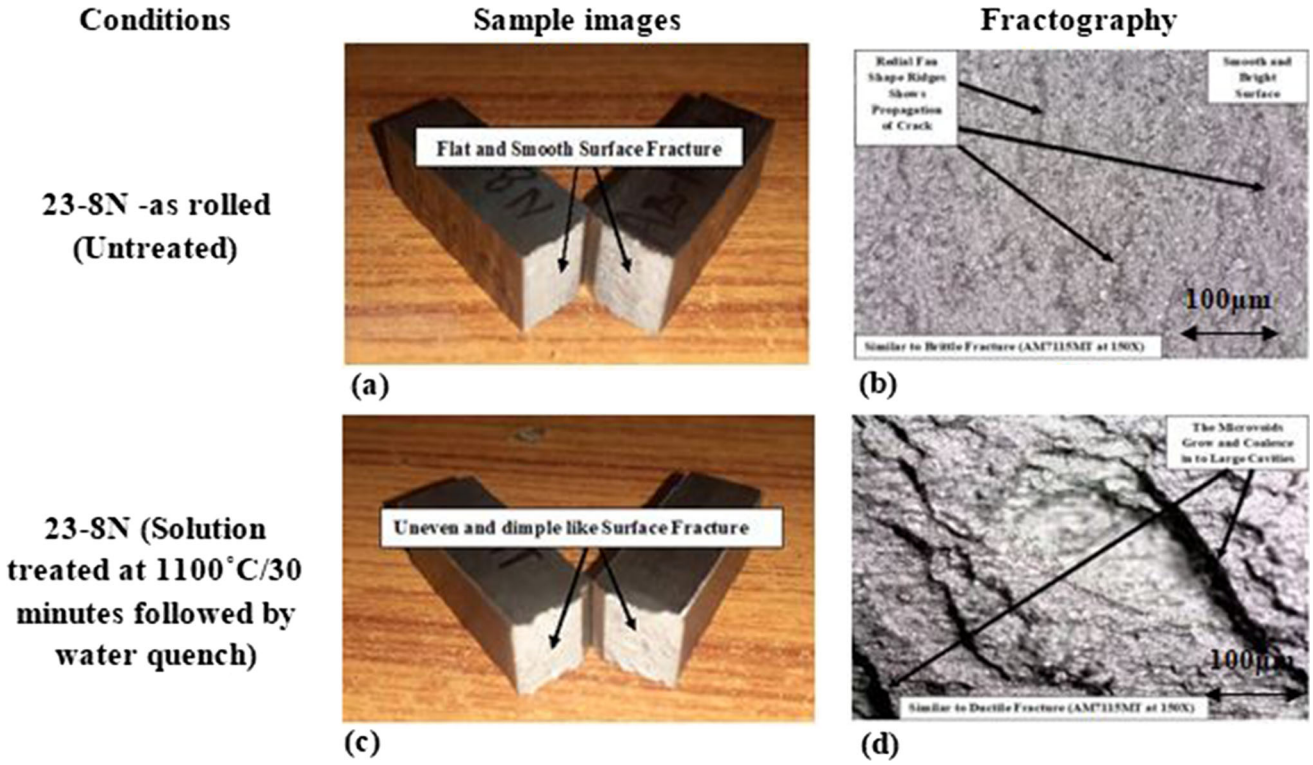


Fig. 5 Impact test specimens and fracture surface images for untreated (a, b) and solution-treated (c, d) specimens, respectively

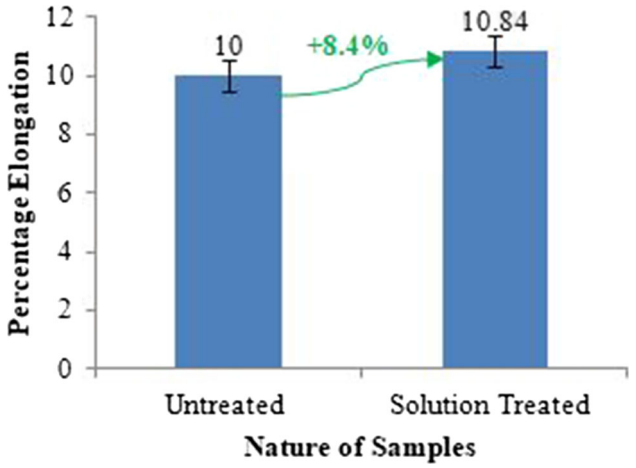


Fig. 6 (%) Elongation of 23-8N untreated and solution-treated specimen

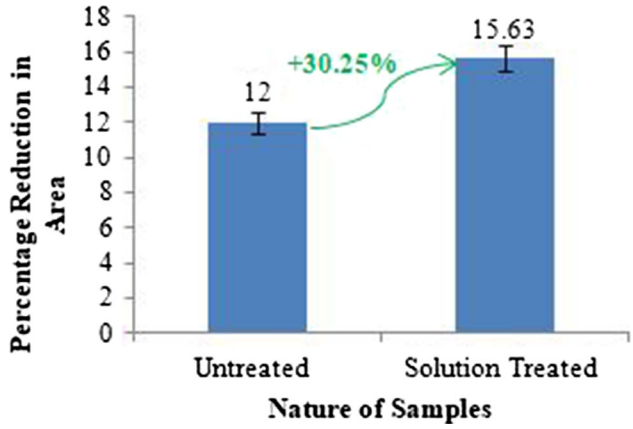


Fig. 7 (%) Reduction in the area of 23-8N untreated and solution-treated specimen

3.1.4 Tensile Test. Three experiments have been conducted and the average value with a deviation (approximate ± 5%) is shown in Fig. 6 and 7.

The percentage elongation and reduction in are is more for solution-treated nitronic steel hence solution-treated steel has more ductility than untreated steel, respectively. The percentage increase in elongation and % reduction in the area is about 8.4 and 30.25% as compared to the untreated specimen, respectively.

The tensile strength is more for an untreated specimen as compared to a solution-treated specimen as shown in Fig. 8 due to the dissolution of carbides at 1100 °C (Ref 15). The mechanical analysis concluded that solution-treated nitronic steel enhances the ductility, toughness, and marginally reduction in hardness. A very similar trend in mechanical properties has also been reported by Gadhikar et al. (2014). The percentage decrease in tensile strength is about 12.2% as compared to the untreated specimen.

3.1.5 Fractography of Tensile Test Specimens. The fractured specimens of untreated and solution-treated steel are shown in Fig. 9(a) and (c), respectively. The microscopic images of the fractured surfaces are taken at 150X for untreated and solution-treated nitronic steel as shown in Fig. 9(b) and (d), respectively.

The fractography analysis showed that the surface of the untreated specimen was smooth and brighter (leads to brittle

fracture) but in solution-treated specimen the surface showed extensive plastic deformation or elongation (leads to ductile fracture).

3.2 Erosion Tests

3.2.1 Air Jet Erosion Test. *3.2.1.1 Selection of Air Jet Erosion Test Parameters.*

a) Selection of Impact Velocity: All the parameters are taken based on the application such as in the wind turbine blade, from the pilot experiments, 90 m/sec is selected due to the high hardness of steel. The velocity (5-20 m/sec) below is not enough to get an even dent on the surface.
b) Selection of Erodent Size: The size of an erodent, which can fly in the desert area, is of the order of 150 μm (max.). Moreover, the prescribed setup can handle up to 150 μm erodent size. So, the 125 μm erodent size is considered in the current experiment.
c) Time duration: The parameters are selected by considering the practical situation and the Pilot study analysis. In the Pilot study, 5 min. time could not get a significant effect on nitronic steel due to its hardness. So, the 10 min. time is considered in the current study.

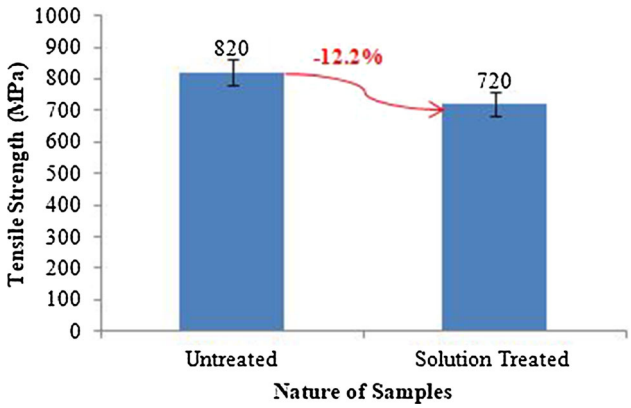


Fig. 8 Tensile strength of 23-8N untreated and solution-treated specimen

From the different literature surveys; the optimum parameters were decided for the test such as the erodent size, impact velocity, time duration, and erodent discharge as 125 μm, 90 m/sec, 10 min, and 5 gm/min, respectively (Ref 23-26).

Three experiments have been carried out at a particular angle and the average value with deviation (approximate ± 5%) is shown in Fig. 10, the air jet erosion rate was calculated for untreated and solution-treated nitronic steel at four different impingement angles as 15, 45, 60, and 90°. As shown in Fig. 10; the dry abrasive jet erosion rate decreases by increasing the impingement angle for both conditions. The air jet erosion rate was lower for solution-treated specimens than the untreated specimen. Selokar et al. performed an air jet erosion test on Martensitic (13-4) and austenitic stainless (23-8N) steel and concluded that the erosion rate is higher in Martensitic steel as compared to nitrogen-alloyed austenitic steel (Ref 27).

Figure 11(a) and (b) shows the micrographs of air jet eroded untreated and solution-treated as-rolled nitronic steel samples. The various type of wear mechanism has been observed during the analysis of micrographs such as crater wear, ploughing, microcutting, and carbide particles pull out.

Through topographic analysis on air jet eroded specimens, it has been observed that light craters were present on both untreated and solution-treated steel. In addition, carbide pull out and small quantities of deep craters were also observed in

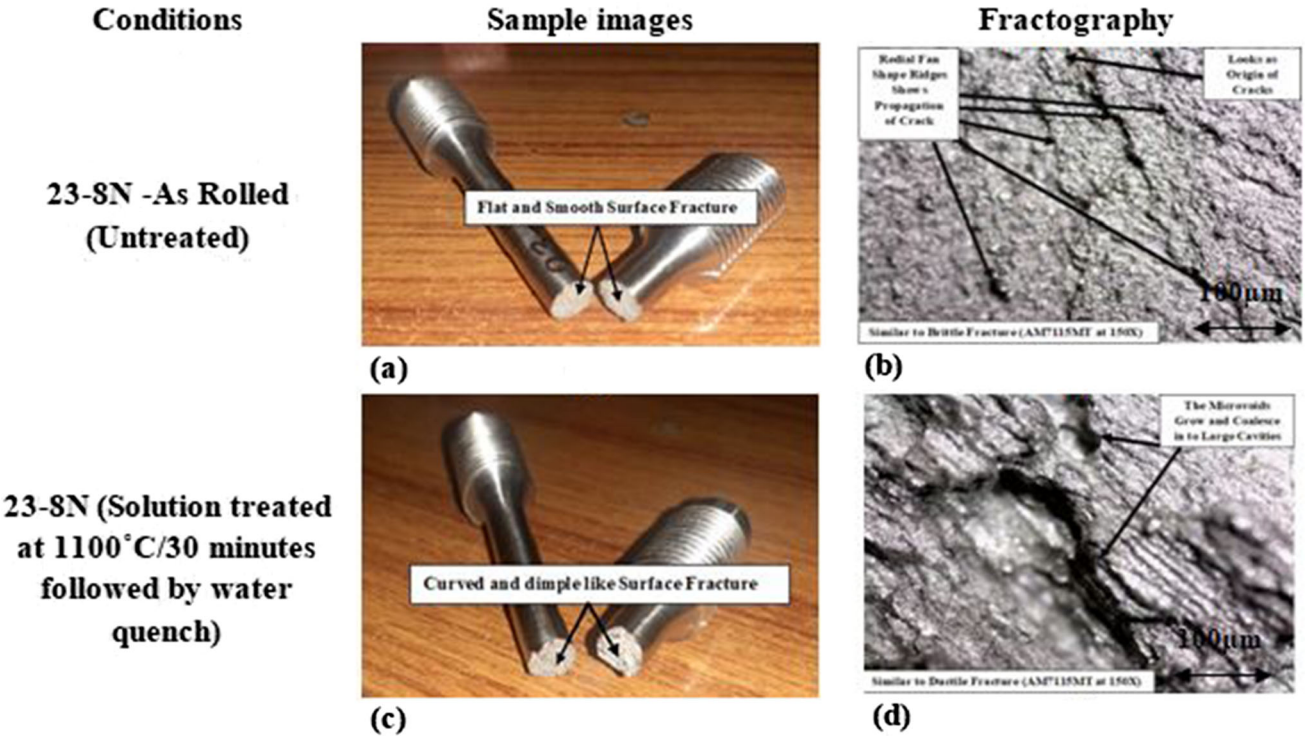


Fig. 9 Tensile test specimens and fracture surface images for untreated (a, b) and; solution-treated (c, d) specimens, respectively

untreated specimen but microcutting and a ploughing mode of material removal were observed in solution-treated specimen.

3.2.2 Slurry Jet Erosion Test. *3.2.2.1 Selection of Slurry Erosion Test Parameters.*

a) Selection of impact velocity: All the parameters are taken based on the application of the hydro-turbine blade, the maximum velocity that may be available in case of hydro-turbine is below 60 m/sec. but in practice, the available velocity is near 30-40 m/sec.
b) Selection of erodent: Clark and Hartwich explained that erosion rate is the function of particle size and observed erosion rate increases with the increase in erodent size from 90 to 300 μm in their experiment (Ref 28). The nitronic steel is very hard so 300 μm erodent size is considered in the current experiment. Moreover, the permissible size of erodent in the case of hydro-turbine is also near 300 μm.
c) Time duration: The parameters are selected by considering the practical situation and the Pilot study analysis. In Pilot study, 5 min. time could not get significant effect on nitronic steel due to its hardness. So, the 10 min. time is considered in current study.

From the different literature surveys; the optimum parameters were decided for the test such as the erodent size, impact velocity, time duration, and erodent discharge as 300 μm, 30 m/sec, 10 min, and 195 gm/min, respectively (Ref 23, 25, 26).

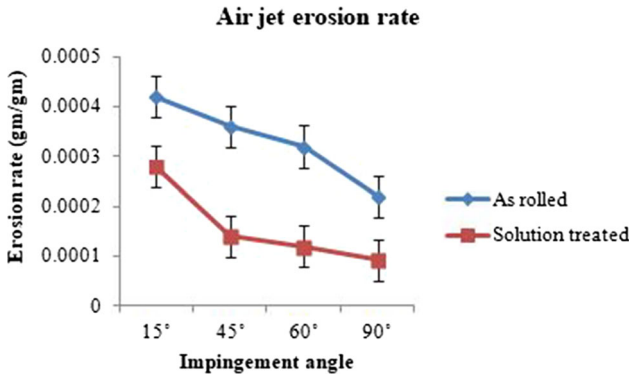


Fig. 10 Plot of abrasive jet erosion rate v/s impingement angle

Three experiments have been carried out at a particular angle and the average value with deviation (approximate ± 5%) is shown in Fig. 12. The slurry erosion rate was calculated for untreated and solution-treated nitronic steel at four different impingement angles such as 30, 45, 60, and 90°. As shown in Fig. 12; the slurry erosion rate decreases by increasing the impingement angle for both conditions. The slurry erosion rate was lower for solution-treated specimens than for untreated specimens (Ref 15).

The ductile materials exhibited maximum loss due to erosion at lower impingement angles and decreased at higher impingement angles. But in the case of brittle materials, the erosion rate increases with an increase in the impingement angle attained the maximum of 90° (Ref 29). Kumar et al. Ref 9 conducted a slurry erosion test on cast and solution-treated nitronic (23-8N) steel and concluded that the solution-treated nitronic steel showed better erosion resistance as compared to cast nitronic steel.

Figure 13(a) and (b) shows the micrographs of slurry jet eroded untreated and solution-treated as-rolled nitronic steel samples. The various type of wear mechanism has been observed during the analysis of micrographs such as crater wear, ploughing, plough lips, and carbide particles pull out. Through topographic analysis on slurry jet eroded specimens, it has been observed that light craters were present on both untreated and solution-treated steel. In addition, carbide pull out, plough lips, and small quantities of deep craters were also observed in untreated specimen but microcutting and a ploughing mode of material removal were observed in solution-treated specimen.

3.2.3 Observations on Erosion Test. The low impingement angle and high impingement angle erosion are shown in Fig. 14 and 15; the variation in these angles is from 15 to 90°.

By considering a low impingement angle, most of the cutting action has been noticed by shearing and in the case of high impingement angle the particles strike at a higher angle maximum of up to 90°; which shows a brittle fashion erosion of material.

There are several impingement angles available in the test rig, many researchers have carried out their research to know the mechanism and effect at various impingement angles on the material. Most of the researchers carried out the variation in

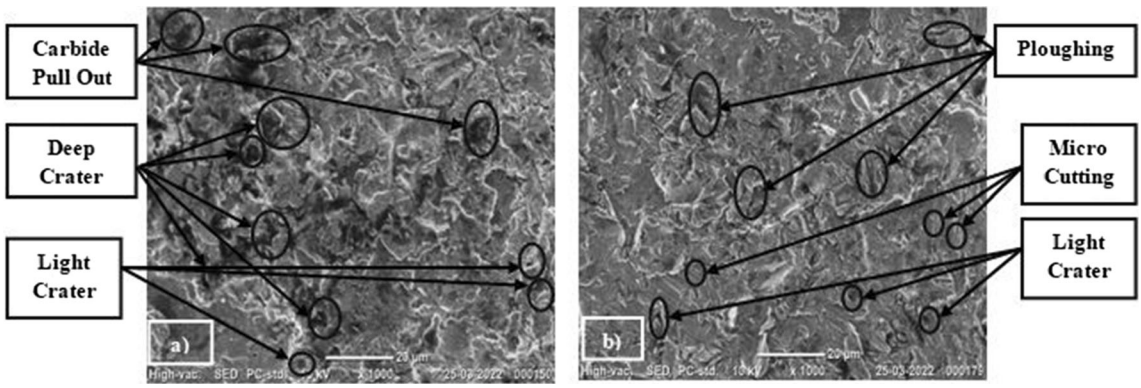


Fig. 11 SEM micrographs of (a) as-rolled untreated and; (b) solution-treated air jet eroded surfaces at 2000X

angle from 15 to 90° to know the erosion rate pattern and revealed their conclusions depending on the behavior of materials (ductile or brittle). The erosion rate was dependent on the type of abrasive particles, size of abrasive particles, the velocity of abrasive particles, contact time of abrasive particles on materials, and angle at which abrasive particles strick on material (Ref 25, 26).

In this paper, the variation in impingement angle has been considered from 15 to 90° to know the pattern of erosion rate. In the slurry jet erosion experiment, the holding of the specimen was difficult at 15° due to the high velocity of slurry at the outlet of the jet; so the impingement angle (30, 45, 60, and 90°) has been considered to know the distribution of plot. From SEM micrographs; different erosion mechanism such as crater wear, formation of lips, ploughing, and removal of hard carbide particles has been observed.

Selokar et al. have also conducted the solid particle erosion experiment on 23-8N steel by considering the varied impingement angle (30, 45, 60, and 90°) and concluded that; at 30°; the erosion loss is mainly due to shear cutting and at 90° the damage was mainly due to embrittlement (Ref 27). Few researchers have also concluded their observation and stated that ductile material shows peak erosion loss at lower impingement angles (15-30°) and brittle materials have peak erosion loss at higher angles (90°) (Ref 30-33).

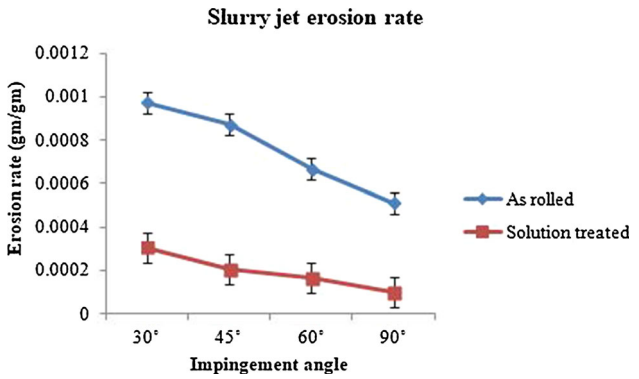


Fig. 12 Plot of slurry jet erosion rate v/s impingement angle

3.3 Corrosion Behavior of Samples

The similar surface conditions were considered in both the cases and the corrosion test has been conducted on without heat-treated and heat-treated samples. Tafel polarization curves of as-rolled untreated and solution-treated nitronic steel samples as shown in Fig. 16 and 17, respectively, are tested in 3.5 wt.% NaCl aqueous solution.

The data obtained from the polarization curve for untreated and solution-treated specimens are depicted in Table 2. The value of I_{corr} is very much related to the corrosion rate. Lower the values of I_{corr} better the corrosion resistance (Ref 34, 35).

From the potentiodynamic polarization curves, the I_{corr} values are obtained by constructing the Tafel plots. It can be seen from Table 2; that solution-treated specimens showed lower corrosion current (I_{corr}) and therefore lower corrosion rate, i.e., better corrosion resistance than the untreated specimen.

The shifting of E_{corr} in the solution-treated specimen is to a more positive value than the untreated specimen. This indicates that solution treatment makes the specimen surface nobler. This does mean that solution-treated specimen offers good corrosion resistance. In general corrosion resistance of the specimens is decided by the I_{corr} values, solution-treated nitronic steel > untreated specimen. However, the solution-treated

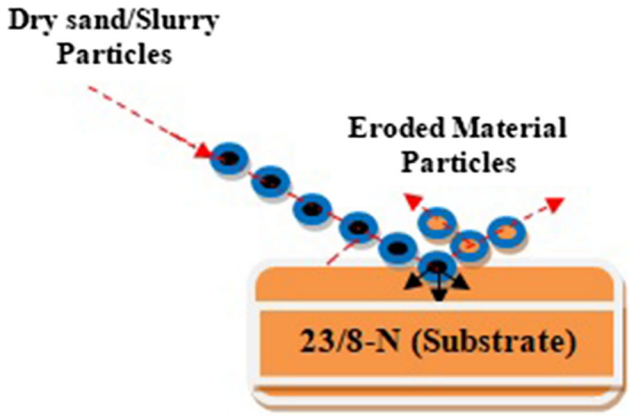


Fig. 14 Low impingement angle erosion on the specimen

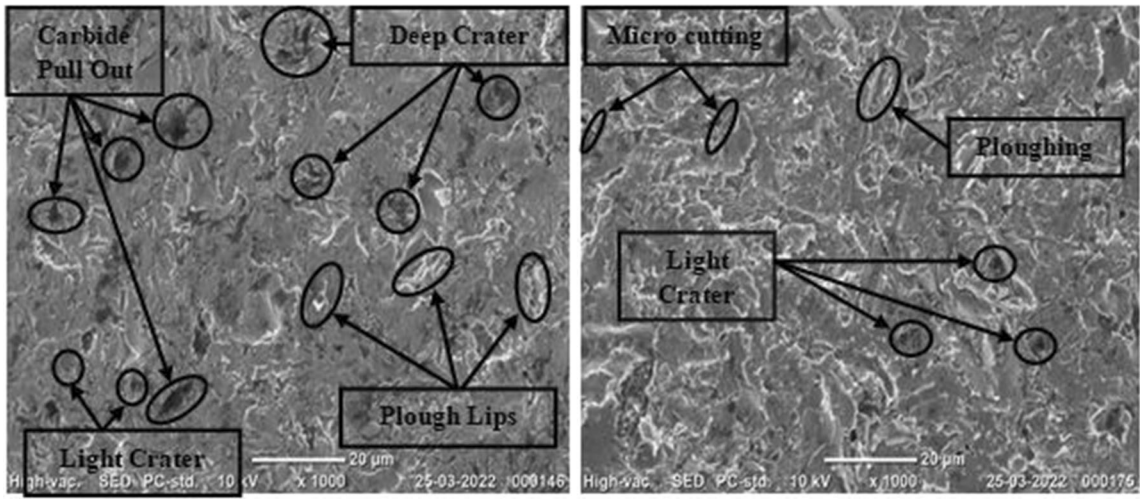


Fig. 13 SEM micrographs of (a) as-rolled untreated and; (b) solution-treated slurry jet eroded surfaces at 2000X

specimen exhibits a wider passive range than the untreated specimen.

Figure 18(a) and (b) shows the micrographs after the poteniodynamic polarization test on untreated and solution-treated specimens, respectively, at 2000X. Figure 18(a) shows a more corrosion rate as compared to Fig. 18(b). As-rolled untreated specimen as shown in Fig. 18(a) revealed that the number of black patches (rusted portion) is more as compared to solution-treated steel as shown in Fig. 18(b). The micrographs conclude that due to the presence of a large number of pores in as-rolled untreated steel than solution-treated specimen. Due to the presence of a large number of pores the corrosion resistance is less as compared to solution-treated steel.

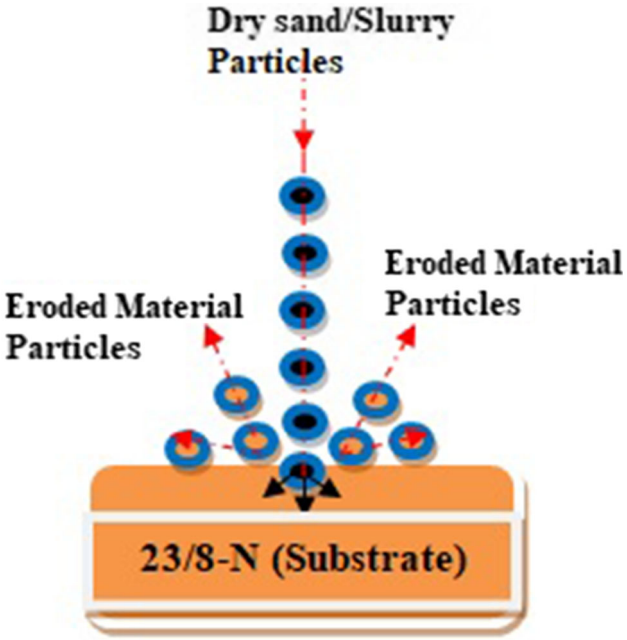


Fig. 15 High impingement angle erosion on the specimen

The micrographs (Fig. 19a and b) have been taken using a scanning electron microscope at 2000X magnification. The deep crater and many pores have been observed in as-rolled steel as shown in Fig. 19(a), whereas there is a reduction in pores and craters in the solution-treated as-rolled steel as shown in Fig. 19(b).

As-rolled means that the material plate is rolled and cooled naturally or there is no heat treatment on steel but after heat treatment of as-rolled steel the mechanical and tribological properties have been increased.

4. Conclusions and Future Scope

The following conclusions have been made:

1. The mechanical properties such as toughness and ductility have been enhanced for solution-treated nitronic steel, but there is a marginal reduction in hardness as compared to untreated steel.
2. The air jet erosion rate is less for as-rolled solution-treated steel as compared to untreated steel.

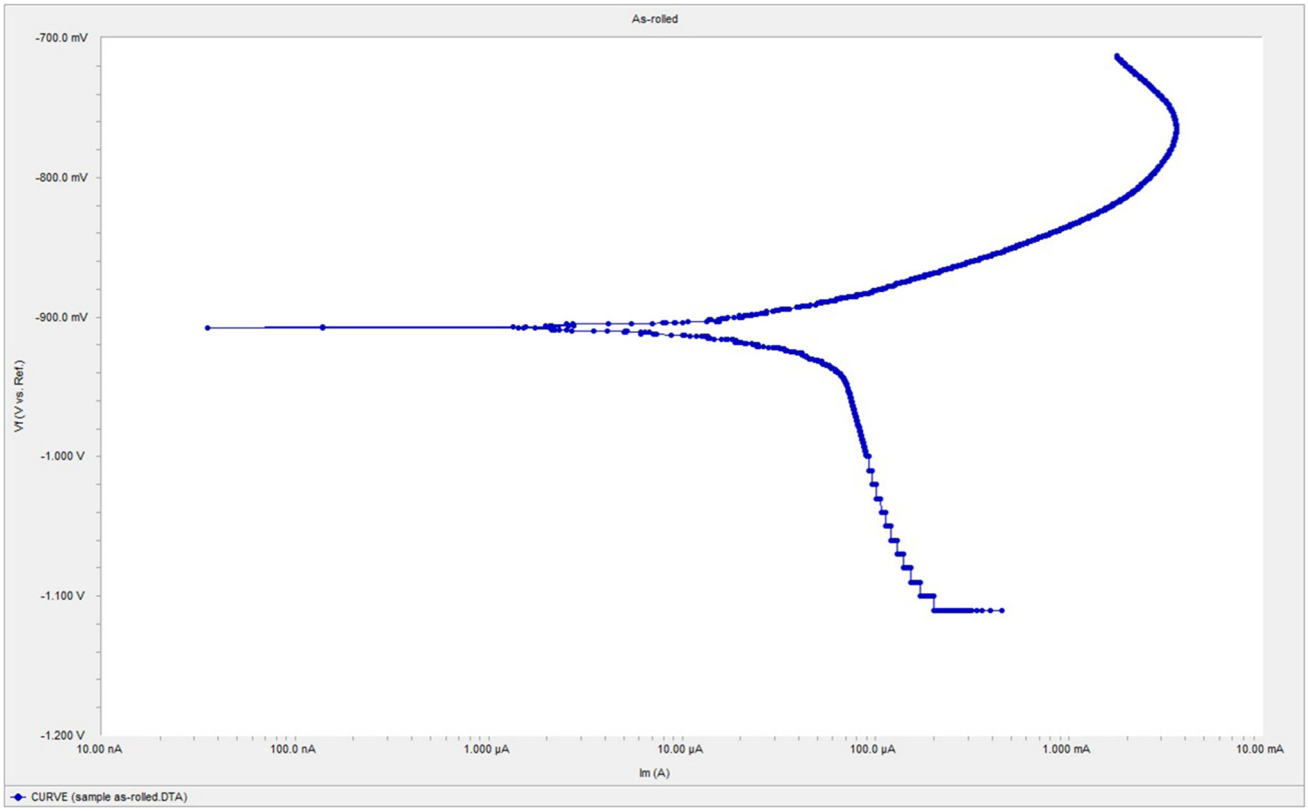


Fig. 16 Potentiodynamic polarization curve for untreated specimen

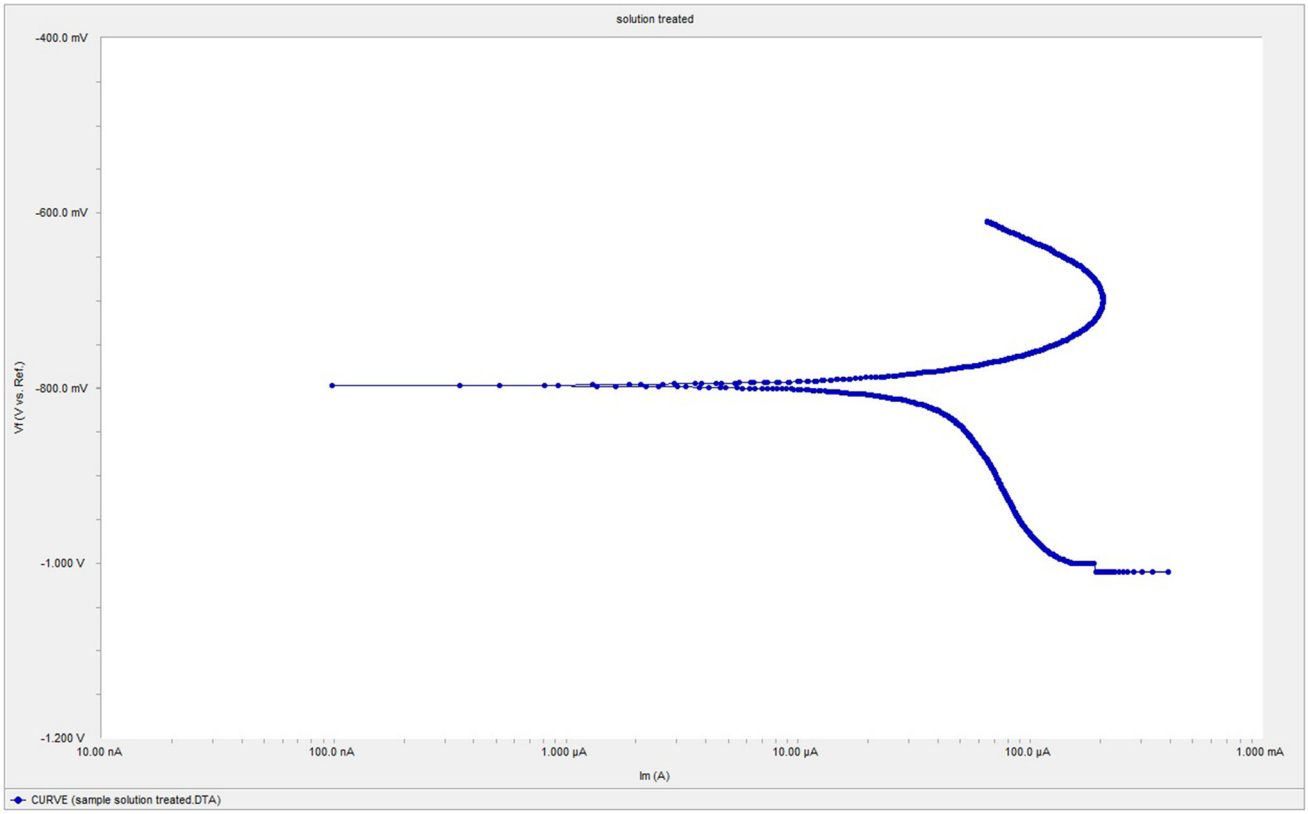


Fig. 17 Potentiodynamic polarization curve for solution-treated specimen

Table 2 Data obtained from potentiodynamic polarization curves

Conditions	As-rolled	Solution-treated
Sample and area, cm^2	4	4
Scan rate, mV/s	0.2	0.2
Sample period, s	1	1
Density, g/cm^3	7.86	7.86
Equivalent wt., gm	27.65	27.70
I_{corr}, $\mu A/cm^2$	0.525	0.425
E_{corr}, mV	− 904	− 799

3. The slurry jet erosion rate is less for as-rolled solution-treated steel as compared to untreated steel.
4. As-rolled solution-treated nitronic steel exhibits excellent corrosion resistance as compared to untreated steel.
5. The different wear mechanisms have been observed from scanning electron microscope (SEM) images such as microcutting, crater wear, ploughing, plough lips, and carbide particles pull out.

This work further can expend by incorporating the optimization techniques on process parameters. The different heat treatment processes can also be applied to as-rolled steel and compared results among them.

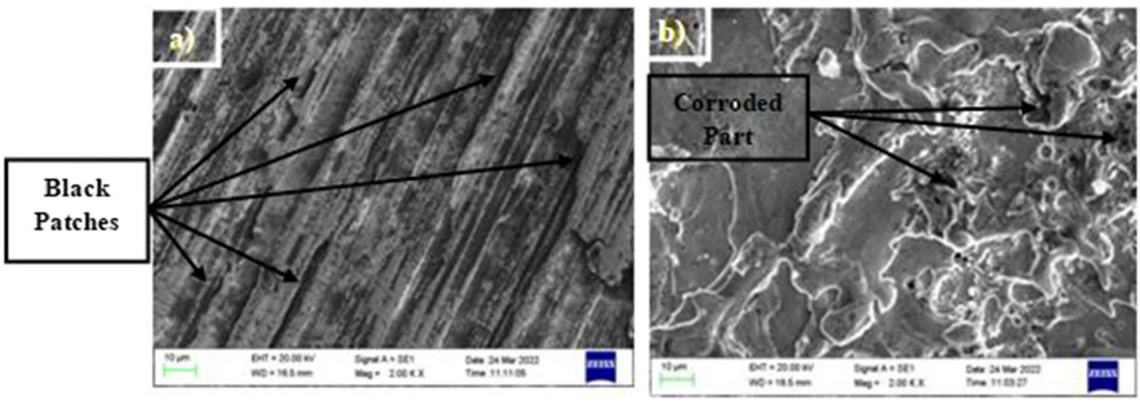


Fig. 18 Scanning electron microscope (SEM) images after corrosion test (a) as-rolled untreated and (b) solution-treated 23-8N steel at 2000X

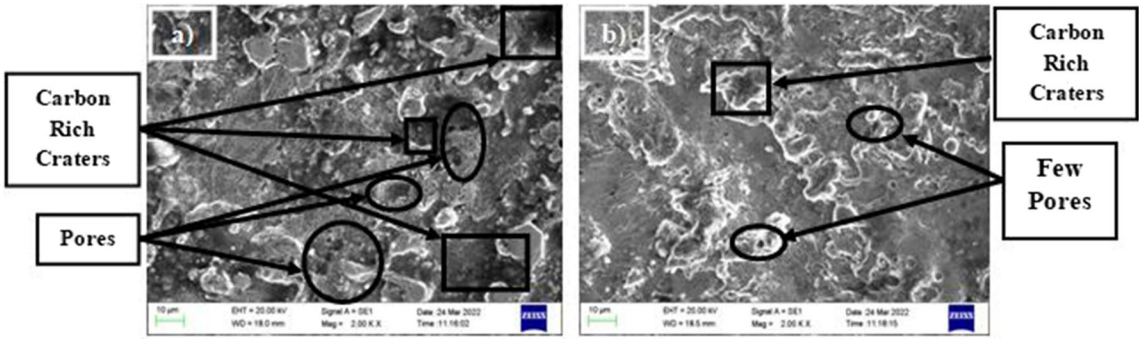


Fig. 19 Microstructure of (a) As-rolled; and (b) Solution-treated as-rolled steel

Acknowledgments

The author wants to thank the JSS Academy of Technical Education, Noida, India, for allowing me to pursue a part-time Ph.D. from the Department of Mechanical Engineering, Delhi Technological University, Delhi 110042, India. The author would like to thank Dr. Amar Patnaik from MNIT Jaipur and Dr. Avinsh Kumar from Star Wire (India) Ltd., Ballabhgarh (Haryana) for helping us in the Tribological Characterization and Providing Material including Mechanical Characterization, respectively.

Author Contributions

MV contributed to conceptualization, investigation, data curation, and writing—original draft. QM contributed to review and editing and supervision. PK contributed to review and editing and supervision.

Funding

No funding has been involved in this work.

Conflict of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. A.A. Gadhikar, C.P. Sharma, D.B. Goel, and A. Sharma, Effect of Heat Treatment on Carbides in 23-8-N Steel, *Met. Sci. Heat Treat.*, 2011, **53**(5), p 293–298.
2. A. Sharma and S.K. Goel, Erosion Behaviour of WC-10Co-4Cr Coating on 23-8-N Nitronic Steel by HVOF Thermal Spraying, *Appl. Surf. Sci.*, 2016, **370**, p 418–426.
3. R. Kumar, D. Bhandari, and K. Goyal, Slurry erosion behaviour of HVOF sprayed coatings on hydro turbine steel: A review. in *IOP Conference Series: Materials Science and Engineering*, vol. 1033, no. 1. IOP Publishing (2021), p. 012064
4. N. Kumar, N. Arora, and S.K. Goel, Weld Joint Properties of Nitrogen-Alloyed Austenitic Stainless Steel Using Multi-pass GMA Welding, *Arch. Civ. Mech. Eng.*, 2020, **20**(3), p 1–13.
5. A. Selokar, D.B. Goel, and U. Prakash, A Comparative Study of Cavitation Erosive Behaviour of 23/8N Nitronic Steel and 13/4 Martensitic Stainless Steel, *Advanced Materials Research*, Vol 585, Trans Tech Publications Ltd., 2012, p 554–558
6. L. Zhao, Z.L. Tian, and Y. Peng, Control of Nitrogen Content and Porosity in Gas Tungsten Arc Welding of High Nitrogen Steel, *Sci. Technol. Weld. Join.*, 2009, **14**(1), p 87–92.
7. H. Sharma, A. Bhardwaj, D. Singh, D. Mittal, R. Kumar, A. Selokar, M. Bilal, S. Akabrid, S.K. Sharmaa, Corrosion Behaviour of Nitronic Steel in Acidic Environment (2021)
8. A.K. Chauhan, D.B. Goel, and S. Prakash, Solid Particle Erosion Behaviour of 13Cr-4Ni and 21Cr-4Ni-N Steels, *J. Alloy. Compd.*, 2009, **467**(1–2), p 459–464.
9. A. Kumar, A. Sharma, and S.K. Goel, Effect of Heat Treatment on Microstructure, Mechanical Properties and Erosion Resistance of Cast 23-8-N Nitronic Steel, *Mater. Sci. Eng. A*, 2015, **637**, p 56–62.
10. N. Kumar, N. Arora, and S.K. Goel, Study on Metallurgical and Mechanical Aspects of GMA Welded Nitronic Steel Under the Influence of Weld Quenching, *J. Manuf. Process.*, 2020, **56**, p 116–130.
11. N. Kumar, N. Arora, and S.K. Goel, Effect of Annealing on Mechanical and Metallurgical Properties of GMA Welded Nitronic Steel, *Mater. Sci. Eng. A*, 2020, **771**, p 138542.
12. M.H. Buszko, A.K. Krella, A. Marchewicz, and G. Gajowiec, Effect of annealing temperature on slurry erosion resistance of ferritic X10CrAlSi18 steel, *Tribol. Int.*, 2021, **153**, p 106648.
13. H. Ji, Z. Cai, W. Pei, X. Huang, and Y. Lu, DRX Behavior and Microstructure Evolution of 33Cr23Ni8Mn3N: Experiment and Finite Element Simulation, *J. Market. Res.*, 2020, **9**(3), p 4340–4355.
14. N. Kumar and N. Arora, Effect of Solution Treatment on Slurry Erosive Wear Performance of Martensitic and Nitrogen Strengthened Austenitic Stainless Steel, *Mater. Lett.*, 2021, **284**, p 128932.
15. A.A. Gadhikar, A. Sharma, D.B. Goel, and C.P. Sharma, Effect of Carbides on Erosion Resistance of 23-8-N Steel, *Bull. Mater. Sci.*, 2014, **37**(2), p 315–319.
16. N. Kumar, N. Arora, and S.K. Goel, Slurry Erosion Study on Nitrogen-Alloyed Austenitic Stainless Steel and Weld Beads, *Tribol. Lett.*, 2020, **68**(3), p 1–16.
17. A.A. Gadhikar, Effect of work hardening ability of nitronic steels on erosion resistance. in *International Conference on Advances in Thermal Systems, Materials and Design Engineering (ATSMDE2017)* (2017)
18. N. Sahu, A. Selokar, and U. Prakash, Thermo-Mechanical Behaviour of 23/8 Austenitic Stainless Steel, *ISIJ Int.*, 2014, **54**(4), p 970–978.
19. V.N. Shukla, D. Kumar, and G. Gupta, Corrosion Studies of Nanostructured AlN Coating Deposited on 23/8N Nitronic Steel by PVD Method, *Mater. Today Proc.*, 2017, **4**(9), p 10216–10220.
20. R. Mariappan, M.A. Prasad, G. Dharmalingam, and D. Sivaprakasham, Microstructure and Mechanical Properties of Hot-Pressed 21-4N oxide-Dispersion-Strengthened Austenitic Stainless Steels, *Metallogr. Microstruct. Anal.*, 2018, **7**(5), p 578–586.
21. S.L. Hernández-Trujillo, V.H. Lopez-Morelos, M.A. García-Rentería, R. García-Hernández, A. Ruiz, and F.F. Curiel-López, Microstructure and Fatigue Behavior of 2205/316L Stainless Steel Dissimilar Welded Joints, *Metals*, 2021, **11**(1), p 93.
22. F. Lai, S. Qu, R. Lewis, T. Slatter, W. Fu, and X. Li, The Influence of Ultrasonic Surface Rolling on the Fatigue and Wear Properties of 23-8N Engine Valve Steel, *Int. J. Fatigue*, 2019, **125**, p 299–313.
23. V.R. Kiragi, A. Patnaik, and T. Singh, Impact of High-Velocity Oxy-Fuel Sprayed TiAlN Surface Coating on Mechanical and Slurry

Erosion Performance of Aluminium Alloys, *Materialwiss. Werkstofftech.*, 2019, **50**(10), p 1250–1261.
24. V.R. Kiragi and A. Patnaik, Erosive Wear Behaviour of Aluminium Alloys: A Comparison Between Slurry and Dry Erosion, *Mater. Res. Express*, 2019, **6**(8), p 086503.
25. V.R. Kiragi, A. Patnaik, T. Singh, and G. Fekete, Parametric Optimization of Erosive Wear Response of TiAlN-coated Aluminium Alloy Using Taguchi Method, *J. Mater. Eng. Perform.*, 2019, **28**(2), p 838–851.
26. A. Babu, G. Perumal, H.S. Arora, and H.S. Grewal, Enhanced Slurry and Cavitation Erosion Resistance of Deep Cryogenically Treated Thermal Spray Coatings for Hydroturbine Applications, *Renew. Energy*, 2021, **180**, p 1044–1055.
27. A. Selokar, R. Kant, D.B. Goel, and U. Prakash, Solid Particle Erosion Behaviour of Martensitic and Nitrogen Alloyed Austenitic Stainless Steel, *Advanced Materials Research*, Vol 1043, Trans Tech Publications Ltd., 2014, p 45–49
28. H.M. Clark and R.B. Hartwich, A Re-examination of the 'Particle Size Effect' in Slurry Erosion, *Wear*, 2001, **248**(1–2), p 147–161.
29. M. Vishnoi, P. Kumar, and Q. Murtaza, A Review on Coating for Hydro-turbine Application by HVOF Process, *Advances in Materials Engineering and Manufacturing Processes*. Springer, Singapore, 2020, p 79–103
30. L. Fuyan and S. Hesheng, The Effect of impingement Angle on Slurry Erosion, *Wear*, 1991, **141**(2), p 279–289.
31. B.S. Chahar, A.K.P. Siddhartha, Erosion wear of ductile materials: a review. in *Proceedings of the International Conference on Advancements and Recent Innovations in Mechanical, Production and Industrial Engineering (ARIMPIE—2016), ELK Asia Pacific Journal, Gaziabad, India* (2016), pp. 15–16
32. H. Arabnejad, A. Mansouri, S.A. Shirazi, B.S. McLaury, Evaluation of solid particle erosion equations and models for oil and gas industry applications. in *SPE Annual Technical Conference and Exhibition*. OnePetro (2015)
33. C. Syamsundar, D. Chatterjee, and M. Kamaraj, Experimental Characterization of Silt Erosion of 16Cr-5Ni Steels and Prediction Using Artificial Neural Network, *Trans. Indian Inst. Met.*, 2015, **68**(4), p 587–599.
34. M. Vishnoi, V. Muthupandi, S.S. Murugan, Characterization of hydrophobic coating developed by micro arc oxidation on AA2014 alloy (2017)
35. M. Vishnoi, P. Kumar and Q. Murtaza, Surface Texturing Techniques to Enhance Tribological Performance: A Review, *Surf. Interfaces*, 2021, **27**, p 101463.

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.

The Journal of Open Source Software

mlpack 4: a fast, header-only C++ machine learning library

Ryan R. Curtin [1], **Marcus Edel** [2], **Omar Shrit** [1], **Shubham Agrawal** [1], **Suryoday Basak** [3], **James J. Balamuta** [4], **Ryan Birmingham** [5], **Kartik Dutt** [6], **Dirk Eddelbuettel** [7], **Rishabh Garg** [8], **Shikhar Jaiswal** [9], **Aakash Kaushik** [1], **Sangyeon Kim** [10], **Anjishnu Mukherjee** [11], **Nanubala Gnana Sai** [1], **Nippun Sharma** [8], **Yashwant Singh Parihar** [12], **Roshan Swain** [1], **and Conrad Sanderson** [13]

1 Independent Researcher **2** Collabora Ltd **3** Pennsylvania State University **4** Departments of Statistics and Informatics, University of Illinois, Urbana-Champaign **5** Emory University **6** Delhi Technological University **7** Department of Statistics, University of Illinois, Urbana-Champaign **8** Indian Institute of Technology Mandi **9** Microsoft Research India **10** NAVER WEBTOON AI **11** George Mason University **12** Department of Computer Science and Engineering, IIT Bombay **13** Data61/CSIRO and Griffith University

DOI: 10.21105/joss.05026

Software

- Review
- Repository
- Archive

Editor: Øystein Sørensen

Reviewers:

- @sandeep-ps
- @zhangjy-ge

Submitted: 25 November 2022
Published: 01 February 2023

License
Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License (CC BY 4.0).

Summary

For over 15 years, the mlpack machine learning library has served as a "swiss army knife" for C++-based machine learning (Curtin et al., 2013). Its efficient implementations of common and cutting-edge machine learning algorithms have been used in a wide variety of scientific and industrial applications. This paper overviews mlpack 4, a significant upgrade over its predecessor (Curtin et al., 2018). The library has been significantly refactored and redesigned to facilitate an easier prototyping-to-deployment pipeline, including bindings to other languages (Python, Julia, R, Go, and the command line) that allow prototyping to be seamlessly performed in environments other than C++.

Statement of Need

The use of machine learning has become ubiquitous in almost every scientific discipline and countless commercial applications (Carleo et al., 2019; Jordan & Mitchell, 2015). There is one important commonality to virtually all of these applications: machine learning is often computationally intensive, due to the large number of parameters and large amounts of training data. This was the main motivator for the original development of mlpack in the C++ language, which allows for efficient close-to-the-metal implementations (Curtin et al., 2013).

But speed is not everything: development and deployment of applications that use machine learning can also be significantly hampered if the overall process is too difficult or unwieldy (Lavin et al., 2022; Paleyes et al., 2020). Furthermore, deployment environments often have computational or engineering constraints that make a full-stack Python solution infeasible (Fischer et al., 2020). As such, it is important that lightweight and easy-to-deploy machine learning solutions are available. This has motivated our refactoring and redesign of mlpack 4: we pair efficient implementations with easy and lightweight deployment, making mlpack suitable for a wide range of deployment environments. A more complete set of motivations can be found in the mlpack vision document (mlpack community, 2021).

mlpack is a general-purpose machine learning library, targeting both academic and commercial use; for instance, data scientists who need efficiency and ease of deployment, or, e.g., by

researchers who need flexibility and extensibility. While there are other machine learning libraries intended to be used from C++, many, such as FAISS (Johnson et al., 2019) and FLANN (Muja & Lowe, 2009), are limited to a few specific algorithms, instead of a full range of machine learning algorithms, like mlpack provides. dlib-ml (King, 2009), on the other hand, does provide a broad toolkit of machine learning algorithms, but its extensibility is somewhat limited as it does not use policy-based design (Alexandrescu, 2001) to provide arbitrary user-defined behavior, and the range of machine learning algorithms provided is smaller than mlpack's.

Functionality

The library contains a wide variety of machine learning algorithms, some of which are new to mlpack 4. The list of algorithms includes linear regression, logistic regression, random forests, furthest-neighbor search (Curtin & Gardner, 2016), accelerated k-means variants (Curtin, 2017), kernel density estimation (Lee & Gray, 2008), and fast max-kernel search (Curtin & Ram, 2014). There is also a module for deep neural networks, which has implementations of numerous layer types, activation functions, and reinforcement learning applications. Details of the available functionality are provided in the online mlpack documentation. The efficiency of these implementations has been shown in various works (Curtin et al., 2013; Fang & Chau, 2016) using mlpack's benchmarking system (Edel et al., 2014).

The algorithms are available via automatically-generated bindings to Python, R, Go, Julia, and the command line. Each of these bindings has a unified interface across the languages; for example, a model trained in Python can be used from Julia or C++ (or any other language with mlpack bindings). The bindings are available in each language's package manager, as well as system-level package managers such as `apt` and `dnf`. Furthermore, ready-to-use Docker containers with the environment fully configured are available on DockerHub, and an interactive C++ notebook interface via the xeus-cling project is available on BinderHub.

Once a user has developed a machine learning workflow in the language of their choice, deployment is straightforward. The mlpack library is now header-only, and directly depends only on three libraries: Armadillo (Sanderson & Curtin, 2016), ensmallen (Curtin et al., 2021), and cereal. When using C++, the only linking requirement is to an efficient implementation of BLAS and LAPACK (required via Armadillo). This significantly eases deployment; a standalone C++ application with only a BLAS/LAPACK dependency is easily deployable to many environments, including standard Linux-based Docker containers, Windows environments, and resource-constrained embedded environments. To this end, mlpack's build system now also contains a number of tools for cross-compilation support, including the ability to easily statically link compiled programs (important for some deployment environments).

Major Changes

Below we detail a few of the major changes present in mlpack 4. For a complete and exhaustive list (including numerous bug fixes and new techniques), the `HISTORY.md` file (distributed with mlpack) can be consulted.

Removed dependencies. In accordance with the vision document (mlpack community, 2021), the majority of the refactoring and redesign work focused on reducing dependencies and compilation overhead. This has motivated the replacement of the Boost C++ libraries, upon which mlpack previously depended, with lightweight alternatives including cereal for serialization. The entire neural network module was refactored to avoid the use of Boost (amounting to an almost complete rewrite). This effort was rewarded handsomely: with mlpack 3, a simple program would often require several gigabytes of memory just for compilation. After refactoring and removing dependencies, compilation generally requires just a few hundred megabytes of memory, and is often an order of magnitude faster.

Interactive notebook environments. mlpack can be used in a Jupyter notebook environment (Kluyver et al., 2016) via the xeus-cling project. This is demonstrated interactively on the mlpack homepage. Examples of C++ notebooks can be found in the mlpack examples repository, and these can easily be run on BinderHub.

New bindings and enhanced availability. Support for the Julia (Bezanson et al., 2017), Go (Pike, 2012), and R languages (R Core Team, 2022; Singh Parihar et al., 2022) has been added via mlpack's automatic binding system. These bindings can be used by installing mlpack from the language's package manager (`Pkg.jl`, `go get`, `install.packages('mlpack')`). Furthermore, since mlpack's reduced dependency footprint has significantly simplified the deployment process, mlpack's Python dependencies are now available for numerous architectures both on PyPI and in `conda-forge`.

Cross-compilation support and build system improvements. mlpack's build configuration now supports easy cross-compilation, for instance via toolchains such as buildroot. By specifying a few flags, a user may produce a working mlpack setup for a variety of embedded systems. This required the implementation of a dependency auto-downloader, which is capable of downloading OpenBLAS and compiling (if necessary) for the target architecture. The auto-downloader can also be enabled and used for any situation, thus easing installation and deployment.

Acknowledgements

Development of mlpack is community-led. It is the product of hard work by over 220 individuals (at the time of writing). We are also indebted to people that have provided bug reports over the years. The development has been supported by Google, via a decade-long participation the Google Summer of Code program, and also by NumFOCUS, which fiscally sponsors mlpack.

References

Alexandrescu, A. (2001). *Modern C++ Design: Generic programming and design patterns applied*. Addison-Wesley.

Bezanson, J., Edelman, A., Karpinski, S., & Shah, V. B. (2017). Julia: A fresh approach to numerical computing. *SIAM Review*, *59*(1), 65–98. https://doi.org/10.1137/141000671

Carleo, G., Cirac, I., Cranmer, K., Daudet, L., Schuld, M., Tishby, N., Vogt-Maranto, L., & Zdeborová, L. (2019). Machine learning and the physical sciences. *Reviews of Modern Physics*, *91*(4), 045002. https://doi.org/10.1103/RevModPhys.91.045002

Curtin, R. R. (2017). A dual-tree algorithm for fast k-means clustering with large k. *SIAM International Conference on Data Mining (SDM '17)*, 300–308. https://doi.org/10.1137/1.9781611974973.34

Curtin, R. R., Cline, J. R., Slagle, N. P., March, W. B., Ram, P., Mehta, N. A., & Gray, A. G. (2013). MLPACK: A scalable C++ machine learning library. *The Journal of Machine Learning Research*, *14*(1), 801–805.

Curtin, R. R., Edel, M., Lozhnikov, M., Mentekidis, Y., Ghaisas, S., & Zhang, S. (2018). Mlpack 3: A fast, flexible machine learning library. *Journal of Open Source Software*, *3*(26), 726. https://doi.org/10.21105/joss.00726

Curtin, R. R., Edel, M., Prabhu, R. G., Basak, S., Lou, Z., & Sanderson, C. (2021). The ensmallen library for flexible numerical optimization. *Journal of Machine Learning Research*, *22*, 1–6.

Curtin, R. R., & Gardner, A. B. (2016). Fast approximate furthest neighbors with data-dependent candidate selection. *International Conference on Similarity Search and Applications (SISAP 2016)*, 221–235. https://doi.org/10.1007/978-3-319-46759-7_17

Curtin, R. R., & Ram, P. (2014). Dual-tree fast exact max-kernel search. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, *7*(4), 229–253. https://doi.org/10.1002/sam.11218

Edel, M., Soni, A., & Curtin, R. R. (2014). An automatic benchmarking system. *NIPS Workshop on Software Engineering for Machine Learning*.

Fang, D., & Chau, D. H. (2016). M3: Scaling up machine learning via memory mapping. *International Conference on Management of Data*, 2249–2250.

Fischer, L., Ehrlinger, L., Geist, V., Ramler, R., Sobiezky, F., Zellinger, W., Brunner, D., Kumar, M., & Moser, B. (2020). AI System Engineering—Key Challenges and Lessons Learned. *Machine Learning and Knowledge Extraction*, *3*(1), 56–83. https://doi.org/10.3390/make3010004

Johnson, J., Douze, M., & Jégou, H. (2019). Billion-scale similarity search with GPUs. *IEEE Transactions on Big Data*, *7*(3), 535–547. https://doi.org/10.1109/tbdata.2019.2921572

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, *349*(6245), 255–260. https://doi.org/10.1126/science.aaa8415

King, D. E. (2009). Dlib-ml: A machine learning toolkit. *Journal of Machine Learning Research*, *10*, 1755–1758.

Kluyver, T., Ragan-Kelley, B., Pérez, F., Granger, B. E., Bussonnier, M., Frederic, J., Kelley, K., Hamrick, J. B., Grout, J., Corlay, S., & others. (2016). Jupyter Notebooks—a publishing format for reproducible computational workflows. *Proceedings of the 20th International Conference on Electronic Publishing*, 87–90. https://doi.org/10.3233/978-1-61499-649-1-87

Lavin, A., Gilligan-Lee, C. M., Visnjic, A., Ganju, S., Newman, D., Ganguly, S., Lange, D., Baydin, A. G., Sharma, A., Gibson, A., & others. (2022). Technology readiness levels for machine learning systems. *Nature Communications*, *13*(1), 1–19. https://doi.org/10.21203/rs.3.rs-133138/v1

Lee, D., & Gray, A. G. (2008). Fast high-dimensional kernel summations using the Monte Carlo Multipole Method. *Advances in Neural Information Processing Systems*, *21*, 929–936.

mlpack community. (2021). *mlpack: A vision for an efficient prototype-to-deployment machine learning library*. https://www.mlpack.org/papers/vision.pdf.

Muja, M., & Lowe, D. G. (2009). Fast approximate nearest neighbors with automatic algorithm configuration. *Proceedings of the 2009 International Conference on Computer Vision Theory and Applications (VISAPP '09)*, 331–340. https://doi.org/10.5220/0001787803310340

Paleyes, A., Urma, R.-G., & Lawrence, N. D. (2020). Challenges in deploying machine learning: A survey of case studies. *ACM Computing Surveys (CSUR)*. https://doi.org/10.1145/3533378

Pike, R. (2012). Go at Google: Language design in the service of software engineering. *Proceedings of the 3rd Annual Conference on Systems, Programming, and Applications: Software for Humanity*, 5–6.

R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Sanderson, C., & Curtin, R. R. (2016). Armadillo: A template-based C++ library for linear algebra. *Journal of Open Source Software*, *1*(2), 26. https://doi.org/10.21105/joss.00026

Singh Parihar, Y., Curtin, R. R., Eddelbuettel, D., Balamuta, J., & others. (2022). *mlpack: "Rcpp" integration for the "mlpack" library*. https://CRAN.R-project.org/package=mlpack

Multiple Feature Selection Frameworks based on Evolutionary Computing and Ensemble Learning for Software Defect Prediction

Ruchika Malhotra
Department of Software Engineering
Delhi Technological University
Delhi, India
ruchikamalhotra@dtu.ac.in

Anmol Budhiraja
Department of Computer Science and Engineering
Delhi Technological University
Delhi, India
anmolbudhiraja1999@gmail.com

Abhinav Kumar Singh
Department of Computer Science and Engineering
Delhi Technological University
Delhi, India
abhinavsngh26@gmail.com

Ishani Ghoshal
Department of Computer Science and Engineering
Delhi Technological University
Delhi, India
ishani.ghoshal11@gmail.com

Shweta Meena
Department of Software Engineering
Delhi Technological University
Delhi, India
shwetameena@dtu.ac.in

2022 5th International Conference on Computational Intelligence and Networks (CINE) | 978-1-6654-6465-9/22/$31.00 ©2022 IEEE | DOI: 10.1109/CINE56307.2022.10037473

***Abstract*—Software defects may cause severe crashes in the system, leading to the software's high maintenance costs. Early identification of these defects would lead to high-quality software, thus saving time and money. This study proposes five feature selection approaches based on evolutionary computing algorithms, each coupled with a majority voting ensemble for Software defect prediction. The objective is to improve the existing process by targeting the metric selection stage. The study was conducted on thirty open-source defect datasets. The proposed feature selection techniques were applied on a within-project defect prediction model and a heterogeneous defect prediction model. The Friedman and the Wilcoxon Signed-rank test concluded that the proposed techniques were promising and generated results comparable to some other state-of-the-art feature selection methodologies.**

Keywords—Defect Prediction, Evolutionary Computing, Artificial Intelligence

I. INTRODUCTION

A software defect arises when the actual results of a software program are different from the expected results. This happens when there is a bug, fault, or error in the code. Most software defects are due to human error (programming faults of the developer). An error-prone software would eventually lead to an increase in the maintenance cost and the error correction. Therefore, it is better to predict the software defects beforehand to handle the flaws before the code is in production. This is where we use Software Defect Prediction (SDP). Many models for defect prediction are used nowadays. One of the most common models is Within Project Defect Prediction wherein the data is partitioned into training set and the testing set, and it is all taken from within-the-project, hence the name. However, for WPDP, it is necessary that the previous defect data from the early models of the project are stored. It fails when the project is in itself the first release or when previous log reports of errors have not been maintained.

Another model which covers a broader set of projects is Cross Project Defect Prediction (CPDP). In this case, an external project serves as the set used for training.The limitation with CPDP is that the projects must have the same metrics and this limitation is overcome by using heterogeneous defect prediction model (HDP) that uses CPDP with projects having different dataset metrics.

The aim of selecting features in such a software defect dataset is to remove undesired information not closely interrelated with the defect proneness of the software, which in turns helps in the reduction of dimensionality of the dataset, improving the training and testing time for the prediction model, and boosting the performance measures. Swarm Intelligence is one the techniques used to optimally reduce this dimensionality of a dataset. It attempts to imitate the collective behaviours of naturally occurring species such as swarms of bees, cuckoos, bats, etc., in order to find optimal strategies for problems [1]-[5].

In this paper, our objective is to present effective feature selection frameworks for SDP based on evolutionary computing algorithms and ensemble learning which will in turn aid in a more precise defect prediction in huge softwares. We propose five feature selection methods, namely, en-BPSO (based on Binary Particle Swarm Optimisation), en-BGA (based on Binary Genetic Algorithm), en-BDFA (based on Binary DragonFly Algorithm), en-BBA (based on Binary Bat Algorithm) and en-BGSA (based on Binary Gravitational Search Algorithm). All the five feature selection approaches use a majority voting-based ensemble classifier-based fitness function. The ensemble classifier uses five popular ML techniques in its constitution. We address the following research questions through our study:

1) RQ1: How do the proposed feature selection approaches perform in WPDP?
2) RQ2: How do the proposed feature selection approaches perform in HDP?

The en-BPSO technique was proposed in [6], however, this study expands the existing study with additional datasets and more comprehensive analysis of results. The main contributions of this study are as follows:

- The current study proposes four more feature selection methods, namely, en-BGA, en-BDFA, en-BBA, and en-BGSA
- This study uses thirty datasets compared to eight in the existing study, thereby enhancing the generalisability of the results.
- The existing study only evaluates the effectiveness of en-BPSO on HDP problems; however, this study has

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on February 16,2023 at 06:23:59 UTC from IEEE Xplore. Restrictions apply.

Fig. 1. Architecture of Proposed Methodology: The HDP model used in the study is [9]. Majority Voting Ensemble Classifier is [6]

been extended to show how the proposed feature selection methods perform on WPDP problems.

- The analysis on HDP problems has also been extended, and the statistical analysis of the proposed techniques for feature extraction has been carried out using the Friedman Test and the Wilcoxon signed-rank Test.

The structure of the paper is as follows. Section 2 discusses the existing related work in the literature. In Section 3, the mechanism of the proposed techniques is detailed. Section 4 discusses the datasets utilised in the study and the experimental design. Section 5 presents the results for each RQ. At last, the threats to validity and the conclusion are highlighted in Sections 6 and 7.

II. Related Work

Various different models for prediction of defects have been utilised in the past. Much research has been done on CPDP and HDP models to get the required results compared to WPDP. Watanabe et al. suggested the reuse of generated metrics of a given project to other target projects [7]. For this, they followed the metric compensation model wherein the target metrics were calculated using the average metrics of the sourceNam et al. introduced TCA+, an enhancement of the Transfer Component Analysis (TCA) approach [8], which transfers data with less difference into a single latent space. TCA+ extended TCA by the addition of normalization pre-processing. Following TCA+, The state-of-the-art model proposed by Nam et al. used transfer learning by applying the maximum-weighted-bipartite matching technique to match the match the source dataset's metrics, and the target dataset's metrics [9].

Son et al. have analyzed various techniques used in defect prediction coupled with numerous optimization techniques (Particle swarm optimization, Ant Colony Optimization, Cuckoo search, Genetic algorithm) to create a mapping of the most suitable techniques [10]. Reference [11] employed feature selection based on evolutionary algorithms to remove the noisy attributes from the dataset. [12] employs cuckoo search algorithm for software defect prediction.

Even though many studies have been conducted on proposing techniques for feature selection based on evolutionary algorithms, no other study, to the the authors' knowledge, has focused on applying evolutionary algorithms in the field of HDP. Also, even though PSO-based hybrid algorithms have been proposed for WPDP, the applicability of some other evolutionary algorithms in this field, especially for feature selection, remains very limited. Thus, we conduct a thorough analysis on thirty datasets to see how evolutionary algorithms, coupled with an ensemble learning classifier, perform in the field of defect prediction, and whether or not, they can further improve the existing SDP techniques.

III. Methodology

In this study, we come up with five feature selection frameworks, each based on evolutionary computing algorithms and ensemble learning. Each of the five feature selection frameworks is built around a majority voting ensemble classifier and uses the same fitness function. The proposed methodology's architecture is presented in Fig. 1.

A. Binary Evolutionary Algorithms

We have used the famous binary implementations of Genetic Algorithm, Bat Algorithm, Dragon Fly Algorithm, Gravitational Search Algorithm and Particle Swarm Optimisation [1]-[5]. Genetic Algorithms (GA) are inspired by natural selection, where we select the individual, which is fittest to carry the reproduction of the next offspring. It performs optimization by using binary-encoded string variables [1]. The Bat algorithm was given in [3]. Bats use sound waves to find the location of objects in their way. It uses a frequency-tuning algorithm to make the solution set more diverse while also using automatic zooming to mimic the variation of the echo that the bats make. The Dragon Fly Algorithm is a swarm optimization algorithm modelled on the coordinated behaviour of Dragonflies [5]. It uses five

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on February 16,2023 at 06:23:59 UTC from IEEE Xplore. Restrictions apply.

parameters to model this, namely Cohesion, Alignment, Separation, Attraction towards a food source and Distraction outward from an enemy. An improved gravitational search algorithm was given in [4]. The suggested algorithm looked at improving the classification technique and transfer function to overcome the stagnation situation. The algorithm is derived from Newton's Laws of Gravitation. Kennedy and Eberhart [2] originally suggested the employing of Particle Swarm Optimization. We start with some random solutions initially. Each random solution, called a particle, is identified by a position vector and a velocity vector. Every particle in the search space moves based on the best position achieved both individually (local best position) and collectively (global best position) when taken all particles together.

In all the Binary Evolutionary Algorithms, the representation is a binary string, and a 0 in the feature mask means that this feature is not picked and 1 that it is picked. The velocity in BPSO defines the probability of switching of polarity of bit in the position vector.

B. *Fitness Function based on Majority Voting Ensemble Classifier*

Making an ensemble classifier using diverse classifiers helps us overcome some flaws of individual classifiers. The fitness function calculates an individual's fitness, and which individuals are better than the others. In our context, every individual is a binary string of 0 and 1. So, if individual A, is better than individual B, it means the features selected by A are better than the those selected by B.

Reference [6] proposes a framework of the fitness function based on MVEC was proposed. The MVEC is based on five diverse classifiers - Random Forests, Naive Bayes, Logistic Regression, Gradient Boosting, and KNN classifier. The binary string is used as a feature mask to select the features from the dataset. Stratified k-fold cross-validation (k=4) is adopted to distribute the complete set of data into training and testing.

All the proposed techniques, i.e., en-BPSO, en-BGA, en-BDFA, en-BBA, and en-BGSA, use the same fitness function. This helps us make a fair comparison of all the proposed techniques.

IV. Experimental Framework

A. *Datasets*

Five groups consisting of 30 open-source defect datasets are used in this study. These five groups are AEEEM, ReLink, MORPH, NASA, and SOFTLAB [13] - [15]. Table 1 describes the datasets used in the study.

There are 61 metrics in each dataset of the AEEEM group, including metrics for churn-of-source-code, change and code entropy, object-oriented metrics (OO), and previous-defect metrics. Wu et al. used the ReLink datasets by increasing the defect data quality [14]. The open-source project datasets in MORPH group consist of twenty metrics that include McCabe's cyclomatic metrics, CK metrics, with some other OO metrics [15]. The NASA dataset, as the name suggests, has proprietary datasets from NASA [13]. Whereas SOFTLAB is from a Turkish software company. Every dataset in the NASA group of datasets describes a system or a sub-system of NASA

TABLE I. Dataset Description

Group	Dataset	# Metrics	# Instances	# Buggy Instances
AEEEM	EQ	61	324	129 (39.81%)
	JDT		997	206 (20.66%)
	LC		691	64 (9.26%)
	ML		1862	245 (13.16%)
	PDE		1492	209 (14.01%)
ReLink	Apache	26	194	98 (50.52%)
	Safe		56	22 (39.29%)
	ZXing		399	118 (29.57%)
MORPH	ant-1.3	20	125	20 (16.00%)
	arc		234	27 (11.54%)
	camel-1.0		339	13 (3.83%)
	poi-1.5		237	141 (59.49%)
	redaktor		176	27 (15.34%)
	skarbonka		45	9 (20.00%)
	tomcat		858	77 (8.97%)
	velocity-1.4		196	147 (75.00%)
	xalan-2.4		723	110 (15.21%)
	xerces-1.2		440	71 (16.14%)
NASA	CM1	37	344	42 (12.21%)
	MW1		264	27 (10.23%)
	PC1		759	61 (8.04%)
	PC3		1125	140 (12.44%)
	PC4		1399	178 (12.72%)
	MC2	39	127	44 (34.65%)
	KC3		200	36 (18.00%)
SOFTLAB	ar1	29	121	9 (7.44%)
	ar3		63	8 (12.70%)
	ar4		107	20 (18.69%)
	ar5		36	8 (22.22%)
	ar6		101	15 (14.85%)

software, and it contains a variety of static code metrics such as readability, size, complexity, etc, and defects marking data. MC2 and KC3 have some metrics different from the rest of the datasets in the group. The projects used from SOFTLAB were obtained from the Promise repository [16]. Halstead and McCabe's cyclomatic metrics are both used in these two groups. however, NASA has some extra complexity metrics like the percentage of comments and parameter count. Having utilized the proposed methods to select features from the source dataset, we then employ this dataset to train the HDP and WPDP model.

B. *Experimental Design*

For quantifying the impact of the proposed feature selection techniques on HDP, we implement the model proposed in [9]. The feature selection is applied only on the source set, and the target set remains as it is. The Logistic Regression (LR) classifier is used here for the purpose of building the HDP and WPDP models. Nam *et al.* and Ghotra *et al.*, in their study, indicated that Logistic based classifiers lead to a better prediction model than other classifiers SDP [9, 17]. After training the data using the source dataset, the predictions are made for the matched target dataset.

Using Stratified 2-fold cross-validation, the dataset (after determining the features using each technique) is split up into the training and testing sets to build model for WPDP. We have used the Logistic Regression classifier here as well, for

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on February 16,2023 at 06:23:59 UTC from IEEE Xplore. Restrictions apply.

the reasons outlined previously, and to ensure that the findings are fair.

C. Performance Measure

We utilise the area under the ROC curve as the performance metric to compare the efficacy of each of the presented strategies. The ROC curve helps us identify how accurate our model is in making the right defect predictions. The ROC curve is plotted with the True Positive Rate (TPR) and the False Positive Rate (FPR) taken on the Y and X axes respectively.

D. Statistical Tests

To statistically evaluate the results, the Friedman Test as well as the Wilcoxon Signed Rank Test are used. [18, 19].

- **Friedman Test:** We use Friedman test to check whether at least two, or more of the results produced, do not give similar medians [18]. It is a non-parametric test that does not require a normality assumption for the results. We purpose following hypotheses:
    - H_0: All methods have the same median.
    - H_1: At least two or three methods have different median.

  We get the test's *p*-value, which we compare with a widely used significance level of 0.05. Based on the comparison we accept or reject the null hypotheses.

  In addition to Friedman's test, we also use Friedman's mean rank to order the various feature selection techniques according to their AUC values.
- **Wilcoxon Signed-Rank Test:** This statistical test, given by Wilcoxon [19], is applied after the Friedman test to check which feature selection method produces statistically different results than when none is used. We check the *p*-value produced from the test at a widely accepted level of significance of 0.05.We propose the null and alternate Hypotheses as follows:
    - H_0: The median of the results is unaffected by feature selection.
    - H_1: Using feature selection affects the median of results.

E. Baselines

We compare our proposed feature selection techniques to three baselines. The first baseline is when we don't apply any feature selection technique and use all the dataset features. This will help us analyze whether there is any need for selecting features or not and verify if we are eliminating redundant features using the proposed techniques. The second baseline is ReliefF feature selection technique. The ReliefF was proposed in [20]. It is a filter-based feature selection technique, but unlike other filter-based algorithms, ReliefF can also detect feature dependencies. As the count of features to be selected need to be specified for feature selection using ReliefF, we select the best 15% features [9]. As the third baseline, we use an embedded feature selection methodology - L1 penalty-based Logistic Regression (LR-L1).

Our aim of comparison with the second and the third baseline is to determine how our proposed techniques perform compared to some other popular feature selection techniques. Thus, we have tried to take two diverse baselines: one being a filter method and the other being an embedded method.

F. Parameter Settings

For every evolutionary algorithm used in the study, we have used 25 agents, with the number of iterations specified as 200. All the parameters have been empirically defined. Table 2 lists the specific parameters for each algorithm.The implementation of all the proposed techniques has been done in python.

TABLE II. PARAMETERS OF EACH TECHNIQUE USED IN THE STUDY

en-BPSO	ω =0.5,$c1$=2,$c2$=2 and V_{max}=4
en-BGA	the probability of mutation to take place = 0.05
en-BDFA	-
en-BGSA	K =25,G_0=1
en-BBA	α =0.9,γ =0.9

V. EXPERIMENTAL RESULTS

This section will be presenting the findings from our paper's proposed techniques. To reduce randomness and variance in results, we have calculated our result 20 times, and the median of those results is presented in the paper. We have chosen AUC as the measure for the performance comparison of the described feature selection methodologies. Friedman's test was followed by Wilcoxon test to check which feature selection techniques do not affect the result much. To compare the overall performance of different methods, Friedman's mean rank, median, and mean of AUC is used.

A. *RQ1: How do the proposed feature selection approaches perform in WPDP.*

We apply the Friedman test which helps us to verify whether the results presented are statistically different or not. The Friedman test results in a *p*-value of 2.7×10^{-7} with a test statistics value of 45.72. This shows that all the methods do not produce similar results. Friedman mean rank of all the feature selection techniques is presented in Table 3.

TABLE III. RESULT SUMMARY OF VARIOUS FEATURE SELECTION TECHNIQUES FOR WPDP

Technique	Mean	Median	Friedman's Mean Rank
None	0.671	0.688	3.017
ReliefF	0.62	0.627	2.977
LR-L1	0.685	0.69	4.233
en-BPSO	0.706	0.712	4.867
en-BDFA	0.721	0.73	5.567
en-BGSA	0.7	0.718	4.9
en-BBA	0.727	0.736	6.05
en-BGA	0.702	0.705	4.4

The en-BBA strategy clearly leads in terms of the Friedman mean rank, and en-BDFA closely follows it. Both methods utilize an evolutionary algorithm for feature selection. Also, on a careful observation, it can be seen that when we use

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on February 16,2023 at 06:23:59 UTC from IEEE Xplore. Restrictions apply.

ReliefF for feature selection, then we get the lowest mean rank, even lower than the case when no feature selection was made. Table 3 also presents the mean and median of all the values. Fig. 2 presents the overall spread of results, with mean and median. We see a similar trend here as en-BBA performs best both in terms of the mean as well as the median and ReliefF provides the lowest values of all the methods considered. For every technique, we apply the Wilcoxon signed-rank pairwise with the 'None' baseline to check which of the techniques produced statistically different results.

Fig. 2. Graphical Summary of Results of Various Feature Selection Techniques for WPDP

From the results presented in Table 4, we can see that except for en-BPSO, all other methods have a *p*-value less than 0.05. So the null hypotheses can be rejected for all the methods except for en-BPSO method. We can conclude, that most of the proposed techniques show promise in terms of WPDP predictions.

TABLE IV. WILCOXON SIGNED-RANK TEST FOR PAIRWISE ANALYSIS WITH 'NONE' FOR WPDP

Technique	Test Statistic	p-value
ReliefF	131	**3.68×10^{-2}**
LR-L1	75.5	**1.24×10^{-3}**
en-BPSO	140	5.71×10^{-2}
en-BDFA	87.5	**2.86×10^{-3}**
en-BGSA	105	**1.50×10^{-2}**
en-BBA	68	**7.15×10^{-4}**
en-BGA	134	**4.28×10^{-2}**

B. *RQ2: How do the proposed feature selection approaches perform in HDP*

Friedman test is applied to the results to check whether there is at least one pair of statistically different results. Friedman's test yields a *p*-value of 4.7×10^{-8}, with test statistics of 49.68, which implies that the results are different from each other. Friedman mean rank, along with mean and median of the different methods is presented in Table 5.

The en-BBA method performs best in terms of Friedman mean rank and is followed by the en-BGSA and the en-BDFA methods. We see the same trend with respect to ReliefF and it gave the poorest results here as well. Each of the remaining feature selection techniques outperformed the 'None' baseline, indicating that feature selection techniques improve the results for HDP.

On close analysis of the median and mean values and Fig. 3, we conclude that en-BPSO provides the highest median value, and Relief-F provides the lowest median value. The en-BGSA provides the highest mean values, and again, the Relief-F provides the lowest mean values.

TABLE V. WILCOXON SIGNED-RANK TEST FOR PAIRWISE ANALYSIS WITH 'NONE' FOR HDP

Technique	Test Statistic	p-value
Relief-F	31	**3.40×10^{-5}**
LR-L1	154	2.64×10^{-1}
en-BPSO	116	**2.82×10^{-2}**
en-BDFA	109.5	**1.14×10^{-2}**
en-BGSA	92.5	**6.87×10^{-3}**
en-BBA	130	**3.50×10^{-2}**
en-BGA	188.5	3.65×10^{-1}

Fig. 3. Graphical Summary of Results of Various Feature Selection Techniques for HDP

For further analysis, Wilcoxon signed-rank test was applied to check which methods perform statistically different than None. From the results presented in Table 6, we can see that except for LR-L1 and en-BGA, the *p*-value for all other methods is less than 0.05. So we can reject the null hypotheses for all the methods except for LR-L1 and the en-BGA technique.

TABLE VI. RESULT SUMMARY OF VARIOUS FEATURE SELECTION TECHNIQUES FOR HDP

Technique	Mean	Median	Freidman's Mean Rank
None	0.612	0.626	3.7
ReliefF	0.558	0.549	2.133
LR-L1	0.617	0.623	4.433
en-BPSO	0.632	0.645	5.233
en-BDFA	0.629	0.637	5.366
en-BGSA	0.636	0.634	5.483
en-BBA	0.631	0.641	5.533
en-BGA	0.621	0.628	4.316

One of the reasons for the evolutionary computing algorithms performing better than other feature selection methods can be that evolutionary algorithms learn about the data and then determine the required count of the features that need to be selected. In contrast, for methods like ReliefF, we need to provide the count of features that need to be selected. Not knowing the optimum count can affect the model performance. When we have fewer features than the optimum number, we would not have enough information to make the correct prediction. In contrast, when we have more than

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on February 16,2023 at 06:23:59 UTC from IEEE Xplore. Restrictions apply.

optimum, we risk adding white noise in the data. The model based method LR-L1, manages to still outperform some proposed classifiers in some cases. However, at an average, the proposed feature selection techniques were significantly better than the baselines as indicated by the results in both HDP and WPDP. In case of both WPDP and HDP, en-BBA gave the best results, also surpassing the en-BPSO technique proposed by the authors in [6]. This indicates that en-BBA has the highest potential to determine the correct count of features for best AUC results. One of the reasons for en-BBA producing good results in both the RQs is that the binary bat algorithm balances both the emmision rate and pulse to evade the local maximum and quickly converge to the global maximum. en-BDFA also provided promising results in both the scenarios, the reason being the five parameters used in the Binary Dragon Fly Algorithm, tuning which enable a good exploration as well as exploitation of the search space.

VI. Conclusion

We proposed five feature selection architectures for defect prediction in this research. The developed techniques were based on popular evolutionary algorithms and used a fitness function that used five diverse classifiers in a majority voting ensemble. The efficacy of the proposed techniques was validated using an HDP model as well as a WPDP model on 30 open-source defect datasets. We used popular feature selection techniques to compare and analyze the proposed techniques. The Friedman test was used to conduct the statistical analysis, which was followed by the Wilcoxon Signed Rank Test. Based on the AUC values, the following results were noted:

- Feature selection helps improve the model's performance in most cases. Only when we use ReliefF for feature selection, we see a downfall in the model's performance.
- Friedman Test gave a *p*-value less than 0.05, in both HDP and WPDP, indicating that all the results were statistically different to each other
- With respect to the Friedman mean rank, the mean and median AUC values for WPDP, the en-BBA approach is superior to the other feature selection methods.
- In case of HDP, the en-BBA technique provided the best results based on the Friedman Mean Rank and mean AUC, and the en-BPSO technique provided the best result based on median AUC.

Overall, the proposed feature selection techniques based on evolutionary computing coupled with the ensemble classifier, provide promising results in both WPDP and HDP scenarios, providing a significant improvement in the accuracy with which the defects can be predicted.

The validity of our experiment findings may be threatened by a number of factors. We use thirty widely used open-source software defect datasets to achieve better generalization. To further generalise the result, however, additional research should assess the experiment using more datasets. Biased measures can influence the result by wrongly assessing the relation between independent and dependent variables leading to wrong conclusions. We use the Friedman test as well as the Wilcoxon test, which do not depend on parent distribution and parameters.

References

[1] K. F. Man, K. S. Tang, and S. Kwong, "Genetic algorithms: concepts and applications [in engineering design]," *IEEE transactions on Industrial Electronics*, vol, 43, no. 5*4*, pp.519-534, 1996

[2] J. Kennedy, and R. Eberhart. "Particle swarm optimization," In *Proceedings of ICNN'95-international conference on neural networks*, vol. 4, pp. 1942-1948. IEEE, 1995.

[3] S. Mirjalili, S. Mirjalili and X. Yang, "Binary bat algorithm," *Neural Computing and Applications*, vol. 25, no. 3-4, pp.663-681, 2013

[4] E. Rashedi and H. Nezamabadi-Pour, "Feature subset selection using improved binary gravitational search algorithm," *Journal of Intelligent & Fuzzy Systems*, vol. *26*, no. 3, pp.1211-1221. 2014

[5] M.M. Mafarja, D. Eleyan, I. Jaber, A. Hammouri, and S. Mirjalili, "Binary dragonfly algorithm for feature selection," In *2017 International conference on new trends in computing sciences (ICTCS)*, pp. 12-17 , October 2017

[6] R. Malhotra, A. Budhiraja, A. K. Singh, and I. Ghoshal, "A Novel Feature Selection Approach based on Binary Particle Swarm Optimization and Ensemble Learning for Heterogeneous Defect Prediction," In *2021 3rd Asia Pacific Information Technology Conference*, pp. 115-121, January 2021

[7] S. Watanabe, H. Kaiya and K. Kaijiri, "Adapting a fault prediction model to allow inter languagereuse," In *Proceedings of the 4th international workshop on Predictor models in software engineering*, pp. 19-24, May 2008

[8] J. Nam, S. J. Pan and S.Kim, "Transfer defect learning," In *2013 35th international conference on software engineering (ICSE)*, pp. 382-391, IEEE. May 2013

[9] J. Nam, W. Fu, S. Kim, T. Menzies, and L. Tan, "Heterogeneous defect prediction," *IEEE Transactions on Software Engineering*, vol. 44, pp. 874-896. 2017

[10] L.H Son, N. Pritam, M. Khari, R. Kumar, P.T.M Phuong, and P.H. Thong,"Empirical study of software defect prediction: a systematic mapping". *Symmetry*, vol. *11*, no. 2, p.212. 2019

[11] S.I. Ayon, "Neural network based software defect prediction using genetic algorithm and particle swarm optimization." In *2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT)*, pp. 1-4. IEEE, 2019.

[12] D. Badvath, A. safali Miriyala, S. chaitanya Gunupudi, and P. venkata Kuricheti, "Prediction of software defects using Deep Learning with improved cuckoo search algorithm," *Concurrency and Computation: Practice and Experience*, vol. 34, no. 26, 2022.

[13] M. Shepperd, Q. Song, Z. Sun, and C. Mair, "Data quality: Some comments on the nasa software defect datasets". *IEEE Transactions on Software Engineering*, vol *39*, no. 9, pp.1208-1215. 2013

[14] R. Wu, H. Zhang, , S. Kim and S. C. Cheung, "Relink: recovering links between bugs and changes". In *Proceedings of the 19th ACM SIGSOFT symposium and the 13th European conference on Foundations of software engineering*, pp. 15-25. September 2011

[15] F. Peters and T. Menzies, "Privacy and utility for defect prediction: Experiments with morph". In *2012 34th International conference on software engineering (ICSE)*, pp. 189-199, June 2012

[16] Promise.site.uottawa.ca. "*PROMISE DATASETS PAGE*." [online] Available at: <http://promise.site.uottawa.ca/SERepository/datasets-page.html>.

[17] K. Gao, T.M. Khoshgoftaar, H. Wang, and N. Seliya, "Choosing software metrics for defect prediction: an investigation on feature selection techniques." *Software: Practice and Experience*, vol. *41*, no. 5, pp.579-606, 2011

[18] M. Friedman, "The use of ranks to avoid the assumption of normality implicit in the analysis of variance". *Journal of the american statistical association*, vol *32*, no. 200, pp.675-701, 1937

[19] F. Wilcoxon. "Individual comparisons by ranking methods," In *Breakthroughs in statistics*, pp.196–202. Springer, 1992.

[20] I. Kononenko. "Estimating attributes: Analysis and extensions of RELIEF." In *European conference on machine learning*, Springer, Berlin, Heidelberg, pp. 171-182, 1994.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on February 16,2023 at 06:23:59 UTC from IEEE Xplore. Restrictions apply.

Materials Today Communications 35 (2023) 105635

Contents lists available at ScienceDirect

Materials Today Communications

journal homepage: www.elsevier.com/locate/mtcomm

Nano ZnO embedded poly (ethylene glycol) diacrylate cross-linked carboxymethyl tamarind kernel gum (CMTKG)/poly (sodium acrylate) composite hydrogels for oral delivery of ciprofloxacin drug and their antibacterial properties

Indu Rani, Sudhir G. Warkar *, Anil Kumar *

Department of Applied Chemistry, Delhi Technological University, Delhi 42, India

ARTICLE INFO

Keywords:
CMTKG
ZnO NPs
ZnO loaded hydrogels
Ciprofloxacin
Drug delivery
Antimicrobials studies

ABSTRACT

In this research work, a novel hydrogel network based on carboxymethyl tamarind kernel gum/poly (sodium acrylate) was synthesized by using poly (ethylene glycol) diacrylate (PEGDA) as a cross-linker. Zinc Oxide nanoparticles (ZnO NPs) were prepared via the hydrothermal synthetic method and developed ZnO NPs embedded within CMTKG/Poly (sodium acrylate) hydrogel for the controlled release studies of ciprofloxacin drug. Various techniques such as FTIR, XRD, FESEM, and TEM were used to characterize the synthesized ZnO NPs, pure hydrogel, and hydrogel nanocomposites. Various parameters such as drug loading (DL %), drug entrapment (DE %), gel content, and porosity were estimated for all the synthesized hydrogel nanocomposites. The results of swelling and rheological studies of hydrogel nanocomposites concluded that the embedded ZnO NPs increased hydrogel's swelling and thermal stability. The water absorption data were analyzed using the Power and Schott model. The result concluded that the Schott function model fitted the dynamic swelling data. The antibacterial action of CMTKG-based hydrogel nanocomposites was studied using *E. coli* (gram-negative) bacteria with the help of the disc diffusion method. The result showed that the incorporation of ZnO NPs enhanced the antimicrobial action of ciprofloxacin-loaded CMTKG-based hydrogels. The kinetic modelling of drug release was done using Higuchi and Korsmeyer - Peppas model. The higher value of regression coefficient (R^2) close to unity indicated that the mechanistic pathway of drug release from the hydrogels was more fitted in the Korsmeyer-Peppas model followed by Fickian diffusion.

1. Introduction

Hydrogels are cross-linked polymeric systems made up of synthetic monomers and natural polymers such as polysaccharides [1]. Hydrogels have a significant amount of porosity, making them efficient materials to absorb drugs in the aqueous mediums via physical or chemical interactions and subsequent release of the drug under physiological conditions [1]. Several stimuli like pH, temperature, and ionic strength affect drug release rate from the hydrogel. The drug released depends on the water uptake capacity of the hydrogels and the mechanism of diffusion of fluid into the matrix of hydrogel [2]. Other than the drug delivery applications, hydrogels can be used in other fields such as agriculture, metal ion sensing, heavy metal ion removal, removal of toxic dyes, and the food industry [3]. Till now, a large number of natural biopolymer-based hydrogels such as carboxymethyl guar gum [4], chitosan [5], tamarind gum [6], carboxymethyl cellulose [7] were used for the controlled release of the drug to the human body but carboxymethyl tamarind kernel gum (CMTKG) based hydrogels have not been much explored. CMTKG is a natural, non-toxic, biodegradable, and biocompatible biopolymer that is a chemically modified form of tamarind gum. It contains the polysaccharide units of D-Xylose, D-Glucose, and D-Galactose [8]. This chemical modification resulted in an enhanced life shell, swelling capacity, and high drug loading capacity [9]. There are only a few numbers of CMTKG-based hydrogel reported for biomedical applications for example Mali et.al., group reported the formation of CMTKG-based hydrogel film cross-linked via citric acid [10]. Alginate and CMTKG hydrogels were synthesized for the delivery of acyclovir as reported earlier [11]. CMTKG-based nanoparticles were synthesized for

* Corresponding authors.
E-mail addresses: sudhirwarkar@dtu.ac.in (S.G. Warkar), anil_kumar@dce.ac.in (A. Kumar).

https://doi.org/10.1016/j.mtcomm.2023.105635
Received 17 June 2022; Received in revised form 10 January 2023; Accepted 13 February 2023
Available online 14 February 2023
2352-4928/© 2023 Elsevier Ltd. All rights reserved.

Table 1
Composition of CMTKG/ Poly (sodium acrylate) hydrogel nanocomposite.

Sample Code	CMTKG (g/10 mL)	AA (mL)	PEGDA (mL)	KPS (g)	ZnO NPs (g)
E1	0.15	5	2	0.025	0
E2	0.15	5	2	0.025	0.03
E3	0.15	5	2	0.025	0.05
E4	0.15	5	2	0.025	0.1

the ocular delivery of tropicamide [12]. Over the past few years, the incorporation of inorganic materials such as nanoparticles in hydrogel has greatly increased. This hydrogel nanocomposite has better thermal and mechanical strength. Bio-composites are formed by the combination of natural polymers with inorganic materials such as ZnO [13], Fe_2O_3 [14], CuO [15], and Ag nanoparticles [16]. Zinc oxide nanoparticles are non-toxic, antibacterial, and eco-friendly [17]. Acrylic acid/PEG/ZnO mucoadhesive nanocomposite was synthesized for the delivery of propranolol HCl [18]. Green synthesis of carboxymethyl cellulose/ZnO nanocomposite hydrogel was reported earlier [19]. Ciprofloxacin is an antibiotic, broad-spectrum, synthetic, chemotherapeutic agent belonging to fluoroquinolones [20]. The main reason to use the ciprofloxacin drug is that it is a superior antibiotic drug to others. The half-life of this drug is 4 h. It is a broad-spectrum antibiotic thus, infections caused by gram negative and positive bacteria can be treated by this drug [21]. Chitosan and hydroxyapatite hydrogel nanocomposite cross-linked with k-carrageenan have been reported for the delivery of the sustained release of ciprofloxacin drug [22]. Poly (ethylene glycol) (PEG) based polymer can be easily synthesized and has great biocompatibility. PEGDA has a C = C double bond at each side of the PEG chain, so it can be used as a cross-linking agent by undergoing free radical polymerization methods with potassium persulfate as an initiator [23]. Only a few numbers of PEGDA cross-linked hydrogels have been reported to date. Hydrogels can also be fabricated from synthetic organic monomers. Out of these, acrylic acid is synthetic, anionic monomer, soluble in water, has a shorter lifetime, and contains an ionizable acidic -COOH group sensitive to various stimuli such as temperature and pH [3]. Here, in this work, CMTKG/Poly (sodium acrylate) hydrogels were synthesized using a PEGDA cross-linker. In this hydrogel system, zinc oxide nanoparticles were added to the hydrogel to enhance its properties such as drug loading, drug entrapment, biocompatibility, and a controlled slow release of the drug. The synthesized ZnO NPs and CMTKG-based hydrogel nanocomposite were characterized using FTIR, XRD, FE-SEM, TEM, and UV-Visible methods. The effect of loading of ZnO NPs was studied on various parameters swelling capacity, porosity, gel content, rate of drug release, and mechanism of drug release from the hydrogel nanocomposite. The drug release kinetic modelling was done for the pure hydrogel and ZnO NPs loaded hydrogel nanocomposite using the Higuchi model and Korsmeyer- Peppas model. Antibacterial studies of drug-loaded pure hydrogel and ZnO NPs loaded hydrogel were done against the *E. coli* bacteria.

2. Experimental sections

2.1. Materials

Carboxymethyl tamarind gum (CMTKG) having 0.20° substitution (courtesy Hindustan Gum Limited, Rewari) acrylic acid (AA, CDH), Poly (ethylene glycol diacrylate) (PEGDA, Mn = 2000, Sigma Aldrich), Sodium hydroxide, Potassium persulfate (KPS, Fischer), Zinc acetate (Sigma Aldrich), were used as received. Ciprofloxacin drug was gifted by M/S Unicure India limited. Buffer solutions of pH 7.4 and pH 1.2 were prepared in the laboratory. Double distilled water was used for carrying out all the experiments.

2.2. Synthesis of CMTKG/Poly (sodium acrylate) hydrogel

A free radical polymerization method was employed to synthesize CMTKG- poly (Sodium acrylate) hydrogel. Initially, 0.15 g of CMTKG was allowed to disperse in 10 mL distilled water for one hour while stirring at 500 rpm. In the beaker, potassium persulfate was added as a free radical initiator after complete dispersion of CMTKG. and stirred for 10 min. After that 5 mL of acrylic acid with 5 mL (2 N), sodium hydroxide was added and again stirred for 20 min. Subsequently, PEGDA was added to the reaction mixture as a cross-linker and stirred for 30 min. Then, the reaction mixture was poured into test tubes and maintained at 60 °C for 2 h in a water bath. The resulting product was then put on Petri plates and dried in an oven at 50 °C temperature. The achieved hydrogel termed as CMTKG/PSA.

2.3. Synthesis of ZnO NPs

The hydrothermal synthetic method was used to make the ZnO NPs [24]. Initially, a solution of 0.05 M Zinc acetate dihydrate ($Zn.(CH3COO)_2{\cdot}2H_2O$) and 0.1 M NaOH were prepared separately in the distilled water. To synthesize, ZnO NPs, the prepared zinc acetate solution was slowly added with continuous stirring to NaOH solution appearing a white color solution which was placed in an oven at 100 °C for 2 h. This solution reacted to give a white color zinc oxide precipitate which upon centrifugation at 2000 rpm for 20 min give ZnO NPs. The particles were dried in the oven at 60 °C for further use.

2.4. Preparation of ZnO NPs loaded CMTKG/Poly (Sodium acrylate) hydrogel nanocomposite

A series of CMTKG/PSA hydrogel nanocomposite were prepared by varying the weight of ZnO NPs. A fixed amount of ZnO NPs 0.03, 0.05, and 0.1 g was mixed into the CMTKG and sodium acrylate reaction mixture subsequent to the addition of PEGDA as a cross-linker in the procedure mentioned in Section 2.2 to obtain hydrogel nanocomposites. The composition of CMTKG/PSA/ZnO hydrogel nanocomposite mentioned in Table 1.

2.5. Swelling studies

To determine the water absorption capacity of hydrogels and ZnO NPs loaded hydrogels, swelling studies were done gravimetrically in which a predetermined amount of purified xerogels was transferred in 100 mL of distilled water, pH 7.4 and pH 1.2 buffer solution. At different time intervals, measurement of weight was done until the constant weight of the hydrogel. The value of the swelling ratio(g/g) was calculated using the given expression in Eq. (1) [25].

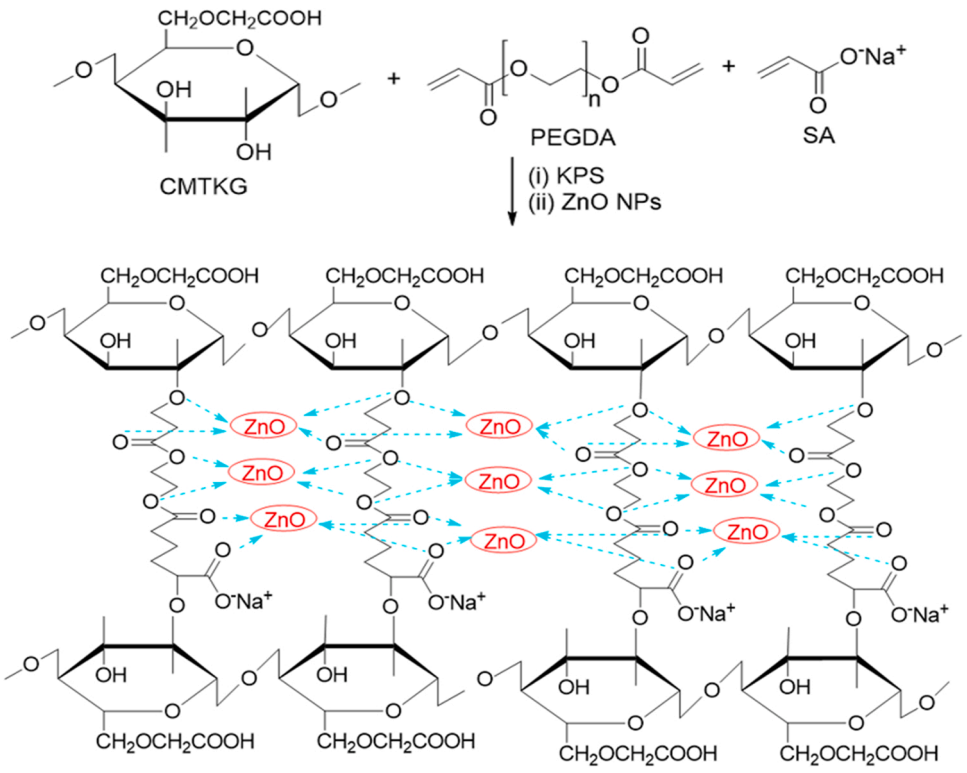
$$SR(g/g) = \frac{W_s - W_0}{W_0} \quad (1)$$

Where Ws represent the mass of hydrogel taken after t time and W_0 denotes the mass of the dry hydrogel sample respectively. 'Ws' is substituted by 'We' in the equilibrium swelling ratio (ESR(g/g)), which is corresponding to the weight of hydrogel taken after the maximum swelling in water.

2.6. Modelling of swelling data

There are many mathematical models which were used to study the swelling capacity of hydrogel quantitatively. Here, the power function model as a time-dependent model was used to study the fractional water uptake using the given Eq. (2) [26].

$$F = \frac{M_t}{M_\infty} = k.t^n \quad (2)$$



Scheme 1. Proposed mechanism for CMTKG/PSA/ZnO hydrogel nanocomposite.

Where M_t denotes the amount of water absorbed or swelling content of hydrogels at t time and M_∞ denotes the maximum water uptake or swelling content at equilibrium. The parameters such as n and k corresponded to the water uptake index and the gel characteristic constant respectively. To further investigate the swelling behavior of hydrogel, the Schott model (second-order) was used. In this model, the value of the water uptake rate constant (k_2) was evaluated using the given Eq. (3) [27].

$$\frac{t}{Mt} = \frac{1}{k_2 M^2_\infty} + \frac{t}{M_\infty} \tag{3}$$

Where M_t denotes the swelling (g/g) or water uptake of the hydrogel at t time and M_∞ denotes the value of theoretically maximum swelling (g/g) of the hydrogel at equilibrium. K_2 is kinetic rate constant of water. The applicability of the given models can be known on basis of the higher value of the regression coefficient (R^2) close to unity.

2.7. Porosity and gel content of hydrogel

The liquid displacement technique was used to determine the porosity of produced hydrogels [18]. A predetermined amount of hydrogel (Wd) was submerged in ethanol till saturation. The initial weight of hydrogel in ethanol was noted as (W_I). The hydrogel sample was taken out after 24 h from ethanol and weighed (W_w) after removing surface liquid with the help of tissue paper. The porosity of the hydrogel was estimated by using Eq. (4) [28].

$$\text{Porosity}(\%) = \frac{W_w - W_d}{W_w - W_I} \times 100 \tag{4}$$

To calculate the gel content present in the hydrogel, the initial weight (W_0) of freshly prepared xerogels was swollen in distilled water for 48 h. It was dried at room temperature after being taken out from distilled water to reach the stage of complete deswelling. Then, the hydrogel was dried in an oven at 40 °C to remove moisture completely. The weight of hydrogel was taken as W_e. The gel content (%) was calculated using the given Eq. (5) [29].

$$\text{Gel Content}(\%) = \frac{W_e}{W_0} \times 100 \tag{5}$$

2.8. Drug loading and entrapment efficiency

The drug loading efficiency of hydrogels is a ratio of the weight of loaded drug to the weight of hydrogels while the drug entrapment efficiency is a ratio of the weight of loaded drug to the weight of the unloaded drug. To evaluate the maximum loading of a drug into the hydrogel, a predetermined amount of xerogel (Wi) was added to 100 mL in a pH 7.4 solution containing a specific amount (W_0) of ciprofloxacin drug. After 48 h, the drug-loaded hydrogel was taken out from the buffer solution and initially dried at room temperature. After that, it was dried in an oven for 48 h at 60 °C temperature and weighed (Wd). Drug loading (DL %) and DE (%) of hydrogel were calculated using Eqs. (6) and (7) respectively [30].

$$\%\text{DL}(\text{mg/g}) = \frac{W_d - W_i}{W_i} \times 100 \tag{6}$$

$$\%\text{DE}(\text{mg/g}) = \frac{W_d - W_i}{W_0} \times 100 \tag{7}$$

2.9. In vitro drug release

The drug release manner of ciprofloxacin drug was studied at a physiological pH of 7.4 at 37 °C temperature. The specific weight of drug incorporated hydrogels was put in 50 mL PBS solution in an orbital shaker incubator with constant shaking at 50 rpm. The absorbance of the drug released was measured using a UV-Visible instrument at a wavelength of 278 nm after removing 2 mL of the supernatant liquid at regular intervals. The total volume of buffer solution was kept constant by adding 2 mL of fresh phosphate buffer solution every time. This absorbance corresponded to the concentration of the released drug as

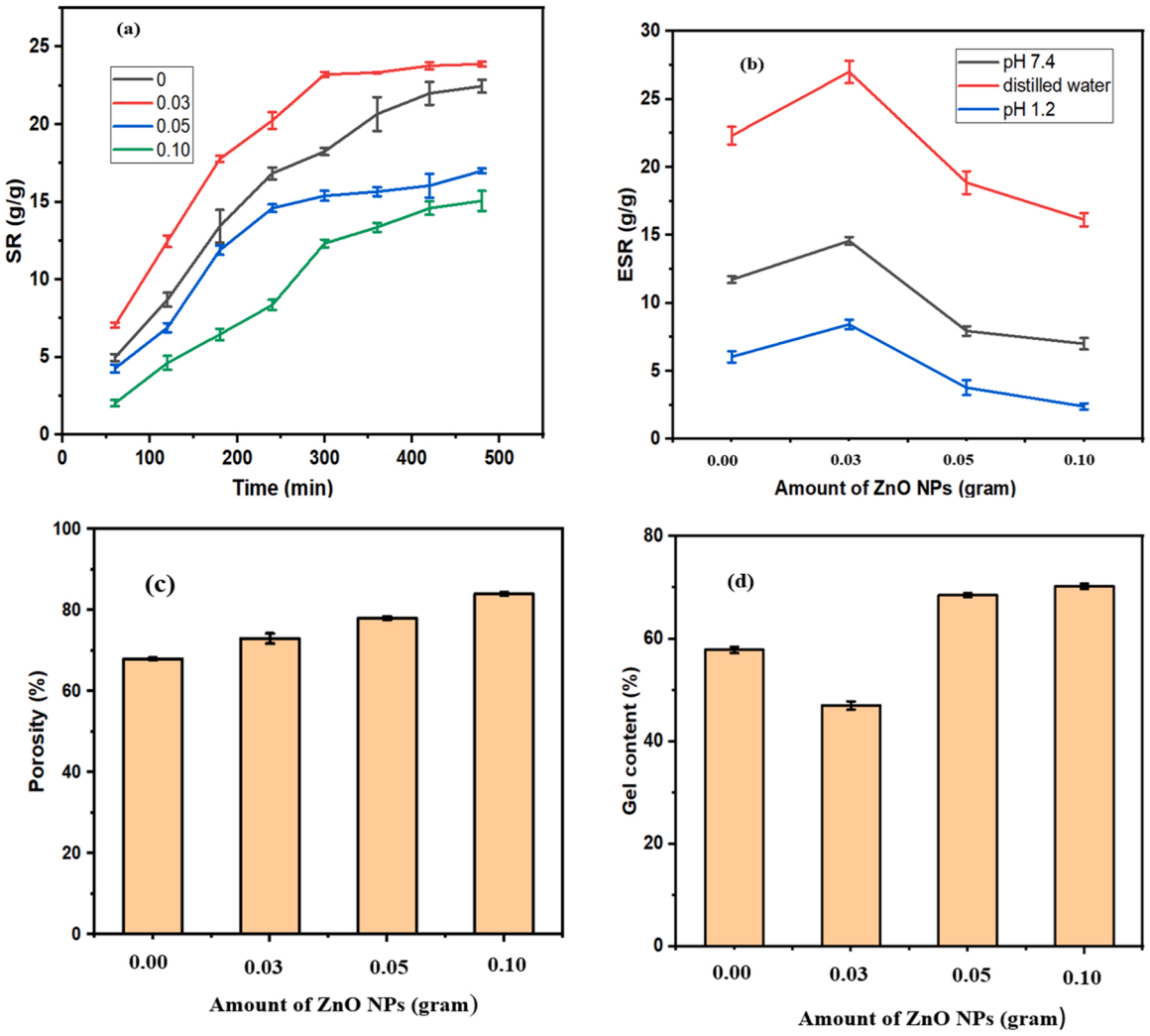


Fig. 1. represent (a) SR (g/g) (b) ESR (g/g) (c) porosity (%) (d) Gel Content (%) for CMTKG/PSA/ZnO hydrogel. The data values are expressed as mean± SD, $p < 0.05$, $n = 3$.

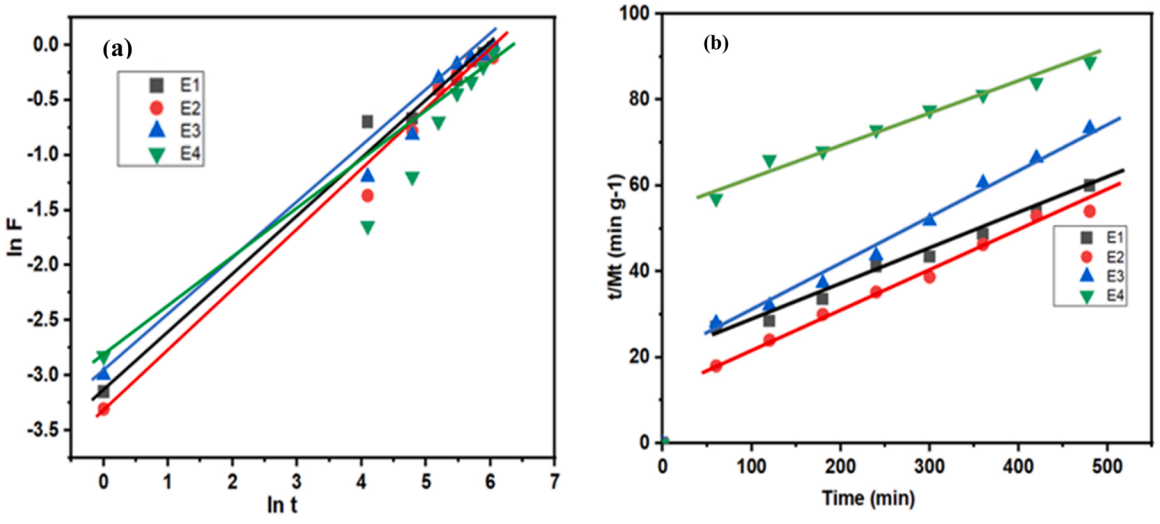


Fig. 2. (a) Power function model (b) Schott function model on swelling data (distilled water) for CMTKG/PSA hydrogel and ZnO loaded hydrogel nanocomposites.

Table 2
Models used to evaluate the dynamic water absorption of CMTKG/PSA hydrogel nanocomposite.

Sample Code	Power function model			Schott kinetic model			
	n	k	R^2	M_∞ (Theo) (g/g)	M∞ (exp) (g/g)	$K_2 \times 10^{-4}$ (g/g min^{-1})	R^2
E1	0.5197	0.0415	0.9981	8.35	8.50	17	0.9932
E2	0.5492	0.035	0.9875	10.78	8.89	7	0.9940
E3	0.5045	0.046	0.9749	8.26	6.54	10	0.9890
E4	0.4620	0.048	0.9430	7.83	5.40	4	0.9931

Table 3
Drug loading (%DL) and drug entrapment (%DE) values into different ZnO NPs loaded CMTKG/PSA hydrogel nanocomposite.

Sample Code	Amount of ZnO NPs	% Drug loading	% Drug entrapment
E1	0.00	10±0.04	58.67±0.89
E2	0.03	21±0.57	65±0.81
E3	0.05	24±0.40	67.50±0.40
E4	1.00	28±0.12	70.65±0.23

Table 4
Detailed analysis of XRD and estimation of various parameters for ZnO NPs.

Sample	d-spacing	FWHM	Crystallite size (nm)
ZnO NPs	2.4657	0.43	20

followed by the Beer-Lambert law. The amount of drug released was calculated using a standard calibration curve for the ciprofloxacin drug. The Cumulative Drug Release (% CDR) was calculated using the given expression in Eq. (8) [21].

$$\text{Drug release}(\%) = \frac{A_t}{A_0} \times 100 \quad (8)$$

Where A_t represents the amount of drug released at any time and A_0 represents the amount loaded in the hydrogel.

2.10. Antibacterial activity

The bactericidal effects of pure hydrogel and ZnO NPs added hydrogel were investigated against *E. coli* NCIMB-1. The disc diffusion test (AATCC-30) was used in this study to determine their antibacterial activity [23]. In a nutshell, 100 mg of various Samples were introduced separately into the well located in the middle of the agar plates for the various bacterial strain cultures under consideration. After that, samples were incubated for 24 h at 37 °C temperature to produce the area of inhibition. Ciprofloxacin drug as a broad-spectrum antibiotic was utilized as a positive control.

2.11. Characterizations

ZnO NPs and CMTKG/PSA/ZnO hydrogel nanocomposite were characterized by FTIR, XRD, FE-SEM, TEM and UV-Visible. The Fourier transform infrared resonance (FTIR) spectra were recorded using the Perkin Elmer Spectrum Model across the frequency range 400–4000 cm^{-1}. Moreover, FESEM imaging of ZnO NPs and ZnO NPs loaded hydrogel was carried out using the ZEISS EVO model along with EDX analysis. TEM imaging of zinc oxide nanoparticles was carried out using TECNAI G 20 (200 KV). Zeta potential measurement of ZnO NPs colloidal solution in ethanol was done using a Zetasizer (Model ZSP, Malvern instrument, UK). The surface plasma resonance (SPR) was recorded by dispersing ZnO NPs loaded hydrogel in ethanol and scanning under a UV–visible spectrophotometer. XRD spectra of ZnO NPs and CMTKG/PSA/ZnO hydrogel were recorded using powder X-ray diffraction with a high-resolution Bruker spectrophotometer over a range of angle $2\theta = 0–80^{\circ}$ with a scan rate of 1 s/step. Rheological measurements were performed using a stress-controlled AR- 500 rheometers (TA Instrument, U.K) using a cone plate geometry (2°, 25 mm). The frequency sweep range from 0.1 to 100 rad/s was used to measure the mechanical spectra in terms of storage modulus((G') and loss modulus (G'').

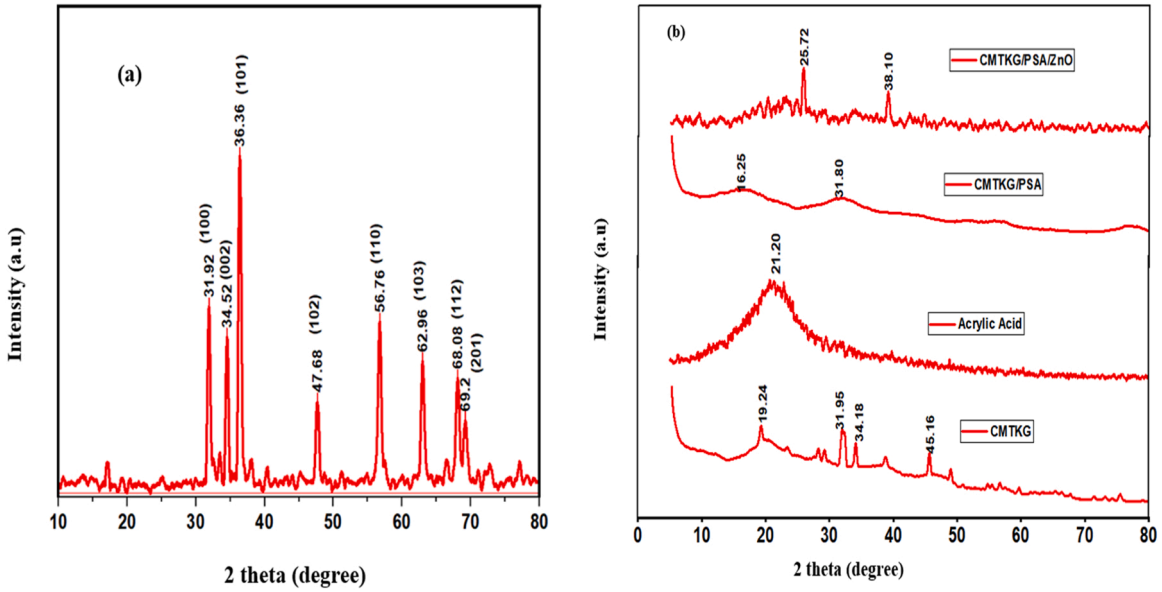


Fig. 3. XRD plot of (a) ZnO NPs (b) CMTKG, acrylic acid, CMTKG/PSA and CMTKG/PSA/ZnO hydrogel nanocomposite.

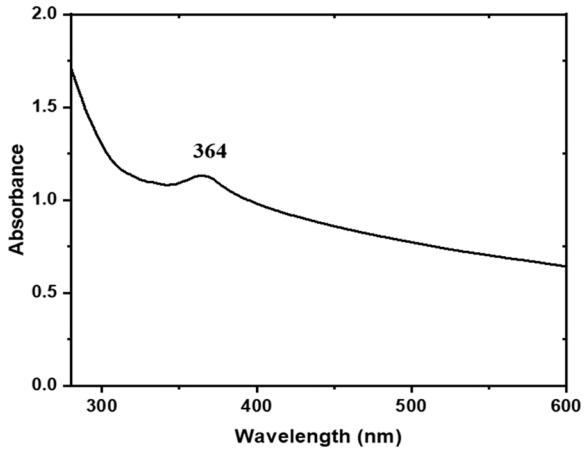


Fig. 4. Surface plasma resonance (SPR) of ZnO NPs.

2.12. Statistical analysis

In this study, the numerical data was expressed as the mean± SD. The data was statistically analyzed by using one way ANOVA followed by Tukey's test with the level of significance set at $p < 0.05$.

3. Results and discussion

3.1. Mechanistic pathway of the reaction

CMTKG/PSA hydrogels was formed by a free radical chain mechanism. Here, potassium persulfate (free radical initiator) was decomposed and produced sulfate radicals at 60 °C [8]. These radicals formed a chain by the removal of the hydrogen atom of the -OH group present on the backbone of CMTKG. These reactive radicals reacted with the C═C double bond present at the terminal position of the PEGDA cross-linker and new radicals formed will react with the C═C bond of sodium acrylate to give a highly crosslinked network as shown in Scheme 1. The hydrogel nanocomposites were formed by the incorporation of ZnO NPs. These particles showed electrostatic interactions with an oxygen atom and carboxyl groups present on the backbone of hydrogels. Also, there may be H-bonding interactions of these particles with the -OH group present on the surface of the CMTKG biopolymer.

3.2. Swelling studies and modeling of swelling data

Swelling studies of hydrogel were done in double distilled water, pH 7.4 and pH 1.2 by placing an accurately weighed quantity of hydrogel. The maximum swelling capacity was noted for hydrogel containing 0.03 g ZnO NPs and exhibited a decreasing trend for the hydrogels with increased concentration of ZnO NPs. The probable cause of maximum swelling for 0.03 g ZnO NPs may be due to the maximum interaction of the oxygen atom of ZnO NPs with the hydrogen atom of the water molecule. ZnO is a polar amphoteric tetrahedrally coordinated structure of Zn^{+2} and O^{-2} ions [26]. The increased concentration of ZnO NPs may lead to additional H-bonding interaction of negatively charged polar atom of oxygen with the hydroxyl group of CMTKG resulting in the decrease of swelling as shown in Fig. 1(a) [26].

It was observed that the ESR(g/g) of hydrogel nanocomposite showed maximum value in distilled water. It may be due to the absence of interfering ions which may interact with the opposite charges present in the hydrogel matrix. At pH 7.4, hydrogel nanocomposite showed higher swelling than at pH 1.2 showed in Fig. 1(b). In the basic medium, there was a large number of negatively charged carboxylate ions due to proper ionization of the -COOH group, the electrostatic repulsion between these groups allowed hydrogel to expand. In an acidic medium, the hydrogel showed the least swelling. It may be due to improper ionization protonation of the -COOH group, H-bonding takes place which would result in a decrease in electrostatic repulsion [3]. The dynamic water adsorption data of hydrogel in distilled water was fitted in the power function model and Schott model as shown in Fig. 2. In the case of the power function model, a linear plot of ln F vs ln t was observed. The calculated parameters such as n, k, and R^2 were showed in Table 2. The value of n corresponds to the diffusion exponent, $n = 0.45$ indicates Fickian diffusion, the value of n between 0.45 and 0.89 for non- Fickian or anomalous diffusion, and $n = 0.89$ for case II diffusion mechanism. The calculated value of n for all samples indicates a non-Fickian swelling behavior of hydrogel. In the case of the Schott model, a linear plot of t/Mt vs t was observed as shown in Fig. 2(b). All parameters related to this model are given in Table 2. It is found that M_∞ (exp) values are nearly the same as M_∞(threo) values. Based on values of R^2, it was concluded that the dynamic swelling data are more fitted in the Schott model.

3.3. Porosity and gel content of the hydrogel

The gel fraction of a hydrogel is related to the cross-linking density which depends on the interaction between various components present

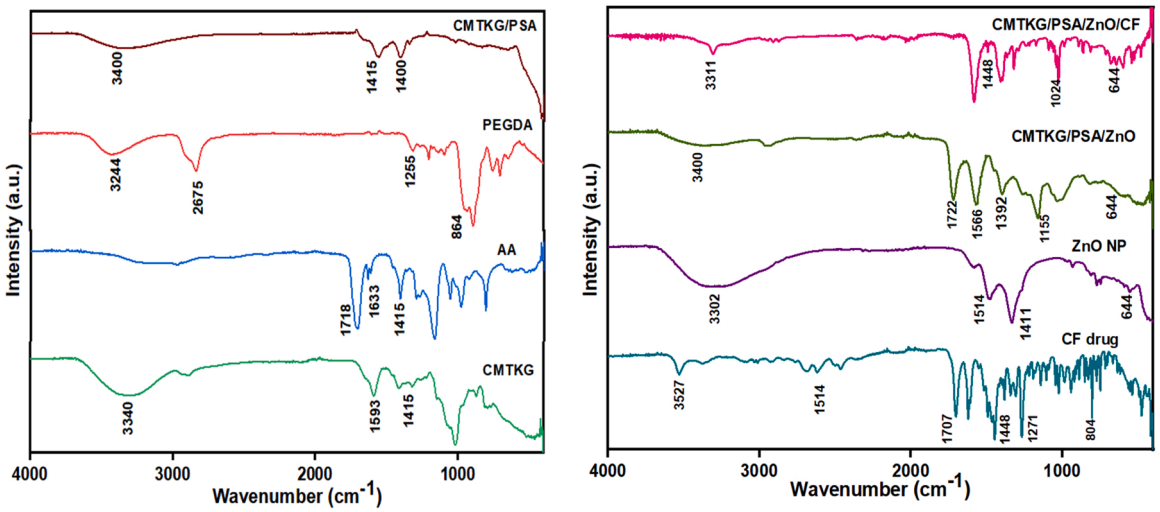


Fig. 5. FTIR plots of (a) CMTKG, AA, PEGDA, and CMTKG/PSA hydrogel (b) CFX drug, ZnO NPs, CMTKG/PSA/ZnO, and CMTKG/PSA/ZnO/CFX hydrogel.

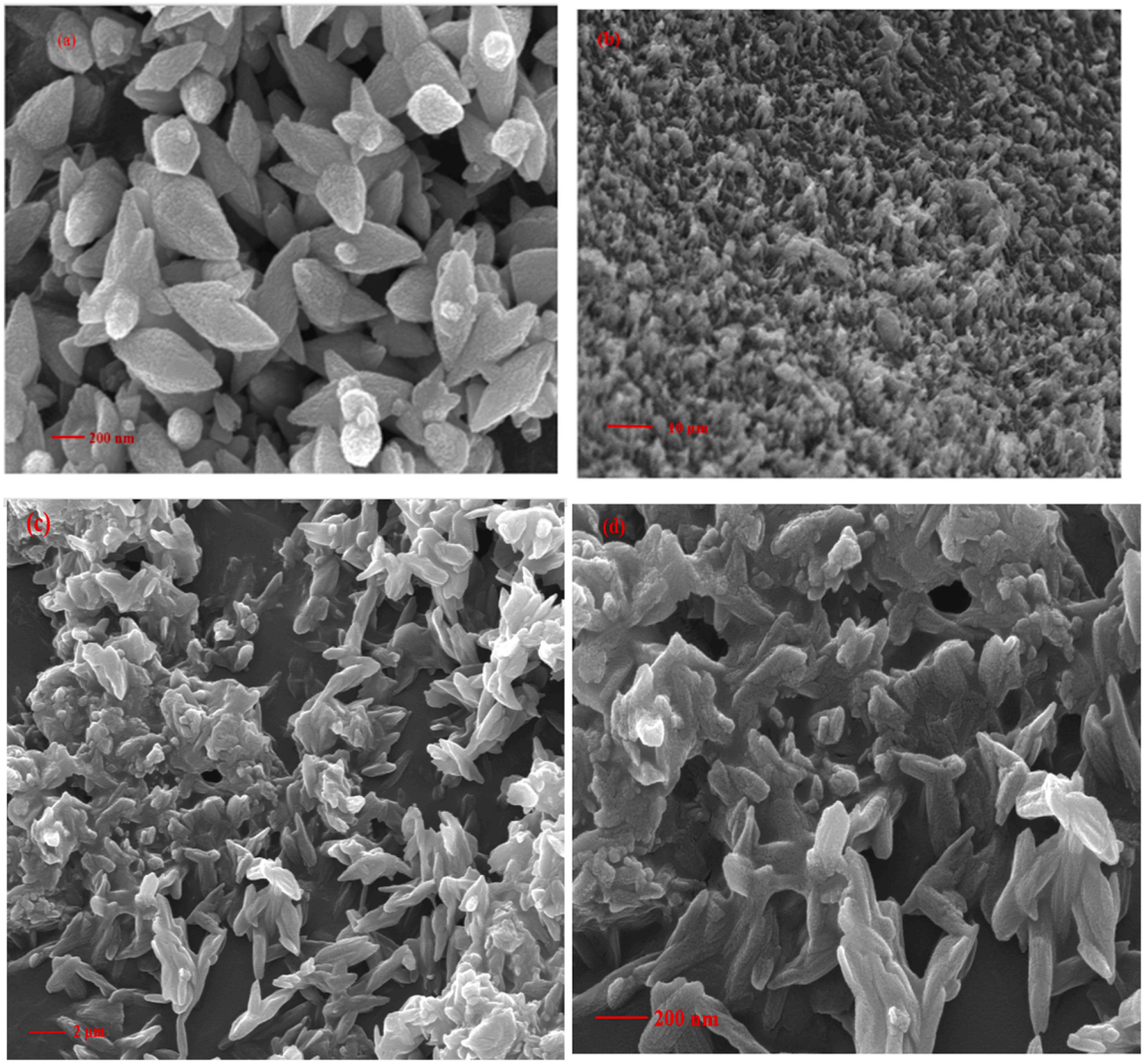


Fig. 6. FE-SEM photographs of (a) ZnO NPs at 200 nm scale (b) CMTKG/PSA at 10 μm scale (c) CMTKG/PSA/ZnO at 2 μm scale (d) CMTKG/PSA/ZnO at 200 nm scale.

in the hydrogel. The gel content of all the synthesized hydrogel nanocomposite was calculated using Eq. (5) as shown in FF Fig. 1(d). It was noticed that the addition of 0.03 g of ZnO NPs, gel content decreased. However, with the further addition of ZnO NPs, gel content decreased. It may be due to the significant interaction of a functional group of nanoparticles with a three-dimensional network of polymeric matrix [32]. In other research, Kalantari reported that with an increase in the amount of CeO_2-NPs, gel content was decreased [28]. Hydrogels have porous structures which provide the space for the encapsulation of drugs. It was observed that the porosity of the synthesized hydrogel increased with an increasing amount of ZnO NPs [29]. The observed range of porosity is from 68 % to 84 % as shown in Fig. 1(c). The porosity of a hydrogel is responsible for its efficacy in absorbing a significant amount of exudate from a wound surface and allowing cells to proliferate. Thus, this hydrogel nanocomposite can also be employed for wound healing [19].

3.4. Drug loading and drug entrapment efficiency

The estimated value of DL (%) and DE (%) for all the hydrogel nanocomposites are shown in Table 3. The drug loading efficiency was calculated with respect to the mass of hydrogel while drug entrapment efficiency was calculated with respect to the mass of the drug in the hydrogel matrix. It was observed that the efficiency of loading and entrapment of the drug was enhanced with an increasing amount of zinc oxide nanoparticles. It may be due to their knotting behavior in which they can form the chelate ring with some -OH and -COO groups present in the polymeric matrix [13]. Secondly, the relative biocompatibility of ZnO NPs and the role of this as a nano-reservoir is responsible for higher loading and entrapment efficiency of hydrogel nanocomposite.

4. Characterizations

4.1. XRD

At room temperature, X-ray powder diffraction (XRD) of ZnO NPs and hydrogel nanocomposite was carried out, as shown in Fig. 6. The XRD pattern of ZnO NPs revealed that they are crystalline and have a hexagonal shape. Debye - Scherrer Eq. (9) was used to calculate the average crystallite size of ZnO NPs [32].

$$D = \frac{0.89\lambda}{\beta\cos\theta} \tag{9}$$

Where λ represents the wavelength of X-ray (0.15418 nm), β is FWHM (Full Width at Half Maximum) in radian, θ indicates diffraction angle and D indicates crystal particle size (nm). X'Pert High Score Plus

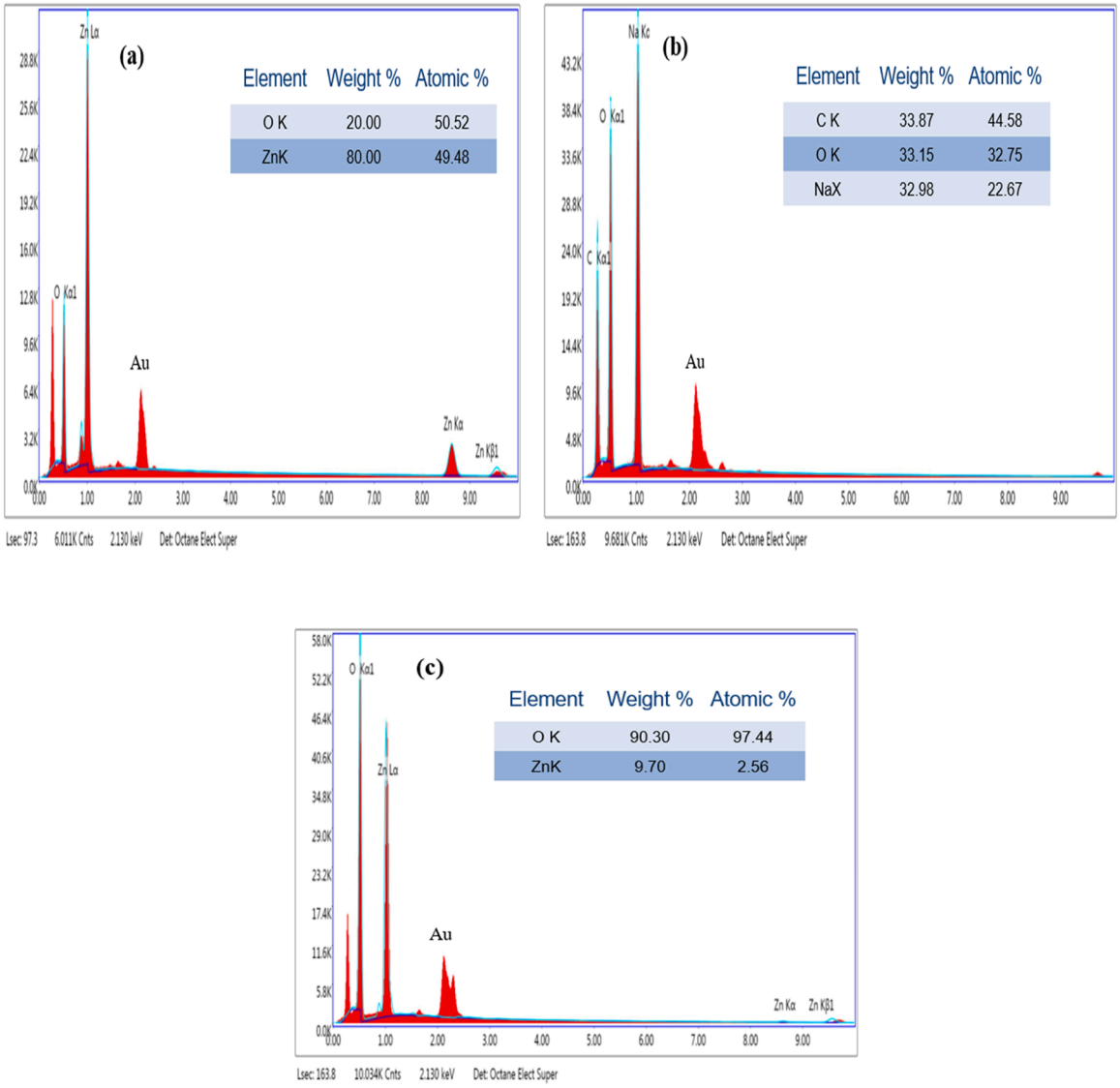


Fig. 7. FESEM-EDX images of (a) ZnO NPs (b) CMTKG/PSA hydrogel (c) CMTKG/PSA/ZnO hydrogel nanocomposite.

software was used to estimate d-spacing values. All the calculated parameters are presented in Table 4. The XRD plot of both ZnO NPs as well as hydrogel nanocomposite is shown in Fig. 3. The diffraction bands at 2θ values of 31.92°, 34.52°, 36.36°, 47.68°, 56.76°, 62.96°, 68.08°, and 69.20° which assigned to (100), (002), (101), (102), (110), (103), (112), and (201) diffractions of zinc oxide nanoparticles [33]. Moreover, XRD of CMTKG, CMTKG/PSA, and CMTKG/PSA/ZnO hydrogel nanocomposite showed in Fig. 3(b). The XRD pattern of CMTKG showed diffraction bands at 19.24°, 31.95°, 34.18°, and 45.16° which may be due to semi- crystalline nature of CMTKG biopolymer. The XRD of acrylic acid monomer showed a broad diffraction peak at 21.20° which are assigned to amorphous nature of acrylic acid. In case of XRD of CMTKG/PSA hydrogels, two broad diffraction peaks at 16.25° and 31.80° indicating the amorphous nature of hydrogel. However, the intensity of peaks of CMTKG decreased after polymerization with amorphous sodium acrylate monomer. Moreover, the XRD pattern of CMTKG/PSA/ZnO hydrogel nanocomposite showed sharp diffraction peaks at 25.72° and 38.10°. Moreover, the significant reduction in the diffraction pattern of ZnO NPs after loading in CMTKG/PSA hydrogels suggest that the crystalline nature of ZnO NPs is converted in amorphous form.

4.2. Surface plasma resonance (SPR)

Surface plasma resonance is a characteristic optical property of metal oxide nanoparticles. The optical property of metallic nanoparticles is dependent on SPR. The plasmon term in SPR signifies the collective oscillation of free electrons within the inorganic nanoparticles and the width of the plasmon peak depends on the type, geometry, and size of the metal [34]. The UV-Visible absorbance spectra of the zinc oxide nanoparticles incorporated hydrogel have been illustrated in Fig. 4. which showed that these particles exhibited a strong, sharp characteristic peak at 364 nm. This confirmed the successful synthesis of hydrogel nanocomposites.

4.3. FTIR analysis

The FTIR spectrum of CMTKG biopolymer in Fig. 5(a). indicated a broad peak at 3340 cm^{-1} for vibration of -OH group while the peaks at 1415 cm^{-1} and 1593 cm^{-1} are related with symmetric and asymmetric vibration of -COO groups respectively. The peaks observed in FTIR of acrylic acid at 1633 cm^{-1} and 1718 cm^{-1} are related to C═C and C-O vibration respectively. The FTIR spectrum of PEGDA showed a band at 2676 cm^{-1} due to vibration of -CH_2 group, whereas the peak at

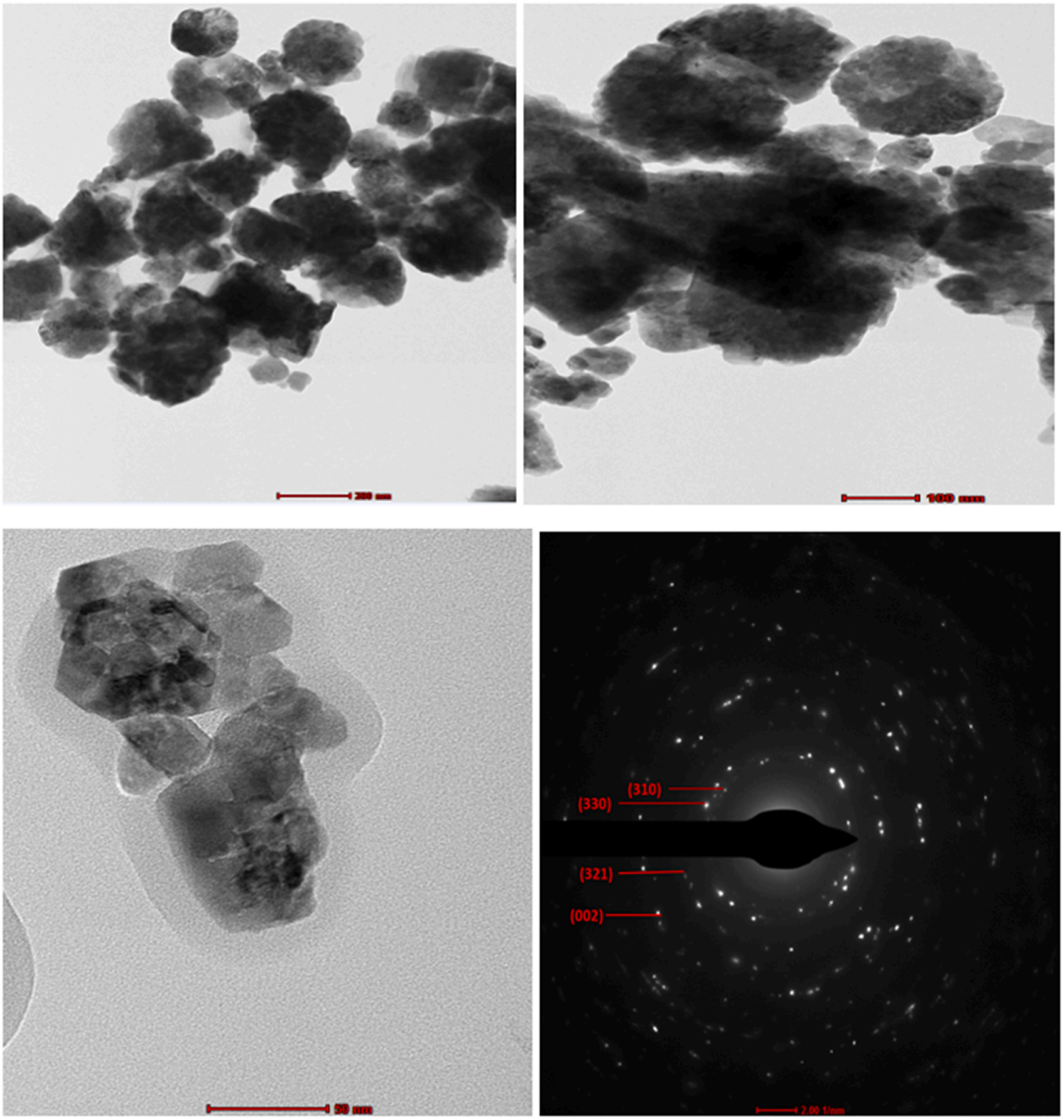


Fig. 8. (a) TEM micrograph of ZnO NPs at 200 nm, 100 nm, and 50 nm scales (b) SAED pattern.

1255 cm^{-1} and 864 cm^{-1} are attributed to the C-O stretching and C═C bending. In FTIR spectrum of CMTKG/PSA hydrogels, there was an -OH stretching band at 3400 cm^{-1}, -COO symmetric vibration band at 1415 cm^{-1}, and O-H bending vibration at 1400 cm^{-1}. However, one common peak at 1415 cm^{-1} related to vibration of -COO groups was noticed in CMTKG, AA, PEGDA and CMTKG/PSA hydrogel. The FTIR spectrum in Fig. 5(b) for ZnO NPs showed a band at 644 cm^{-1} assigned to the stretching vibration of the Zn-O bond. However, due to the hygroscopic nature of ZnO NPs, a broad peak at 3302 cm^{-1} was observed which are assigned to the stretching of the hydroxyl group. The bands observed at 1514 and 1411 cm^{-1} in zinc oxide nanoparticles correspond to the stretching vibration and asymmetric vibration of the -C═O group respectively [35]. The FTIR spectra in Fig. 5(b) of ciprofloxacin drug showed the characteristic band at 1707 cm^{-1} corresponded to C= O group, 3527 cm^{-1} due to vibration of O-H bond, 1448 cm^{-1} corresponded to C- N stretching vibration. However, the presence of carboxylic acid was confirmed by the presence of the peak at 1271 cm^{-1} due to the bending vibration of the O-H group [36,37]. FTIR spectra of ZnO loaded hydrogels showed a broad peak at 3400 cm^{-1} due to vibration of the -OH group. Other bands at 1392 cm^{-1} and 1566 cm^{-1} assigned to the stretching of C-O and C═O stretching vibration of carboxylic acid. A peak at 644 cm^{-1} related to vibration of ZnO bond was noticed in CMTKG/PSA/ZnO hydrogel indicated the loading of nanoparticles in hydrogel matrix. Moreover, both peaks at 644 cm^{-1} and 1448 cm^{-1} related to characteristics of ZnO NPs and ciprofloxacin drug respectively are also present in CMTKG/PSA/ZnO/CFX hydrogel nanocomposite. However, some common peaks were noticed in the FTIR spectrum of ZnO, CFX drug, and drug loaded CMTKG/PSA/ZnO nanocomposite. All the related peaks of ZnO nanoparticles are present in ZnO loaded hydrogels which confirmed the formation of hydrogel nanocomposite. In FTIR spectra of without drug and drug-loaded hydrogel nanocomposite, it was seen that all the related peaks of drug were present in drug loaded hydrogel nanocomposite which signified the loading of ciprofloxacin drug in the hydrogel matrix.

4.4. FE-SEM analysis

The surface morphology of ZnO NPs, CMTKG/PSA, and CMTKG/

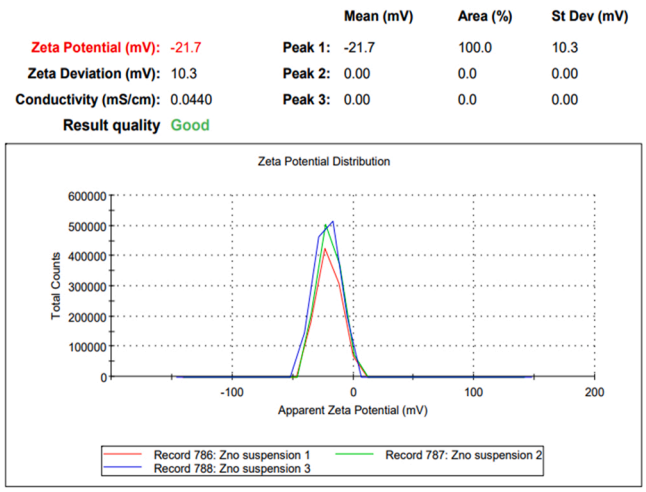


Fig. 9. Zeta potential of synthesized ZnO nanoparticles.

PSA/ZnO hydrogel nanocomposite were studied by the FE-SEM analysis [38]. The morphology of zinc oxide nanoparticles formed depends on the pH of the reaction. The FE-SEM micrograph of ZnO NPs was shown in Fig. 6(a) which indicated that they have flower shaped morphology. However, the FE-SEM analysis of CMTKG/PSA hydrogels showed lots of interspatial voids and spongy surfaces as showed in Fig. 6(b). However, the surface morphology of CMTKG/PSA/ZnO nanocomposite showed that flower shaped ZnO NPs are embedded in the polymeric hydrogel matrix. Moreover, the average size of ZnO nanoparticles was estimated using ImageJ software which is found equal to 118 nm. Moreover, the EDX was used to analyze the composition of various elements present in the synthesized products. It revealed that the presence of Zn and O elements in ZnO NPs while Na, C, and O are the main constituent of CMTKG/PSA hydrogel Fig. 7).

4.5. TEM analysis

The TEM image of synthesized ZnO nanoparticles in Fig. 8 showed that particles have aggregated spherical shapes. The average particle size of ZnO nanoparticles was estimated using ImageJ software. The estimated size of ZnO NPs was found to be 37 nm. However, the estimated particle size determined by XRD was equal to 20 nm which is in good agreement with the size determined by TEM analysis [39]. In the SAED pattern, specific bright spots were observed which correspond to the hexagonal wurtzite polycrystalline structure of ZnO NPs. The SAED pattern showed the diffraction peaks at (002), (310), (330), and (321) corresponded to ZnO nanoparticles. The interplanar spacing d (A°) was calculated for different regions of the SAED pattern using ImageJ software. The estimated value of d spacing (angstrom unit) is 3 Å, 2.38 Å, 1.95 Å and 1.46 Å for (310), (330), (321), and (002) planes respectively. However, a strong diffraction peak at (002) plane in the XRD of ZnO NPs is consistent with their SAED pattern. Thus, both the characterization confirmed the wurtzite structure and synthesis of ZnO nanoparticles.

4.6. Zeta potential measurement

The zeta potential was used to investigate metallic nanoparticle surface charges and stability. The zeta potential of produced zinc oxide nanoparticles was measured using ethanol as a dispersant as shown in Fig. 9. In the present study, it is measured as 21.7 mV which indicated its strongly anionic behavior as well as the stability of the colloidal suspension.

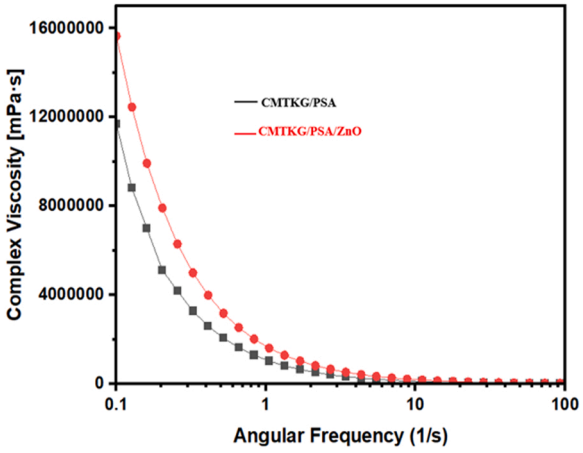


Fig. 11. A plot of complex viscosity vs angular frequency.

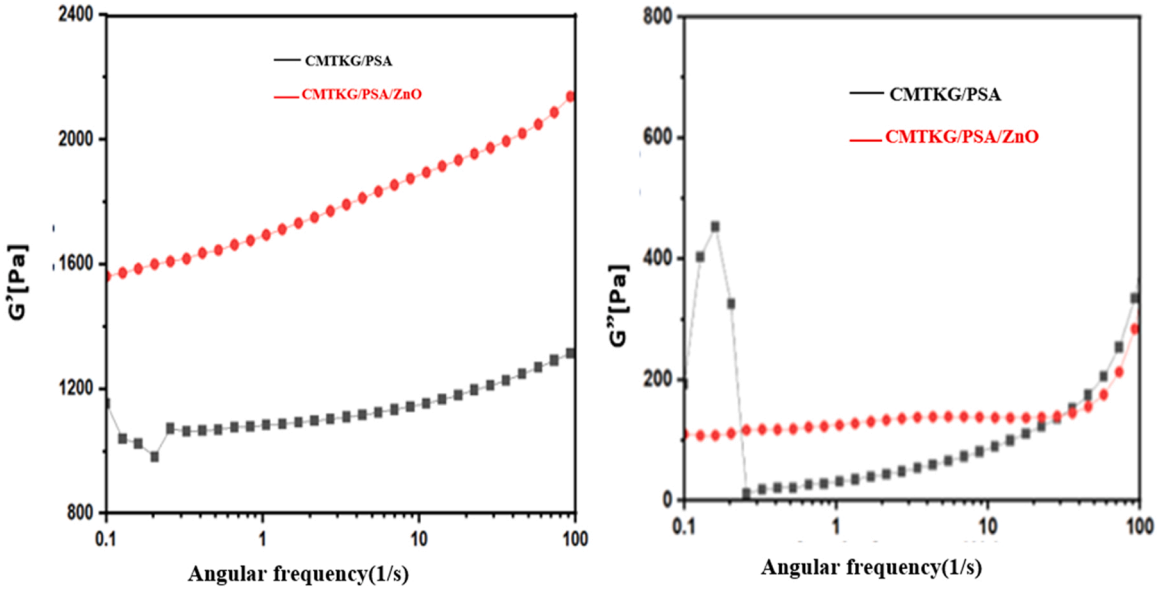


Fig. 10. (a) Storage modulus (G′) (b) Loss modulus (G′′) of pure hydrogels and ZnO NPs embedded hydrogel nanocomposite.

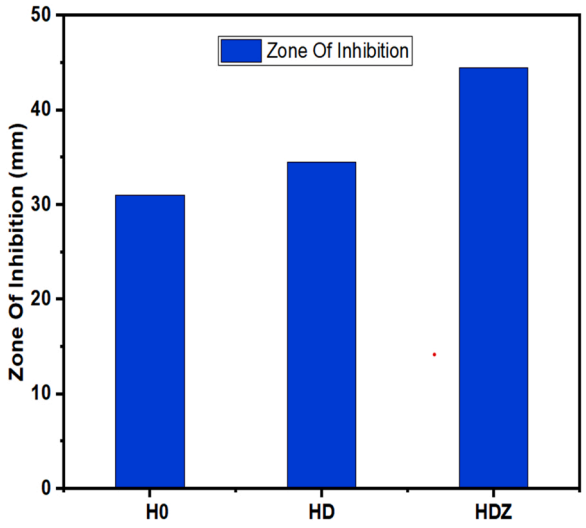


Fig. 12. Size of the zone of inhibition (mm) exhibited by CMTKG/PSA hydrogel (H0), CMTKG/PSA-CF (HD), and CMTKG/PSA/ZnO-CF (HDZ) hydrogel nanocomposite.

4.7. Rheological analysis

The rheological behavior of CMTKG/PSA and CMTKG/PSA/ZnO hydrogel nanocomposite was studied based on change of storage modulus (G′), loss modulus (G″), and complex viscosity as a function of angular frequency. The storage modulus corresponds to the elastic component while the loss modulus corresponded to the viscous component. The storage modulus(G′) of pure hydrogel and hydrogel nanocomposite is illustrated in Fig. 10(a). It is observed that G′ of hydrogel nanocomposite is higher than pure hydrogel. It indicated that the loading of ZnO NPs increased the cross-linking strength of hydrogel due to the physical interactions of these particles with the hydrogel polymeric network [40]. It is also seen that along with the storage modulus, the loss modulus(G″) increased with an increase in the angular frequency from 1 Hz to 100 Hz as shown in Fig. 10(b). At the lower frequency region, the value of loss modulus for pure hydrogel is higher than ZnO loaded hydrogels. The reason may be that there is a higher

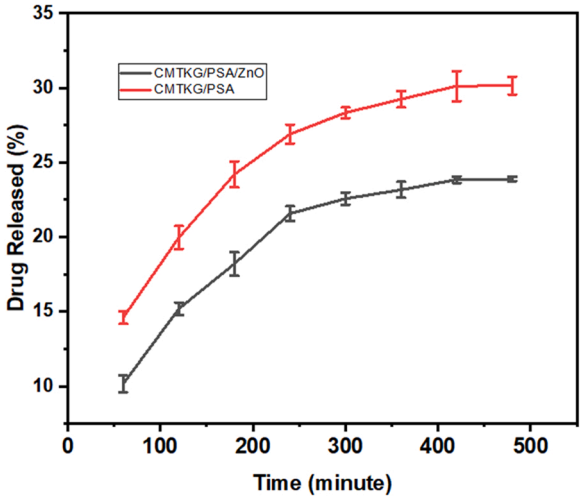


Fig. 14. The drug release (%) with time of CMTKG/PSA and CMTKG/PSA/ZnO hydrogel nanocomposite.

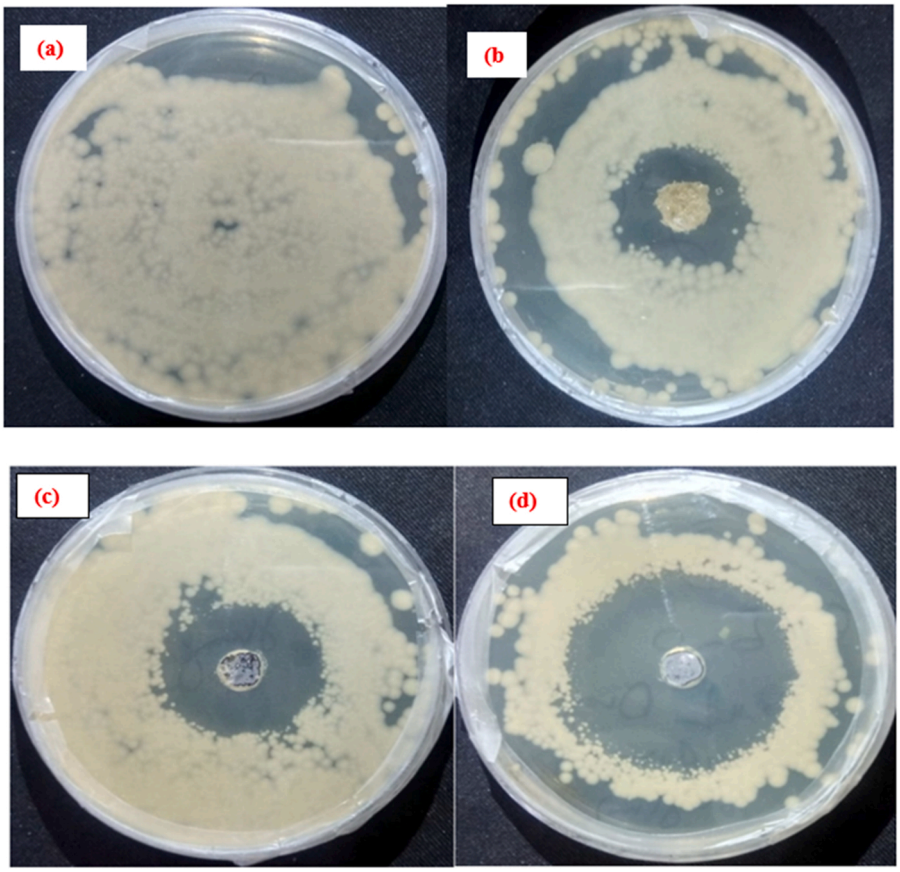


Fig. 13. Antibacterial actions (a) blank sample(b) CMTKG/PSA (c) CMTKG/PSA-CF (d) CMTKG/PSA/ZnO-CF against *E. coli* bacteria.

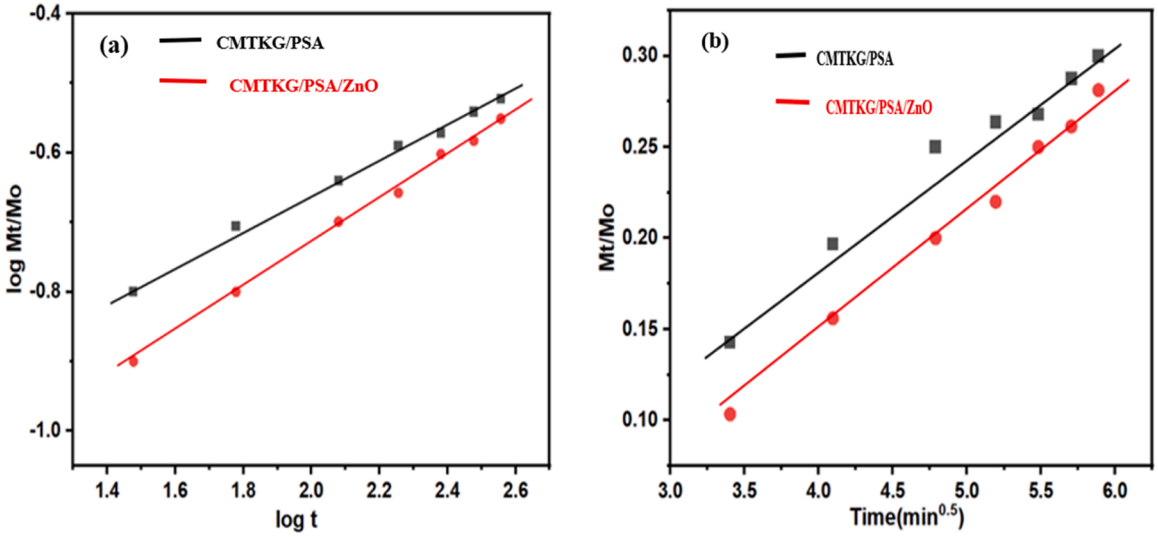


Fig. 15. (a) Korsmeyer-Peppas model (b) Higuchi model for CMTKG/PSA and CMTKG/PSA/ZnO hydrogel.

Table 5
Model followed for Drug release studies using ciprofloxacin loaded CMTKG/ PSA hydrogel (E1) and CMTKG/PSA/ZnO (E2) hydrogel nanocomposite.

Sample Code	Different parameters associated with models for ciprofloxacin drug release				
	Korsmeyer -Peppas Model			Higuchi Model	
	K	n	R^2	K_h	R^2
E1	0.069	0.25	0.9951	0.060	0.9885
E2	0.042	0.32	0.9932	0.069	0.9921

possibility of the formation of a physically cross-links network in absence of zinc oxide nanoparticles at a lower frequency region [41]. However, the presence of zinc oxide nanoparticles provided more stability, rigidity, and thickness to the hydrogel matrix. It showed the pseudo elastic behavior of hydrogel nanocomposite [42].

The decrease in value of complex viscosity with an increase in the angular frequency changes from 0.1 to 100 rad/s clearly showed in Fig. 11 that hydrogels are pseudoplastic fluid and showed thixotropic behavior [43]. It was also found that the complex viscosity of ZnO embedded hydrogel is higher than pure hydrogel because of the higher value of storage modulus (G') for hydrogel nanocomposite. Hence, ZnO embedded hydrogel nanocomposites are better drug delivery carriers than pure hydrogel. This is also confirmed by the values of DL (%) and DE (%) efficiency given in Table 2.

5. Antibacterial studies

Antibacterial performance of CMTKG/PSA, CMTKG/PSA-CF, and CMTKG/PSA/ZnO-CF were studied via a disc diffusion method against gram-negative *E. coli* (NCIMB-1) bacteria as shown in Fig. 13. The zone of inhibition (mm) for pure hydrogel, drug-loaded hydrogel, and hydrogel nanocomposite was measured as showed in Fig. 12. Ciprofloxacin is an antibiotic drug therefore enhanced the bacteriostatic behavior of hydrogels. ZnO nanoparticles themselves exhibited antibacterial properties by interacting with the bacterial surface [44]. The loading of antibacterial ZnO NPs and ciprofloxacin drug increased the antibacterial action of xerogels which showed the highest radius of zone of inhibition (Fig. 13).

6. Drug release studies

The release mode of ciprofloxacin drug from CMTKG/PSA hydrogels and ZnO loaded hydrogel nanocomposite is illustrated in Fig. 14. The drug release behavior of hydrogels was studied in 7.4 at 37 °C temperature using a UV-Visible spectrophotometer. The drug release was performed for up to 9 h and it was seen that after 6 h, a plateau region of the drug release curve was observed. The release of ciprofloxacin drug is thus relying on the swelling of hydrogel [45]. The result concluded that ZnO embedded hydrogel nanocomposite showed more controlled release than pure hydrogel. This is due to the knot-tying function of ZnO nanoparticles [31]. Thus, a chelated ring was formed via the interaction of zinc oxide nanoparticles with hydroxyl and carboxyl groups of carboxymethyl tamarind gum [46]. As a result, it was difficult for the ciprofloxacin drug to release from such tightened network of hydrogel nanocomposite. Thus, ZnO embedded hydrogel showed a slow and controlled release of ciprofloxacin drugs.

6.1. Drug release mechanism

The release of drugs from the hydrogel matrix was studied with the help of Higuchi and Korsmeyer-Peppas models(Fig. 15). The Higuchi model is dependent on the Fickian diffusion process. It is useful to explain the release of hydrophilic and hydrophobic drugs from the polymeric system. The drug release mechanism in the case of Korsmeyer- Peppas model is dependent on the diffusion exponent value [47]. The various parameters such as K, K_h, n, and R^2 are calculated using these models as shown in Table 5. Several physical and chemical phenomena affect the drug release kinetics such as swelling of hydrogel, mesh size, porosity, and chemical interaction of drug with polymers. The Higuchi model and Korsmeyer -Peppas models are expressed in given Eqs. (10) and (11) respectively [48].

$$\frac{M_t}{M_\infty} = K_h.t^{0.5} \tag{10}$$

$$\frac{M_t}{M_\infty} = K.t^n \tag{11}$$

Where $\frac{M_t}{M_\infty}$ represents the fraction of released drugs at t time with respect to infinite time. Mt represents the amount of drug released at t time and M_∞ represents the amount of drug released at equilibrium. The K and K_h are rate constant for drug release corresponded to Korsemeyer-Pepass and Higuchi model. The value of diffusion exponent n decides the mechanism of drug release. In the case, the value of n less or equal to 0.5 indicated Fickian, $1 > n > 0.5$ non Fickian, and $n = 1$ case II anomalous diffusion [18]. Since the calculated value of n for pure hydrogel (E1) and

ZnO loaded hydrogels (E2) is less than 0.5, drug release from the hydrogels followed Fickian law of diffusion. It clearly showed that drug release from hydrogel depends on purely diffusivity phenomena and for both samples, the solvent diffusion is higher than polymeric chain relaxation [49]. The value of R^2 is compared for pure hydrogel and ZnO-loaded hydrogels. It is found that the drug release data are fitted in both models.

7. Conclusion

In this research work, ZnO NPs loaded CMTKG/PSA hydrogel nanocomposites were synthesized using a PEGDA cross-linker. ZnO NPs were made via a hydrothermal method while the synthesis of CMTKG/PSA hydrogels was done via a free radical mechanism. The synthesized ZnO NPs and hydrogel nanocomposites were characterized using FTIR, XRD, FE-SEM, and TEM. The swelling studies of hydrogels were performed at different pH. The dynamic swelling data were fitted in the Korsmeyer-Peppas model and Schott model. The result concluded that data was more fitted in the Schott model with a higher R^2 value close to unity. The synthesized hydrogel and ZnO-loaded hydrogel nanocomposites were used in in-vitro release studies of the ciprofloxacin drug. The various parameters such as porosity, drug loading, and entrapment efficiency of hydrogel improved in presence of ZnO NPs. The rheological studies of hydrogels and ZnO loaded hydrogels were performed which concluded that the elasticity of hydrogels was enhanced in the presence of ZnO NPs. The effect of incorporation of ZnO NPs into drug-loaded hydrogels was studied on antibacterial properties against *E. coli* bacteria. The result revealed that due to the incorporation of antibacterial ZnO NPs, the antibacterial action of drug-loaded hydrogels increased. The in vitro drug release results concluded that CMTKG/PSA/ZnO hydrogels nanocomposite led to more controlled slow drug release. The drug release kinetics concluded that the release of the drug from hydrogels was followed by Korsmeyer-Peppas and Higuchi model. In brief, CMTKG/PSA/ZnO hydrogel nanocomposite could be a better candidate for the oral delivery of ciprofloxacin drugs.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

AK, SGW, and IR would like to thank Delhi Technological University (DTU) for providing instrumental facilities. IR also thanks Council for scientific and industrial research (CSIR) 08/133(0036)/2019-EMR-I, India for Senior Research Fellowship (SRF).

References

[1] G.R. Bardajee, F. Mizani, S.S. Hosseini, pH sensitive release of doxorubicin anticancer drug from gold nanocomposite hydrogel based on poly(acrylic acid) grafted onto salep biopolymer, J. Polym. Res. 24 (3) (2017), https://doi.org/10.1007/s10965-017-1197-4.

[2] L. Ali, M. Ahmad, M.N. Aamir, M.U. Minhas, H.H. Shah, M.A. Shah, Cross-linked pH-sensitive pectin and acrylic acid based hydrogels for controlled delivery of metformin, Pak. J. Pharm. Sci. 33 (4) (2020) 1483–1491, https://doi.org/10.36721/PJPS.2020.33.4.REG.1483-1491.1.

[3] N. Ajaz, I.U. Khan, I. Khalid, R.U. Khan, H.A. Khan, S. Asghar, S.H. Khalid, Y. Shahzad, A.M. Yousaf, T. Hussain, N. Sabir, In vitro and toxicological assessment of dexamethasone sodium phosphate loaded pH sensitive Pectin-g-poly(AA)/PVP semi interpenetrating network, Mater. Today Commun. 25 (June) (2020), 101325, https://doi.org/10.1016/j.mtcomm.2020.101325.

[4] S.K. Ghosh, A. Das, A. Basu, A. Halder, S. Das, S. Basu, M.F. Abdullah, A. Mukherjee, S. Kundu, Semi-interpenetrating hydrogels from carboxymethyl guar gum and gelatin for ciprofloxacin sustained release, Int. J. Biol. Macromol. 120 (2018) 1823–1833, https://doi.org/10.1016/j.ijbiomac.2018.09.212.

[5] V.F. Lotfy, A.H. Basta, Optimizing the chitosan-cellulose based drug delivery system for controlling the ciprofloxacin release versus organic/inorganic crosslinker, characterization and kinetic study, Int. J. Biol. Macromol. 165 (2020) 1496–1506, https://doi.org/10.1016/j.ijbiomac.2020.10.047.

[6] M. Rawooth, D. Qureshi, M. Hoque, M.G. Prasad, B. Mohanty, M.A. Alam, A. Anis, P. Sarkar, K. Pal, Synthesis and characterization of novel tamarind gum and rice bran oil-based emulgels for the ocular delivery of antibiotics, Int. J. Biol. Macromol. 164 (2020) 1608–1620, https://doi.org/10.1016/j.ijbiomac.2020.07.231.

[7] Z. Naderi, J. Azizian, E. Moniri, N. Farhadyar, Synthesis and characterization of carboxymethyl cellulose/β-cyclodextrin/chitosan hydrogels and investigating the effect of magnetic nanoparticles (Fe3O4) on a novel carrier for a controlled release of methotrexate as drug delivery, J. Inorg. Organomet. Polym. Mater. 30 (4) (2020) 1339–1351, https://doi.org/10.1007/s10904-019-01301-1.

[8] S.G. Khushbu, Warkar, A. Kumar, Synthesis and assessment of carboxymethyl tamarind kernel gum based novel superabsorbent hydrogels for agricultural applications, Polymer 182 (July) (2019), 121823, https://doi.org/10.1016/j.polymer.2019.121823.

[9] Khushbu, S.G. Warkar, Potential applications and various aspects of polyfunctional macromolecule- carboxymethyl tamarind kernel gum, Eur. Polym. J. 140 (September) (2020), 110042, https://doi.org/10.1016/j.eurpolymj.2020.110042.

[10] K.K. Mali, S.C. Dhawale, R.J. Dias, Synthesis and characterization of hydrogel films of carboxymethyl tamarind gum using citric acid, Int. J. Biol. Macromol. 105 (2017) 463–470, https://doi.org/10.1016/j.ijbiomac.2017.07.058.

[11] S. Jana, R. Sharma, S. Maiti, K.K. Sen, Interpenetrating hydrogels of O-carboxymethyl Tamarind gum and alginate for monitoring delivery of acyclovir, Int. J. Biol. Macromol. 92 (2016) 1034–1039, https://doi.org/10.1016/j.ijbiomac.2016.08.017.

[12] H. Kaur, M. Ahuja, S. Kumar, N. Dilbaghi, Carboxymethyl tamarind kernel polysaccharide nanoparticles for ophthalmic drug delivery, Int. J. Biol. Macromol. 50 (3) (2012) 833–839, https://doi.org/10.1016/j.ijbiomac.2011.11.017.

[13] X. Sun, C. Liu, A.M. Omer, W. Lu, S. Zhang, X. Jiang, H. Wu, D. Yu, X.K. Ouyang, pH-sensitive ZnO/carboxymethyl cellulose/chitosan bio-nanocomposite beads for colon-specific release of 5-fluorouracil, Int. J. Biol. Macromol. 128 (2019) 468–479, https://doi.org/10.1016/j.ijbiomac.2019.01.140.

[14] R. Patwa, O. Zandraa, Z. Capáková, N. Saha, P. Sáha, Effect of iron-oxide nanoparticles impregnated bacterial cellulose on overall properties of alginate/casein hydrogels: Potential injectable biomaterial for wound healing applications, Polymer 12 (11) (2020) 1–21, https://doi.org/10.3390/polym12112690.

[15] Y. Ahmadian, A. Bakravi, H. Hashemi, H. Namazi, Synthesis of polyvinyl alcohol/CuO nanocomposite hydrogel and its application as drug delivery agent, Polym. Bull. 76 (4) (2019) 1967–1983, https://doi.org/10.1007/s00289-018-2477-9.

[16] K. Virk, K. Sharma, S. Kapil, V. Kumar, V. Sharma, S. Pandey, V. Kumar, Synthesis of gum acacia-silver nanoparticles based hydrogel composites and their comparative anti-bacterial activity, J. Polym. Res. 29 (4) (2022) 1–15, https://doi.org/10.1007/s10965-022-02978-8.

[17] M.T. Khorasani, A. Joorabloo, A. Moghaddam, H. Shamsi, Z. Mansoorimoghadam, International Journal of Biological Macromolecules Incorporation of ZnO nanoparticles into heparinised polyvinyl alcohol/ chitosan hydrogels for wound dressing application, Int. J. Biol. Macromol. 114 (2018) 1203–1215, https://doi.org/10.1016/j.ijbiomac.2018.04.010.

[18] G.A. Mahmoud, A.E.H. Ali, A.I. Raafat, N.A. Badawy, M.F. Elshahawy, Development of (acrylic acid/ polyethylene glycol)-zinc oxide mucoadhesive nanocomposites for buccal administration of propranolol HCl, Radiat. Phys. Chem. 147 (July 2017) (2018) 18–26, https://doi.org/10.1016/j.radphyschem.2018.01.020.

[19] R. Priyadarshi, B. Kumar, J.W. Rhim, Green and facile synthesis of carboxymethylcellulose/ZnO nanocomposite hydrogels crosslinked with Zn2+ ions, Int. J. Biol. Macromol. 162 (2020) 229–235, https://doi.org/10.1016/j.ijbiomac.2020.06.155.

[20] D.H. Hanna, G.R. Saad, Bioorganic Chemistry Encapsulation of ciprofloxacin within modified xanthan gum- chitosan based hydrogel for drug delivery, Bioorg. Chem. 84 (October 2018) (2019) 115–124, https://doi.org/10.1016/j.bioorg.2018.11.036.

[21] K. Prusty, A. Biswal, S.B. Biswal, S.K. Swain, Synthesis of soy protein/polyacrylamide nanocomposite hydrogels for delivery of ciprofloxacin drug, Mater. Chem. Phys. 234 (May) (2019) 378–389, https://doi.org/10.1016/j.matchemphys.2019.05.038.

[22] G.R. Mahdavinia, M.H. Karimi, M. Soltaniniya, B. Massoumi, In vitro evaluation of sustained ciprofloxacin release from κ-carrageenan-crosslinked chitosan/hydroxyapatite hydrogel nanocomposites, Int. J. Biol. Macromol. 126 (2019) 443–453, https://doi.org/10.1016/j.ijbiomac.2018.12.240.

[23] K. McAvoy, D. Jones, R.R.S. Thakur, Synthesis and characterisation of photocrosslinked poly(ethylene glycol) diacrylate implants for sustained ocular drug delivery, Pharm. Res. 35 (2) (2018), https://doi.org/10.1007/s11095-017-2298-9.

[24] D.A.M.O. Osman, M.A. Mustafa, Synthesis and characterization of zinc oxide nanoparticles using zinc acetate dihydrate and sodium hydroxide, J. Nanosci. Nanoeng. 1 (4) (2015) 248–251 [Online]. Available: ⟨http://www.aiscience.org/journal/jnnhttp://creativecommons.org/licenses/by-nc/4.0/⟩.

[25] S. Jana, A. Banerjee, K.K. Sen, S. Maiti, Gelatin-carboxymethyl tamarind gum biocomposites: In vitro characterization & anti-inflammatory pharmacodynamics, Mater. Sci. Eng. C. 69 (2016) 478–485, https://doi.org/10.1016/j.msec.2016.07.008.

[26] S.K. Bajpai, M. Jadaun, S. Tiwari, Synthesis, characterization and antimicrobial applications of zinc oxide nanoparticles loaded gum acacia/poly(SA) hydrogels, Carbohydr. Polym. 153 (2016) 60–65, https://doi.org/10.1016/j.carbpol.2016.07.019.

[27] M.V. Nagarpita, P. Roy, S.B. Shruthi, R.R.N. Sailaja, Synthesis and swelling characteristics of chitosan and CMC grafted sodium acrylate-co-acrylamide using modified nanoclay and examining its efficacy for removal of dyes, Int. J. Biol. Macromol. 102 (2017) 1226–1240, https://doi.org/10.1016/j.ijbiomac.2017.04.099.

[28] K. Kalantari, E. Mostafavi, B. Saleh, P. Soltantabar, T.J. Webster, Chitosan/PVA hydrogels incorporated with green synthesized cerium oxide nanoparticles for wound healing applications, Eur. Polym. J. 134 (June) (2020), 109853, https://doi.org/10.1016/j.eurpolymj.2020.109853.

[29] S. Batool, Z. Hussain, M.B.K. Niazi, U. Liaqat, M. Afzal, Biogenic synthesis of silver nanoparticles and evaluation of physical and antimicrobial properties of Ag/PVA/starch nanocomposites hydrogel membranes for wound dressing application, J. Drug Deliv. Sci. Technol. 52 (April) (2019) 403–414, https://doi.org/10.1016/j.jddst.2019.05.016.

[30] S. Ganguly, N.C. Das, Synthesis of a novel pH responsive phyllosilicate loaded polymeric hydrogel based on poly (acrylic acid-co-N-vinylpyrrolidone) and polyethylene glycol for drug delivery: modelling and kinetics study for the sustained release of an antibiotic drug, RSC Adv. 5 (24) (2015) 18312–18327.

[31] A. Tanwar, P. Date, D. Ottoor, ZnO NPs incorporated gelatin grafted polyacrylamide hydrogel nanocomposite for controlled release of ciprofloxacin, Colloid Interface Sci. Commun. 42 (December 2020) (2021), 100413, https://doi.org/10.1016/j.colcom.2021.100413.

[32] I. Gholamali, M. Yadollahi, Doxorubicin-loaded carboxymethyl cellulose/Starch/ZnO nanocomposite hydrogel beads as an anticancer drug carrier agent, Int. J. Biol. Macromol. 160 (2020) 724–735, https://doi.org/10.1016/j.ijbiomac.2020.05.232.

[33] S.R. Brintha, M. Ajitha, Synthesis and characterization of ZnO nanoparticles via aqueous solution, sol-gel and hydrothermal methods, IOSR J. Appl. Chem. 8 (11) (2015) 66–72, https://doi.org/10.9790/5736-081116672.

[34] M.A. Islam, I. Khan, ZnO/Ag composite nanoparticles for surface plasmon resonance based sensor application in UV-Vis region, Int. J. Adv. Res. Technol. 1 (5) (2012).

[35] S. Nilavazhagan, D. Anbuselvan, A. Santhanam, N. Chidhambaram, Effect of an alkali hydroxide concentration on the structural, optical, and surface morphological properties of ZnO nanoparticles, Appl. Phys. A Mater. Sci. Process. 126 (4) (2020) 1–8, https://doi.org/10.1007/s00339-020-3462-3.

[36] R. Nithya, N. Meenakshi Sundaram, Biodegradation and cytotoxicity of ciprofloxacin-loaded hydroxyapatite-polycaprolactone nanocomposite film for sustainable bone implants, Int. J. Nanomed. 10 (October 2015) (2015) 119–127, https://doi.org/10.2147/IJN.S79995.

[37] S.A. Sadeek, W.H. El-Shwiniy, W.A. Zordok, A.M. El-Didamony, Spectroscopic, structure and antimicrobial activity of new Y(III) and Zr(IV) ciprofloxacin, Spectrochim. Acta - Part A Mol. Biomol. Spectrosc. 78 (2) (2011) 854–867, https://doi.org/10.1016/j.saa.2010.12.048.

[38] A.C. Mohan, B. Renjanadevi, Preparation of zinc oxide nanoparticles and its characterization using scanning electron microscopy (SEM) and X-ray diffraction (XRD), Procedia Technol. 24 (2016) 761–766, https://doi.org/10.1016/j.protcy.2016.05.078.

[39] V. Srivastava, D. Gusain, Y.C. Sharma, Synthesis, characterization and application of zinc oxide nanoparticles (n-ZnO), Ceram. Int. 39 (8) (2013) 9803–9808, https://doi.org/10.1016/j.ceramint.2013.04.110.

[40] S. Murali, S. Kumar, J. Koh, S. Seena, P. Singh, A. Ramalho, A.J. Sobral, Bio-based chitosan/gelatin/Ag@ZnO bionanocomposites: synthesis and mechanical and antibacterial properties, Cellulose 26 (9) (2019) 5347–5361, https://doi.org/10.1007/s10570-019-02457-2.

[41] M. Arab, M. Jallab, M. Ghaffari, E. Moghbelli, M.R. Saeb, Synthesis, rheological characterization, and antibacterial activity of polyvinyl alcohol (PVA)/ zinc oxide nanoparticles wound dressing, achieved under electron beam irradiation, Iran. Polym. J. (Engl. Ed.) (2021) 1019–1028, https://doi.org/10.1007/s13726-021-00952-7.

[42] K. Prusty, S. Patra, S.K. Swain, Nano ZnO imprinted dextran hybrid poly (N-isopropylacrylamide)/poly ethylene glycol composite hydrogels for in vitro release of ciprofloxacin, Mater. Today Commun. 26 (December 2019) (2021), https://doi.org/10.1016/j.mtcomm.2020.101869.

[43] A. Nanodispersion, F. Cuomo, M. Cofelice, and F. Lopez, Rheological Characterization of Hydrogels from, 2019. ⟨https://doi.org/10.3390/polym11020259⟩.

[44] A. Sirelkhatim, S. Mahmud, A. Seeni, N.H.M. Kaus, L.C. Ann, S.K.M. Bakhori, H. Hasan, D. Mohamad, Review on zinc oxide nanoparticles: antibacterial activity and toxicity mechanism, Nano-Micro Lett. 7 (3) (2015) 219–242, https://doi.org/10.1007/s40820-015-0040-x.

[45] M.K. Younis, A.Z. Tareq, I.M. Kamal, Optimization of swelling, drug loading and release from natural polymer hydrogels, IOP Conf. Ser. Mater. Sci. Eng. 454 (1) (2018), https://doi.org/10.1088/1757-899X/454/1/012017.

[46] Z. Zare-Akbari, H. Farhadnejad, B. Furughi-Nia, S. Abedin, M. Yadollahi, M. Khorsand-Ghayeni, PH-sensitive bionanocomposite hydrogel beads based on caborboxymethyl cellulose/ZnO nanoparticle as drug carrier, Int. J. Biol. Macromol. 93 (2016) 1317–1327, https://doi.org/10.1016/j.ijbiomac.2016.09.110.

[47] M.A. Lungan, M. Popa, S. Racovita, G. Hitruc, F. Doroftei, J. Desbrieres, S. Vasiliu, Surface characterization and drug release from porous microparticles based on methacrylic monomers and xanthan, Carbohydr. Polym. 125 (2015) 323–333, https://doi.org/10.1016/j.carbpol.2015.02.058.

[48] J. David, B. Mahanty, Optimized ciprofloxacin release from citric acid crosslinked starch - PVA hydrogel film: modelling with mixture design, J. Polym. Res. (2021) 1–11, https://doi.org/10.1007/s10965-020-02397-7.

[49] I. Permanadewi, A.C. Kumoro, D.H. Wardhani, N. Aryanti, Modelling of controlled drug release in gastrointestinal tract simulation, J. Phys. Conf. Ser. 1295 (1) (2019), https://doi.org/10.1088/1742-6596/1295/1/012063.

arXiv:2302.00889v1 [math.CV] 2 Feb 2023

ON A CLASS OF NON-UNIVALENT FUNCTIONS ASSOCIATED WITH A PARABOLIC REGION

MRIDULA MUNDALIA AND S. SIVAPRASAD KUMAR

Abstract. In the present investigation, we introduce and study the geometric properties of a class of analytic functions, associated with a parabolic region majorly lying in the left-half plane. Further we establish radius and majorization results for the class under study with pictorial illustrations of some of its special cases. Also we derive some sufficient conditions for the class under consideration.

1. Introduction

Let $\mathcal{A}$ be the class of analytic functions $f(z)$ defined on the open unit disc $\mathbb{D} = \{z : |z| < 1\}$ with the normalization $f(0) = 0$ and $f'(0) = 1$. Assume $\mathcal{S} \subset \mathcal{A}$ to be the class of univalent functions. Let $f(z)$ and $g(z)$ be two analytic functions, then $f(z)$ is said to be subordinate to $g(z)$, symbolically $f \prec g$, if there exist a Schwarz function $w(z)$ in $\mathbb{D}$ with $w(0) = 0$, such that $f(z) = g(w(z))$. Additionally, if $g(z)$ is univalent in $\mathbb{D}$, then $f \prec g$ if and only if $f(\mathbb{D}_r) \subset g(\mathbb{D}_r)$, where $\mathbb{D}_r = \{z : |z| < r < 1\}$. Recall that a function $f \in \mathcal{A}$ is starlike if $f(\mathbb{D})$ is starlike with respect to 0. Analytically, a function $f \in \mathcal{A}$ is starlike if

$$\frac{zf'(z)}{f(z)} \prec \frac{1+z}{1-z}.$$

We denote this class by $\mathcal{S}^*$. Ma and Minda [17] gave a unified representation for various subclasses of starlike functions by replacing the superordinate function $(1+z)/(1-z)$ with a more general function $\phi(z)$, and the corresponding class is denoted by $\mathcal{S}^*(\phi)$. Here $\phi(z)$ is chosen such that it is univalent, starlike with respect to $\phi(0) = 1$ and $\operatorname{Re}\phi(z) > 0$ with $\phi'(0) > 0$, also $\phi(\mathbb{D})$ is symmetric about real axis. Several Ma-Minda subclasses have been studied previously (See Table 1). In contrast to $\mathcal{S}^*(\phi)$, Uralegaddi [28] introduced and studied the class

$$\mathcal{M}(\beta) = \{f \in \mathcal{A} : \operatorname{Re}(zf'(z)/f(z)) < \beta,\ \beta > 1\}.$$

Note that $\mathcal{M}(\beta) \nsubseteq \mathcal{S}^*$ and also contains non-univalent functions. In 2006, Ravichandran et al. [25] computed the radius of starlikeness for the class $\mathcal{M}(\beta)$. Motivated by the above class, Kumar et al. [15] made a systematic study of the class $\mathcal{F}(\psi)$ containing non-univalent functions, given by

$$\mathcal{F}(\psi) := \left\{ f \in \mathcal{A} : \frac{zf'(z)}{f(z)} - 1 \prec \psi(z) \right\},$$

where $\psi(z)$ is an analytic univalent function such that $\psi(\mathbb{D})$ is starlike with respect to 0 and $\psi(0) = 0$. In general $\mathcal{F}(\psi) \nsubseteq \mathcal{S}^*(\phi)$. Further if $\phi(z) = 1 + \psi(z) \prec (1+z)/(1-z)$, then $\mathcal{F}(\psi)$ reduces to $\mathcal{S}^*(1+\psi)$. Here below we give a list of functions $\psi_i(z)$, $i = 1, \ldots, 4$, which are considered in

2010 *Mathematics Subject Classification.* 30C45, 30C80.

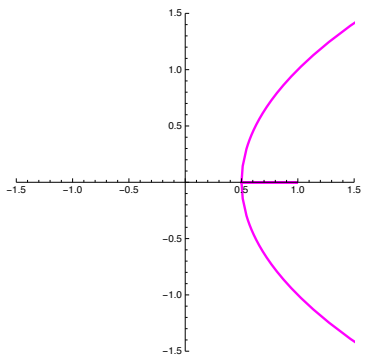
Key words and phrases. Univalent functions, Starlike functions, Logarithmic function, Radius problems, Differential inequality.

The first author is supported by Delhi Technological University, New Delhi.

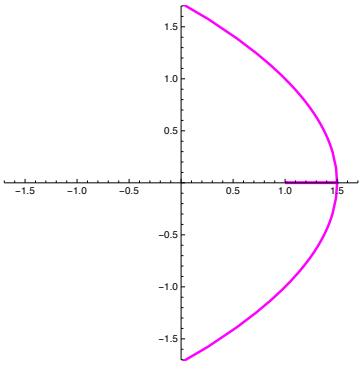
context of the above class $\mathcal{F}(\psi)$,

$$\psi_i(z) = \begin{cases} \gamma z(1+\eta z)^{-2} & , i = 1 \\ z(1-\alpha z^2)^{-1} & , i = 2 \\ z(1-z)^{-1}(1+\beta z)^{-1} & , i = 3 \\ (A-B)^{-1}\log((1+Az)/(1+Bz)) & , i = 4, \end{cases} \tag{1.1}$$

where $A = \alpha e^{i\tau}, B = \alpha e^{-i\tau}$ with $\tau \in (0, \pi/2]$, and $\alpha, \beta, \eta \in (0, 1], \gamma > 0$ (see [12, 15, 16, 19]). Kumar and Gangania in [15], introduced the class $\mathcal{S}_\gamma(\eta) = \mathcal{F}(\psi_1)$ and obtained the radius of starlikeness. Cho et al. [2] dealt with certain sharp radius problems for the class $\mathcal{BS}(\alpha) = \mathcal{F}(\psi_2)$. Infact Masih et al. [19] studied the class $\mathcal{S}_{cs}(\beta) = \mathcal{F}(\psi_3)$, where $0 \leq \beta < 1$, discussed the growth theorem and established sharp estimates of logarithmic coefficients for $0 \leq \beta \leq 1/2$. Further for $1/2 < \beta \leq 1$, the class $\mathcal{S}_{cs}(\beta)$ contains non-univalent functions, infact for $0 \leq \beta \leq 1/2$ the class $\mathcal{S}_{cs}(\beta) \subset \mathcal{S}^*$. In 2022, Kumar et al. [16] introduced the class $\mathcal{F}(A, B) = \mathcal{F}(\psi_4)$ and established some radii results.



(A)



(B)

FIGURE 1. Graphs representing different parabolas with a common focus (1,0) (a) $1 + \mathcal{P}_{0,\pi}(\partial\mathbb{D})$, (b) $1 + \mathcal{P}_{0,0}(\partial\mathbb{D})$.

Motivated essentially by the above classes and observations, we now study a subclass of $\mathcal{A}$ containing non-univalent functions. For $\tau, \theta \in (-\pi, \pi]$, the transformation

$$\omega = ((2e^{i\theta/2}\sqrt{2}/\pi)(\tan^{-1}(e^{i(\tau-\pi/2)}\sqrt{z})))^2$$

maps the boundary of $\mathbb{D}$ onto a parabola (see **Fig.** 1), given by

$$\begin{aligned} \mathcal{P}_{\tau,\theta}(z) &:= \frac{2e^{i(\theta+\pi)}}{\pi^2}\left(\log\left(\frac{1+e^{i\tau}\sqrt{z}}{1-e^{i\tau}\sqrt{z}}\right)\right)^2 \\ &= \frac{8e^{i(\theta+\pi)}}{\pi^2}\sum_{n=1}^{\infty}\left(\frac{e^{2i\tau n}}{n}\sum_{k=0}^{n-1}\frac{1}{2k+1}\right)z^n, \end{aligned}$$

where the branch of $\sqrt{z}$ is chosen so that $\operatorname{Im}\sqrt{z} \geq 0$. Note that the function $\mathcal{P}_{\tau,\theta}(z)$ is univalent in $\mathbb{D}$ and $1 + \mathcal{P}_{0,\pi}(z)$ (see **Fig.** 1(a)) is the parabolic function introduced independently by Ma-Minda [18] and Ronning [26], and used it extensively in context of parabolic starlike functions (see [10, 26]). Further, Kanas [11] discussed some differential subordination techniques and later studied admissibility results [9] involving $1 + \mathcal{P}_{0,\pi}(z)$. Note that the boundary of the function $\mathcal{P}_{0,0}(z)$ is a horizontal parabola with an opening in the left-half plane (see **Fig.** 1(b)). In the present study, we consider the function $1 + \mathcal{P}_{0,0}(z)$, in context of non-univalent functions, which is completely different from the way Ronning, Ma-Minda and Kanas (see [9, 11, 18, 26]) handled parabolic regions. Obviously, the other choices of τ and θ leads to oblique parabolic

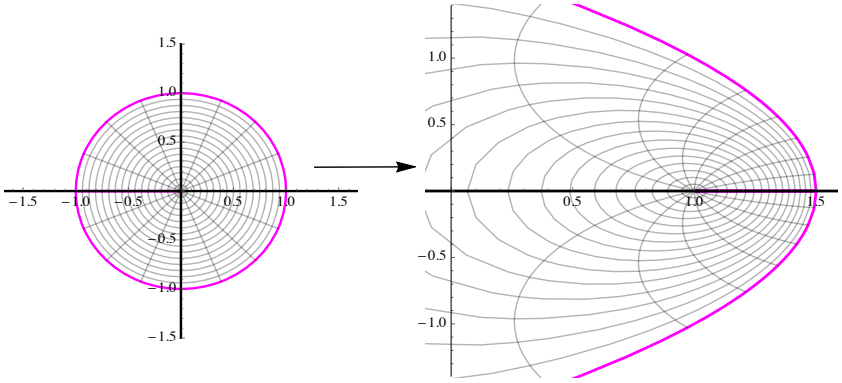


FIGURE 2. Image of open unit disk $\mathbb{D}$ under the mapping $\mathcal{LP}(z)$.

regions, which is beyond the scope of our present study and therefore it is skipped here, as it needs to be handled differently.

For brevity, let us assume $\mathcal{P}_0(z) := \mathcal{P}_{0,0}(z)$ and $\mathcal{LP}(z) := 1 + \mathcal{P}_0(z)$, then the horizontal parabolic region $\mathcal{LP}(z)$ (see **Fig.** 2), is given by

$$\Omega_{\mathcal{LP}} := \{\omega \in \mathbb{C} : (\operatorname{Im}\omega)^2 < 3 - 2\operatorname{Re}\omega \text{ or } |1-\omega| < 2 - \operatorname{Re}\omega\}.$$

Since the major part of the above region lies in the left half plane, it is interesting to find the optimal radius of the domain disc for which it is fully mapped into the right half plane. Here below we define a class of analytic functions consisting of non-univalent functions.

Definition 1.1. Let the class $\mathcal{F}_{\mathcal{LP}}$ consist of functions $f \in \mathcal{A}$ satisfying the subordination

$$\frac{zf'(z)}{f(z)} \prec \mathcal{LP}(z) = 1 - \frac{2}{\pi^2}\left(\log\left(\frac{1+\sqrt{z}}{1-\sqrt{z}}\right)\right)^2.$$

Note that if $f_0 \in \mathcal{F}_{\mathcal{LP}}$, then it can be expressed as

$$f_0(z) = z\left(\exp\int_0^z \frac{\mathcal{P}_0(t)}{t}dt\right), \tag{1.2}$$

which acts as an extremal function for many radius results. Evidently, the subclasses of starlike functions $\mathcal{S}^*_{SG}$, $\mathcal{S}^*_{\mathcal{RL}}$ and $\mathcal{SL}^*(\alpha)$ $(0 \le \alpha < 1)$ (see Table 1) are contained in $\mathcal{F}_{\mathcal{LP}}$, however $\mathcal{F}_{\mathcal{LP}} \nsubseteq \mathcal{S}$. In section 2, we examine some geometrical properties of the function $\mathcal{LP}(z)$.

2. Main Results

2.1. **Geometric Properties of the function $\mathcal{LP}(z)$.** In Lemma 2.1, we establish the maximum and minimum bounds of real part of the function $\mathcal{P}_0(z)$.

Lemma 2.1. *Let $z \in \mathbb{D}_r = \{z : |z| = r\}$, then for each $0 \le r < 1$ and $\alpha \in (-\pi, \pi]$, we have*

$$\mathcal{P}_0(r) \le \operatorname{Re}\mathcal{P}_0(re^{i\alpha}) \le \mathcal{P}_0(-r).$$

Proof. Suppose $z = re^{i\alpha}$, where $-\pi < \alpha \leq \pi$, then for $|z| = r < 1$,

$$\begin{aligned}\operatorname{Re}(\mathcal{P}_0(z)) &= -\frac{2}{\pi^2}\left\{\operatorname{Re}\left(\log\left(\frac{1+\sqrt{z}}{1-\sqrt{z}}\right)\right)^2\right\} \\ &= -\frac{2}{\pi^2}\left(\log\left(\sqrt{\frac{\mu_1(r,c)}{\mu_2(r,c)}}\right)\right)^2 + \frac{2}{\pi^2}\left(\tan^{-1}\left(\frac{2\sqrt{1-c^2}\sqrt{r}}{1-r}\right)\right)^2 \\ &=: \mathcal{G}(r,c).\end{aligned}$$

where $c := \cos(\alpha/2)$ and

$$\mu_i(r,c) := \begin{cases} 1+r+2c\sqrt{r}, & i=1, \\ 1+r-2c\sqrt{r}, & i=2. \end{cases}$$

Observe that $c \in [-1,1]$, infact it is easy to check that $\partial\mathcal{G}(r,c)/\partial c = 0$ if and only if $c = 0$, also $\partial^2\mathcal{G}(r,0)/\partial c^2 < 0$, which leads to

$$\max_{c\in[-1,1]} \mathcal{G}(r,c) = \mathcal{G}(r,0) = \mathcal{P}_0(-r) = 1 + \frac{2}{\pi^2}\left(\tan^{-1}\left(\frac{2\sqrt{r}}{1-r}\right)\right)^2. \tag{2.1}$$

Moreover, for each $R \leq r < 1$, equation (2.1) leads to $\mathcal{G}(r,0) \geq \mathcal{P}_0(r) = \mathcal{G}(r,1)$. Since $\mathcal{G}(r,0)$ is an increasing function, whereas $\mathcal{G}(r,1)$ is a decreasing function of r, this leads to the inequality $\mathcal{G}(r,1) < \mathcal{G}(r,0)$, for each $r < 1$. Hence the required bound is achieved. ∎

As a consequence of Lemma 2.1 and [15, Theorem 2.1 & Corollary 2.2], we obtain the Growth and Covering Theorems for the class $\mathcal{F}_{\mathcal{LP}}$.

Theorem 2.2. *Let $f \in \mathcal{F}_{\mathcal{LP}}$, then the following holds*

I. (Growth Theorem) For $|z| = r < 1$, let

$$\max_{|z|=r} \operatorname{Re}\mathcal{P}_0(z) = \mathcal{P}_0(-r) \text{ and } \min_{|z|=r} \operatorname{Re}\mathcal{P}_0(z) = \mathcal{P}_0(r),$$

then for $|z| = r < 1$ the following sharp inequality holds

$$r\exp\left(\int_0^r \frac{\mathcal{P}_0(t)}{t}dt\right) \leq |f(z)| \leq r\exp\left(\int_0^r \frac{\mathcal{P}_0(-t)}{t}dt\right).$$

II. (Covering Theorem) Suppose $\min_{|z|=r} \operatorname{Re}\mathcal{P}_0(z) = \mathcal{P}_0(r)$ and $f \in \mathcal{F}_{\mathcal{LP}}$. Let f_0 be given by (1.2), then $f(z)$ is a rotation of f_0 or $\{w \in \mathbb{D} : |w| \leq -f_0(-1)\} \subset f(\mathbb{D})$, where $-f_0(-1) = \lim_{r\to 1} -f_0(-1)$.

Remark 2.3. (See **Fig.** 3) If $f(z)$ is of the form $f(z) = z + a_2z^2 + a_3z^3 + \ldots$, belongs to $\mathcal{F}_{\mathcal{LP}}$, then

$$\left|\arg\left(\frac{zf'(z)}{f(z)} - 2\right)\right| > \frac{3\pi}{4}.$$

It can be verified that the equation of tangent corresponding to $\Gamma : y^2 = 1+2(1-x)$ is given by $y = \pm(x-2)$. Infact these tangents intersect the parabola Γ at the points $(1, \pm 1)$. Therefore it can be observed that the convex region $\Omega_{\mathcal{LP}}$ lies in the sector $|\arg(\omega - 2)| > 3\pi/4$. Hence this gives a sharp argument estimate for functions lying in the class $\mathcal{F}_{\mathcal{LP}}$.

Remark 2.4. Due to Lemma 2.1, for $|z| = r < 1$, we have $\mathcal{LP}(r) \leq \operatorname{Re}\mathcal{LP}(z) \leq \mathcal{LP}(-r)$ and infact $\max_{|z|=r}|\mathcal{LP}(z)| = |\mathcal{LP}(r)| = |\mathcal{P}_0(r)|$.

2.2. **Radius Problems for the class $\mathcal{F}_{\mathcal{LP}}$.** Based on the definition of the class $\mathcal{F}_{\mathcal{LP}}$ and pictorial representation of $\mathcal{LP}(\partial\mathbb{D})$ (see **Fig.** 2), we have $\max_{|z|\leq 1} \operatorname{Re}(\mathcal{LP}(z)) = \mathcal{LP}(-1) = 3/2$. This means $\operatorname{Re} zf'(z)/f(z) < 3/2$, thus $f \in \mathcal{F}_{\mathcal{LP}}$ may or may not be a univalent function. Therefore it is an interesting problem to establish the largest radius $r_0 < 1$ such that each $f \in \mathcal{F}_{\mathcal{LP}}$ is starlike in $|z| \leq r_0$. In this section, we study some radius results for the class $\mathcal{F}_{\mathcal{LP}}$ along with the classes $\mathcal{S}^*(\phi)$ and $\mathcal{F}(\psi)$ for some special choices of $\phi(z)$ and $\psi(z)$, as mentioned in Table 1 (see Appendix) and equation (1.1), respectively. Here below we provide a lemma that yields a maximal disc that can be subscribed within the parabolic region $\Omega_{\mathcal{LP}}$.

Lemma 2.5. *Suppose $a < 3/2$ and assume that $\zeta(\eta)$ is defined as follows:*

$$\zeta = \zeta(\eta) = \log\left(\frac{\sqrt{\eta}}{\sqrt{1-\eta}}\right) \text{ with } \eta = \frac{e^{-\pi\sqrt{1-2a}}}{1+e^{-\pi\sqrt{1-2a}}},$$

then $\mathcal{LP}(\mathbb{D})$ satisfies the following inclusion

$$\mathcal{D}(a, r_a) := \{\omega \in \mathbb{C} : |\omega - a| < r_a\} \subset \Omega_{\mathcal{LP}},$$

where

$$r_a = \begin{cases} \sqrt{\left(a - \frac{3}{2} + \frac{2\zeta^2}{\pi^2}\right)^2 + \frac{4\zeta^2}{\pi^2}}, & a \leq \frac{1}{2} \\ \frac{3}{2} - a, & \frac{1}{2} < a < \frac{3}{2}, \end{cases}$$

Proof. We obtain a maximal disc centered at $(a, 0)$, where $a < 3/2$, that can be inscribed within $\Omega_{\mathcal{LP}}$. The distance from center $(a, 0)$ to the boundary $f(\partial(\mathbb{D}))$ is given by square root of

$$\mathcal{D}_a(X) := \left(a + \frac{2}{\pi^2}\left(\log\left(\frac{\sqrt{X^2}}{\sqrt{1-X^2}}\right)\right)^2 - \frac{3}{2}\right)^2 + \frac{4}{\pi^2}\left(\log\left(\frac{\sqrt{X^2}}{\sqrt{1-X^2}}\right)\right)^2,$$

where $X = \cos t$. Now the critical points of $\mathcal{D}_a(X)$ are

$$X' := \begin{cases} \pm\frac{e^{\frac{1}{2}\pi\sqrt{1-2a}}}{\sqrt{1+e^{\pi\sqrt{1-2a}}}}, \pm\frac{e^{-\frac{1}{2}\pi\sqrt{1-2a}}}{\sqrt{1+e^{-\pi\sqrt{1-2a}}}}, & \text{if } a < 1/2, \\ \pm 1/\sqrt{2}, & \text{if } 1/2 \leq a < 3/2. \end{cases}$$

It can be verified that $\mathcal{D}_a''(X) > 0$ at $X = X'$, whenever $a < 3/2$. Therefore, $X = X'$ is the point of minima for $\mathcal{D}_a(X)$, which leads us to the optimal disk centered at a with radius r_a. ∎

Theorem 2.6. *Suppose $0 \leq \alpha < 1$ and $-1 < B < A \leq 1$, then for $f \in \mathcal{A}$, the sharp $\mathcal{F}_{\mathcal{LP}}-$radii for the classes $\mathcal{S}_p^*$, $\mathcal{S}_s^*$, Δ^*, $\mathcal{S}_\varrho^*$, $\mathcal{S}_\rho^*$, $\mathcal{S}_\wp^*$, $\mathcal{BS}^*(\alpha)$, $\mathcal{S}_{\alpha,e}^*$ and $\mathcal{S}^*(A,B)$ (see Table 1 in Appendix) are respectively given by*

(i) $\mathcal{R}_{\mathcal{F}_{\mathcal{LP}}}(\mathcal{S}_p^*) = \tanh^2(\pi/4)$.
(ii) $\mathcal{R}_{\mathcal{F}_{\mathcal{LP}}}(\mathcal{S}_s^*) = \pi/6$.
(iii) $\mathcal{R}_{\mathcal{F}_{\mathcal{LP}}}(\Delta^*) = 5/12$.
(iv) $\mathcal{R}_{\mathcal{F}_{\mathcal{LP}}}(\mathcal{S}_\varrho^*) = (\cosh^{-1}(3/2))^2$.
(v) $\mathcal{R}_{\mathcal{F}_{\mathcal{LP}}}(\mathcal{S}_\rho^*) = \sinh(1/2)$.
(vi) $\mathcal{R}_{\mathcal{F}_{\mathcal{LP}}}(\mathcal{S}_\wp^*) \approx 0.3517\ldots$.
(vii) For $0 < \alpha < 1$, $\mathcal{R}_{\mathcal{F}_{\mathcal{LP}}}(\mathcal{BS}^*(\alpha)) = R_{\mathcal{BS}}$, where

$$R_{\mathcal{BS}} = \begin{cases} 1/2, & \alpha = 0 \\ (\sqrt{1+\alpha} - 1)/\alpha, & 0 < \alpha < 1. \end{cases}$$

(viii) $\mathcal{R}_{\mathcal{F}_{\mathcal{LP}}}(\mathcal{S}_{\alpha,e}^*) = R_{\alpha,e}$, *where*

$$R_{\alpha,e} = \begin{cases} \log(1 - 1/2(\alpha - 1)), & 0 \leq \alpha < 1 - 1/2(e-1) \\ 1, & 1 - 1/2(e-1) \leq \alpha < 1. \end{cases}$$

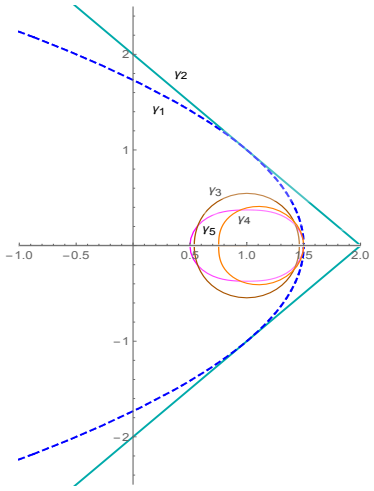
In particular, $\mathcal{R}_{\mathcal{F}_{\mathcal{LP}}}(\mathcal{S}_e^*) = \log(3/2)$.
(ix) $\mathcal{R}_{\mathcal{F}_{\mathcal{LP}}}(\mathcal{S}^*(A,B)) =: \tilde{R}$, *where*

$$\tilde{R} = \begin{cases} 1/(2A-3B), & \text{when } ((-1 < B \leq (2A-1)/3) \wedge (-1 < A < 0)) \\ & \quad \vee ((-1 < B < (2A-1)/3) \wedge (0 \leq A \leq 1)), \\ 1, & \text{when } ((2A-1)/3 < B < A \leq 1) \wedge (-1 < A < 0)) \\ & \quad \vee (((2A-1)/3 \leq B < A \leq 1) \wedge (0 \leq A \leq 1)). \end{cases}$$

Proof. For part (i), as $f \in \mathcal{S}_p^*$, then due to the geometry of the function $1+\mathcal{P}_{0,\pi}(z) = 1+2/\pi^2(\log((1+\sqrt{z})/(1-\sqrt{z})))^2$ it can be observed that

$$\max_{|z|=r} \operatorname{Re}(1+\mathcal{P}_{0,\pi}(z)) = 1+\mathcal{P}_{0,\pi}(r).$$

Now for $f(z)$ to lie in the class $\mathcal{F}_{\mathcal{LP}}$, we must have



$\gamma_1 : \mathcal{LP}(z)$
$1-2/\pi^2(\log((1+\sqrt{z})/(1-\sqrt{z}))^2$
$\gamma_2 : |\arg(u-2)| = 3\pi/4$
γ_3 : Modified Sigmoid function
$2/(1+e^{-z})$
γ_4 : Cardioid function
$1+ze^z$ in $|z|<r^*$,
$r^* \approx 0.351734$
γ_5 : Booth Lemniscate for $\alpha = 0.8$
$1+5z/(5-4z^2)$ in $|z|<r_0$,
$r_0 = |z| < (1/4)(3\sqrt{5}-5)$.

FIGURE 3. Pictorial boundary description of inclusion results pertaining to the parabolic function $\mathcal{LP}(z)$

$$1+\frac{2}{\pi^2}\left(\log\left(\frac{1+\sqrt{r}}{1-\sqrt{r}}\right)\right)^2 \leq \frac{3}{2},$$

which holds provided $r \leq \mathcal{R}_{\mathcal{F}_{\mathcal{LP}}}(\mathcal{S}_p^*)$. Clearly equality in (i) is attained for the function $\tilde{f}(z)$ such that $z\tilde{f}'(z_0)/\tilde{f}(z_0) = 1+\mathcal{P}_{0,\pi}(z_0)$ at $z_0 = \mathcal{R}_{\mathcal{F}_{\mathcal{LP}}}(\mathcal{S}_p^*)$. Further observe that, if $p \in \mathcal{P}[A,B] = \{p(z) : p(z) = 1+c_1z+c_2z^2\ldots \prec (1+Az)/(1+Bz), -1 \leq B < A \leq 1\}$, then for $|z| = r < 1$, it is a known fact that

$$\left|p(z) - \frac{1-ABr^2}{1-B^2r^2}\right| < \frac{|A-B|r}{1-B^2r^2}. \tag{2.2}$$

Then, in view of (2.2), for part (ix) $f \in \mathcal{S}^*(A,B)$ lies in $\mathcal{F}_{\mathcal{LP}}$, if

$$\frac{(A-B)r+1-ABr^2}{1-B^2r^2} \leq \frac{3}{2}.$$

Equivalently, we can say that $(1-Br)((2A-3B)r-1) \leq 0$, provided $r \leq \mathcal{R}_{\mathcal{F}_{\mathcal{LP}}}(\mathcal{S}^*(A,B))$. Extremal function in this case is $\hat{f} \in \mathcal{A}$ satisfying $z_0\hat{f}'(z_0)/\hat{f}(z_0) = (1+Az_0)/(1+Bz_0)$, where $z_0 = \tilde{R}$. For part (iv), as $f \in \mathcal{F}_{\mathcal{LP}}$, then proceeding as before, f lies in $\mathcal{S}_\varrho^*$ if for $|z| = r$ we have $3/2 \geq \cosh\sqrt{r}$, provided $r = \mathcal{R}_{\mathcal{F}_{\mathcal{LP}}}(\mathcal{S}_\varrho^*)$. Sharpness holds for the function $f_\varrho \in \mathcal{A}$ defined as $zf_\varrho'(z)/f_\varrho(z) = \cosh\sqrt{z}$. Further we know that $\max_{|z|=r}\operatorname{Re}(1+\sin z) = 1+\sin r$, then for part (ii) it is enough to find an $r<1$ satisfying the equation $\sin r = 1/2$, thus $r = \mathcal{R}_{\mathcal{F}_{\mathcal{LP}}}(\mathcal{S}_s^*) = \pi/6$. Sharpness holds for the function $f_s(z)$ given by $zf_s'(z)/f_s(z) = 1+\sin z$. In all the subsequent parts, the proofs follow along the same lines, therefore they are omitted. ∎

Let $\mathcal{P}_\alpha$ consist of functions of the form $p(z) = 1+c_1z+c_2z^2+\ldots$, satisfying $\operatorname{Re} p(z) > \alpha$ for $0 \leq \alpha < 1$, then we say $p(z)$ is a Carathéodory function of order α. Denote $\mathcal{P}(0) =: \mathcal{P}$, commonly known as the class of Carathéodory functions. Further, assume $\mathfrak{P}_{\mathcal{LP}}$ to be the class of functions of the form $p(z) = 1+c_1z+c_2z^2+\ldots$, such that $p(z) \prec \mathcal{LP}(z)$.

Theorem 2.7. *Let $0 \leq \alpha < 1$ and $0 \leq \gamma \leq \gamma_\alpha$, where $\gamma_\alpha = \tanh^2(\pi\sqrt{1-\alpha}/2\sqrt{2})$. If $p \in \mathfrak{P}_{\mathcal{LP}}$, then $p \in \mathcal{P}_\alpha$, i.e $p(z)$ is a Carathéodory function of order α, in the disc $|z| < \gamma_\alpha$.*

Proof. Since $p \in \mathfrak{P}_{\mathcal{LP}}$, then by definition of subordination and Schwarz Lemma, there exists an analytic function $w(z)$ with $|w(z)| \leq |z| < 1$ and $w(0) = 0$, such that $p(z) = \mathcal{LP}(w(z))$. Suppose $w(z) = Re^{i\theta}$ $(-\pi < \theta \leq \pi)$, then $|w(z)| = R \leq |z| = r < 1$. On applying Lemma 2.1 for $p \in \mathfrak{P}_{\mathcal{LP}}$, we get $\operatorname{Re} p(z) \geq \mathcal{LP}(r)$. Further $p \in \mathfrak{P}_{\mathcal{LP}}$ is Carathéodory of order α $(0 \leq \alpha < 1)$ if $\mathcal{LP}(r) \geq \alpha$, provided $r \leq \gamma_\alpha = \tanh^2(\pi\sqrt{1-\alpha}/2\sqrt{2})$. The function $f_0(z)$ given by (1.2), is the extremal. ∎

Upon replacing $p(z)$ with $zf'(z)/f(z)$ in Theorem 2.7, we deduce the next result.

Corollary 2.8. *Let $0 \leq \alpha < 1$ and $0 \leq \gamma \leq \gamma_\alpha$, where γ_α is as defined in Theorem 2.7. If $f \in \mathcal{F}_{\mathcal{LP}}$, then $f(z)$ is starlike of order α in the disc $|z| < \gamma_\alpha$. This result is sharp.*

Remark 2.9. Put $\alpha = 0$ in Theorem 2.7, we get a sharp $\mathcal{P}-$ radius for the class $\mathfrak{P}_{\mathcal{LP}}$. Infact for the class $\mathcal{F}_{\mathcal{LP}}$, Corollary 2.8 gives sharp radius of starlikeness $\gamma_0 = \tanh^2(\pi/2\sqrt{2})$. Moreover, $r = \gamma_0 < 1$ serves as the sharp radius of univalence for the class $\mathcal{F}_{\mathcal{LP}}$.

Theorem 2.10. *Assume $0 < \alpha \leq 1$, then the sharp $\mathcal{S}^*(1+\alpha z)-$ radius for the class $\mathcal{F}_{\mathcal{LP}}$ is the unique positive root $r_\alpha = \tanh^2(\pi\sqrt{\alpha}/2\sqrt{2})$ of the equation*

$$2\left(\log((1+\sqrt{r})/(1-\sqrt{r}))\right)^2 - \alpha\pi^2 = 0, \tag{2.3}$$

where α is the radius of the disc $\{\omega : |\omega - 1| < \alpha\}$.

Proof. In view of Remark 2.4, for the circle $|z| = r < 1$, we have

$$\max_{|z|\leq r<1} |\mathcal{LP}(z)| = 1 - \frac{2}{\pi^2}\left(\log\left(\frac{1+\sqrt{r}}{1-\sqrt{r}}\right)\right)^2 = \mathcal{LP}(r), \tag{2.4}$$

which is a decreasing function. Infact $\mathcal{LP}(r) = 0$ if and only if $r = \tanh^2(\pi/2\sqrt{2}) \approx 0.6469\ldots$. As $f \in \mathcal{F}_{\mathcal{LP}}$, then there exists a Schwarz's function $w(z)$ with $w(0) = 0$, so that

$$\frac{zf'(z)}{f(z)} = \mathcal{LP}(w(z)).$$

Assume $w(z) = Re^{i\theta}$ where $R \leq r < 1$. Now observe that for $0 < \alpha \leq 1$, equation (2.4) yields

$$|\mathcal{LP}(z) - 1| \leq |\mathcal{LP}(R) - 1| \leq |\mathcal{LP}(r) - 1| = |\mathcal{P}_0(r)| \leq \alpha,$$

provided $r \leq \tanh^2(\pi\sqrt{\alpha}/2\sqrt{2}) = r_\alpha$. Further, at $z_0 = r_\alpha$, the function $f_0(z)$ (defined in (1.2)) such that $zf_0'(z)/f_0(z) = \mathcal{LP}(z)$, works as the extremal function. ∎

As a consequence of Theorem 2.10, $\mathcal{S}^*(1+\alpha z)-$ radii for some well-known Ma-Minda subclasses of starlike functions, namely, $\mathcal{S}_e^*$, $\mathcal{S}_s^*$, $\mathcal{S}_\varrho^*$, $\mathcal{S}_\wp^*$, $\mathcal{S}_\rho^*$, $\mathcal{S}_{SG}^*$ and $\mathcal{S}_{N_e}^*$ (see Table 1 in Appendix) are stated in Corollary 2.11. Moreover, sharpness of Corollary 2.11 is illustrated by **Fig.** 4.

Corollary 2.11. *Let $f \in \mathcal{A}$ belong to $\mathcal{F}_{\mathcal{LP}}$, then the following radii are sharp for the class $\mathcal{F}_{\mathcal{LP}}$, (see **Fig.** 4)*

(i) The $\mathcal{S}_e^*-$radius is $r_1 = \tanh^2(\lambda\pi)$, where $\lambda = (1/2)\sqrt{(e-1)/2e}$.
(ii) The $\mathcal{S}_s^*-$radius is $r_2 = \tanh^2(\pi/\lambda)$, where $\lambda = 2\sqrt{2\csc 1}$.
(iii) The $\mathcal{S}_\varrho^*-$radius is $r_3 = \tanh^2(\pi\lambda/2)$, where $\lambda = \sin(1/2)$.
(iv) The $\mathcal{S}_\wp^*-$radius is $r_4 = \tanh^2(\pi/2\sqrt{2e})$.
(v) The $\mathcal{S}_\rho^*-$radius is $r_5 = \tanh^2(\pi\sqrt{\lambda}/2)$, where $\lambda = (1/2)\sinh^{-1} 1$.
(vi) The $\mathcal{S}_{SG}^*-$radius is $r_6 = \tanh^2(\lambda\pi/2\sqrt{2})$, where $\lambda = \sqrt{(e-1)/(e+1)}$.
(vii) The $\mathcal{S}_{N_e}^*-$radius is $r_7 = \tanh^2(\pi/2\sqrt{3})$.

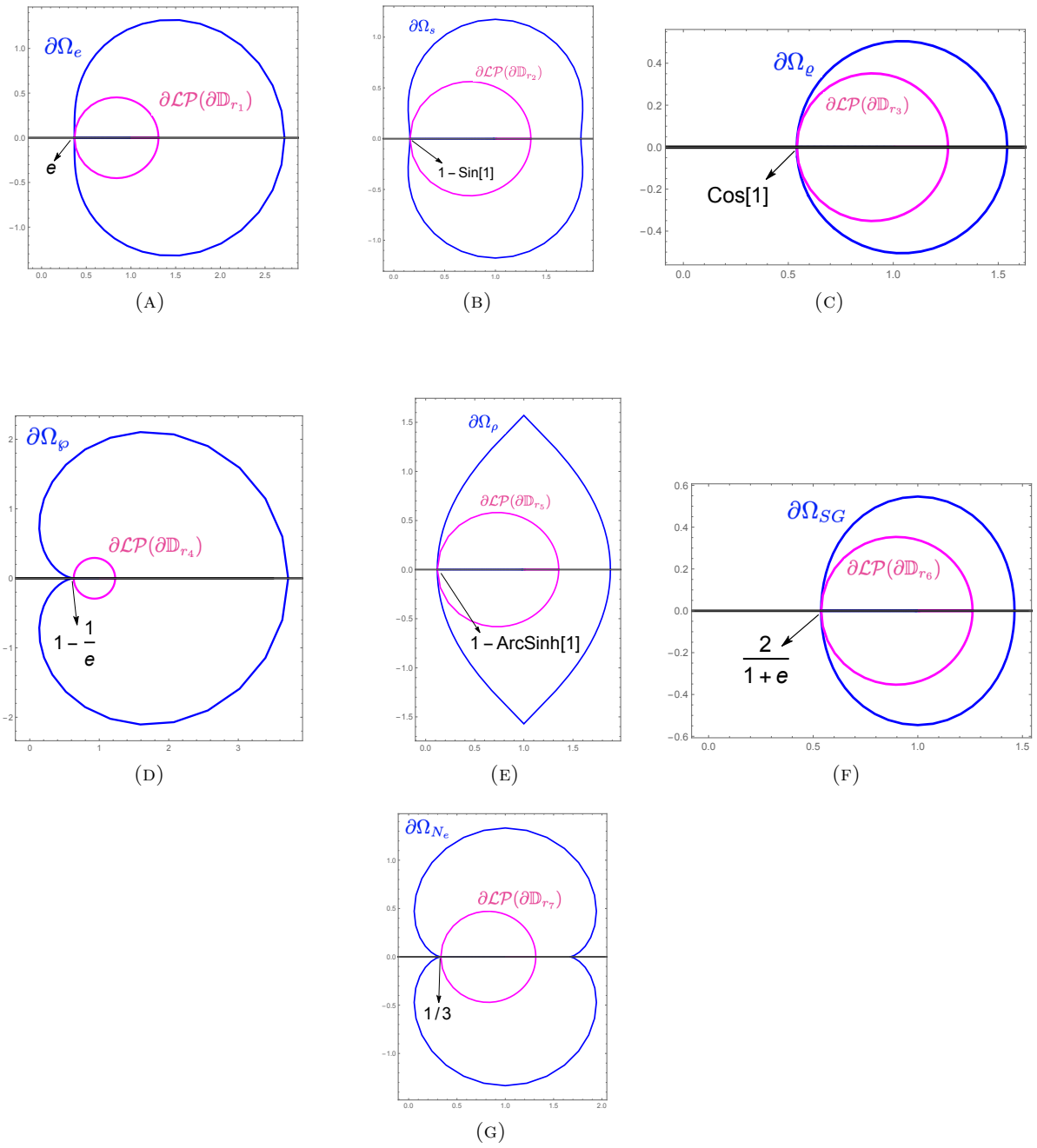


FIGURE 4. Images depicting sharpness of radius result in Corollary 2.11.

In the next corollary we obtain the sharp $\mathcal{S}^*(\beta)-$ radius for the class $\mathcal{F}_{\mathcal{LP}}$.

Corollary 2.12. *Suppose* $0 \leq \beta < 1$ *and* $f \in \mathcal{F}_{\mathcal{LP}}$, *then sharp* $\mathcal{S}^*(\beta)-$ *radius is* $\tanh^2(\pi\sqrt{\beta}/2\sqrt{2})$.

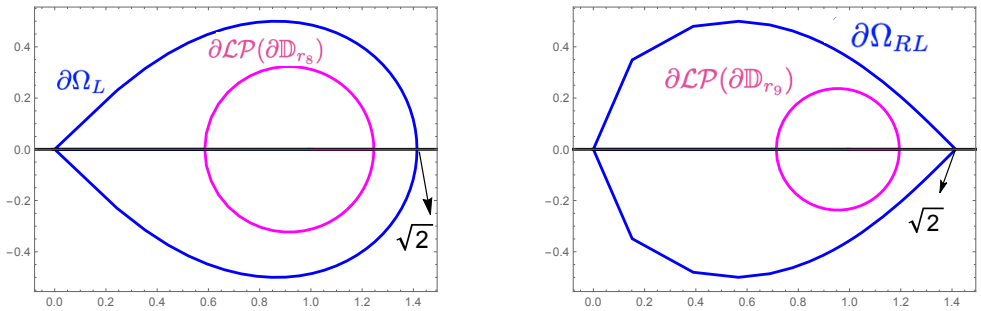
Remark 2.13. On replacing β with $1-\alpha$ $(0 < \alpha \leq 1)$, in Corollary 2.12 we get the radius of starlikness of order α obtained in Corollary 2.8. Moreover, from Corollary 2.12, we obtain the sharp $\mathcal{S}^*_\alpha-$radius for the class $\mathcal{F}_{\mathcal{LP}}$, where $\mathcal{S}^*_\alpha = \{f \in \mathcal{A} : |zf'(z)/f(z) - 1| < 1-\alpha\}$.

Corollary 2.14. *Let* $\eta = \sqrt{2}-1$ *and suppose* $f \in \mathcal{F}_{\mathcal{LP}}$, *then the following holds (see* ***Fig.*** *5)*

(i) $f \in \mathcal{S}^*_{\mathcal{L}}$ *in* $|z| < \tanh^2\left(\pi\sqrt{\eta}/2\sqrt{2}\right) \approx 0.376\ldots$.

(ii) $f \in \mathcal{S}^*_{\mathcal{RL}}$ *in* $|z| < \tanh^2(\pi \sqrt[4]{\sqrt{2}\,\eta(1-\sqrt{2\eta})}/2\sqrt{2}) \approx 0.283\ldots$.

Remark 2.15. From **Fig.** 5, it is evident that the radii obtained in Corollary 2.14 can be further improved.



(A) (B)

FIGURE 5. Above figures correspond to Corollary 2.14, with radii r_8 and r_9 given by: (a) $r_8 \approx 0.376\ldots$ (b) $r_9 \approx 0.283\ldots$.

We now define the class $\mathfrak{F}$ constructed with the help of ratios of two analytic functions $f, g \in \mathcal{A}$, studied by Mundalia and Kumar [21].

$$\mathfrak{F} = \left\{ f \in \mathcal{A} : \operatorname{Re}\left(\frac{f(z)}{g(z)}\right) > 0 \ \&\ \operatorname{Re}\left(\frac{(1-z)^{1+A} g(z)}{z}\right) > 0, -1 \le A \le 1 \right\}.$$

In patricular, for $A = -1$ and $A = 1$, the class $\mathfrak{F}$ reduces to the following classes, respectively,

$$\mathfrak{F}_1 = \left\{ f \in \mathcal{A} : \operatorname{Re}\left(\frac{f(z)}{g(z)}\right) > 0 \ \&\ \operatorname{Re}\left(\frac{g(z)}{z}\right) > 0 \right\}$$

and

$$\mathfrak{F}_2 = \left\{ f \in \mathcal{A} : \operatorname{Re}\left(\frac{f(z)}{g(z)}\right) > 0 \ \&\ \operatorname{Re}\left(\frac{(1-z)^2 g(z)}{z}\right) > 0 \right\}.$$

For proving the next theorem, we require the following lemma of Ravichandran et al. [24].

Lemma 2.16. *If* $p \in \mathcal{P}_n[A, B]$, *then for* $|z| = r$

$$\left| p(z) - \frac{1 - ABr^{2n}}{1 - B^2 r^{2n}} \right| \le \frac{|A - B| r^n}{1 - B^2 r^{2n}}.$$

Particularly, if $p \in \mathcal{P}_n(\alpha)$, *then*

$$\left| p(z) - \frac{1 + (1-2\alpha) r^{2n}}{1 - r^{2n}} \right| \le \frac{2(1-\alpha) r^n}{1 - r^{2n}}.$$

Theorem 2.17. *Let* $-1 \le A \le 1$, *and suppose* $f \in \mathcal{F}_{\mathcal{LP}}$, *then the sharp* $\mathfrak{F}-$*radius is given by*

$$\mathcal{R}_{\mathfrak{F}}(\mathcal{F}_{\mathcal{LP}}) = \frac{1}{2A+3}\left(\sqrt{A^2 + 12A + 28} - (5 + A)\right) =: R_{\mathfrak{F}}.$$

Proof. Since $f \in \mathfrak{F}$, then by definition of the class $\mathfrak{F}$, we have $f(z) = p_1(z) g(z)$ and $g(z) = z p_2(z)(1 - z)^{-(1+A)}$, where for each $i = 1, 2$, $p_i : \mathbb{D} \to \mathbb{C}$ are analytic functions such that $p_i(0) = 1$ and $\operatorname{Re} p_i(z) > 0$. This leads to $f(z) = z p_1(z) p_2(z)(1 - z)^{-(1+A)}$, and as a consequence of logarithmic differentiation, we obtain

$$\frac{z f'(z)}{f(z)} = \frac{1 + Az}{1 - z} + \frac{z p_1'(z)}{p_1(z)} + \frac{z p_2'(z)}{p_2(z)}.$$

For each $-1 \le A \le 1$, Lemma 2.16 leads to,

$$\left| \frac{z f'(z)}{f(z)} - \frac{1 + Ar^2}{1 - r^2} \right| \le \frac{(5 + A) r}{1 - r^2} = R. \tag{2.5}$$

Further, for each $|z|= r \leq R_{\mathfrak{F}}$, one can observe that

$$\frac{1}{2} \leq 1 \leq a = \frac{1+Ar^2}{1-r^2} \leq \frac{1+AR_{\mathfrak{F}}^2}{1-R_{\mathfrak{F}}^2} < \frac{3}{2}. \tag{2.6}$$

Infact inequalities (2.5) and (2.6) yields the inequality,

$$\frac{(5+A)r}{1-r^2} \leq \frac{3}{2} - \frac{1+Ar^2}{1-r^2},$$

provided $r \leq R_{\mathfrak{F}}$. Due to Lemma 2.5, it is clear that the disc $|u-a|< R$ lies in $\Omega_{\mathcal{LP}}$. Further, at $z_0 = R_{\mathfrak{F}}$ the function $f_{\mathfrak{F}}(z)$ defined as $f_{\mathfrak{F}}(z) = z(1+z)^2/(1-z)^{3+A}$ acts as the extremal function. ∎

Corollary 2.18. *Let $f \in \mathcal{F}_{\mathcal{LP}}$, then sharp $\mathfrak{F}_1-$ radius and $\mathfrak{F}_2-$ radius for the class $\mathcal{F}_{\mathcal{LP}}$ are respectively given as*

(i) $\mathcal{R}_{\mathfrak{F}_1}(\mathcal{F}_{\mathcal{LP}}) = \sqrt{17} - 4 \approx 0.123...$
(ii) $\mathcal{R}_{\mathfrak{F}_2}(\mathcal{F}_{\mathcal{LP}}) = (\sqrt{41} - 6)/5 \approx 0.080...$

Theorem 2.19. *Let $\delta = (\pi\sqrt{\beta-1}/\sqrt{2})$, where $1 < \beta < 3/2$, and suppose $f \in \mathcal{F}_{\mathcal{LP}}$, then $\mathcal{M}(\beta)-$ radius is $r_\beta = 1 + 2\left(\cot\delta\right)^2 - 2|\sec\delta/(\tan^2\delta)|$.*

Proof. From Lemma 2.1, it can be viewed that

$$\operatorname{Re}\mathcal{LP}(z) \leq \mathcal{LP}(-r) = 1 - \frac{2}{\pi^2}\left(\log\left(\frac{1+i\sqrt{r}}{1-i\sqrt{r}}\right)\right)^2 = 1 + \frac{2}{\pi^2}\left(\tan^{-1}\left(\frac{2\sqrt{r}}{1-r}\right)\right)^2.$$

As $f \in \mathcal{F}_{\mathcal{LP}}$, then assume that $zf'(z)/f(z) = p(z)$. Due to the above inequality $\operatorname{Re} p(z) \leq \mathcal{LP}(-r)$. Moreover, $\mathcal{LP}(-r) \leq \beta$ provided $r \leq r_\beta$, where r_β is the root of the equation $(1-\beta)\pi^2+2(\tan^{-1}(2\sqrt{r}/(1-r)))^2 = 0$ for $1 < \beta < 3/2$. Equality here occurs for the function $f_0 \in \mathcal{A}$, given by (1.2). ∎

If $f(z)$ and $g(z)$ be analytic functions in $|z|< r$, then $f(z)$ is said to be majorized by $g(z)$, denoted as $f(z) \ll g(z)$, in $|z|< r$, if $|f(z)|\leq |g(z)|$ in $|z|< r$. Equivalently, a function $f(z)$ is said to be be majorized by $g(z)$, if there exists an analytic $\Psi(z)$ with $|\Psi(z)|\leq 1$ in $\mathbb{D}$ and $f(z) = \Psi(z)f(z)$ for all $z \in \mathbb{D}$. For recent update on majorization for starlike and convex function, see [4, 5]. In the next theorem, we determine sharp majorization radius for the class $\mathcal{F}_{\mathcal{LP}}$.

Theorem 2.20. *Let $f \in \mathcal{A}$ and suppose that $g \in \mathcal{F}_{\mathcal{LP}}$. Further assume that $f(z)$ is majorized by $g(z)$ in $\mathbb{D}$, i.e $f(z) \ll g(z)$, then for $|z|\leq r_m \approx 0.4220\ldots$,*

$$|f'(z)|\leq |g'(z)|,$$

where r_m is the unique positive root of the following equation

$$2\pi^2 r - (1-r^2)(\pi^2 - 2(\log((1+\sqrt{r})/(1-\sqrt{r})))^2) = 0. \tag{2.7}$$

Proof. Suppose $0 \leq r < r^* = \tanh^2(\pi/2\sqrt{2}) \approx 0.646\ldots$, then due to Remark 2.9 we conclude that, $g \in \mathcal{F}_{\mathcal{LP}}$ qualifies to be a Ma-Minda type function in $|z|< r^*$. Further let $w(z)$ be a Schwarz function in $\mathbb{D}$ with $w(0) = 0$, then by definition of subordination,

$$\frac{zg'(z)}{g(z)} = \mathcal{LP}(w(z)).$$

Note that for each $|z|= r < 1$, the inequality $|\mathcal{LP}(w(z))|\leq |\mathcal{LP}(r)|$ holds. Now for $|z|= r < r^*$, we obtain

$$\left|\frac{g(z)}{g'(z)}\right| = \frac{|z|}{|\mathcal{LP}(z)|} \leq \frac{r}{1-|\mathcal{P}_0(r)|} = \frac{r}{\mathcal{LP}(r)}. \tag{2.8}$$

As $f(z)$ is majorized by $g(z)$ in $\mathbb{D}$, we find from the definition of majorization,

$$f(z) = \psi(z)g(z).$$

Upon differentiating the above equality and suitable rearrangement of terms, we obtain

$$f'(z) = g'(z)\left(\psi'(z)\frac{g(z)}{g'(z)} + \psi(z)\right). \tag{2.9}$$

Additionally, as a result of the Schwarzian inequality, we have

$$|\psi'(z)|\leq \frac{1-|\psi(z)|^2}{1-|z|^2}, \quad (z\in\mathbb{D}). \tag{2.10}$$

Moreover, from equations (2.8)-(2.10), we deduce

$$|f'(z)|\leq \left(|\psi(z)|+\frac{r(1-|\psi(z)|^2)}{(1-r^2)(\mathcal{LP}(r))}\right)|g'(z)|.$$

Substituting $|\psi(z)|=\sigma$ $(0\leq\sigma\leq 1)$, results in

$$|f'(z)|\leq \Psi(r,\sigma)|g'(z)|,$$

where

$$\Psi(r,\sigma)=\sigma+\frac{r(1-\sigma^2)}{(1-r^2)(\mathcal{LP}(r))}.$$

We need to determine $r_m\leq r^*$ so that

$$r_m=\max\{r\in[0,r^*]:\Psi(r,\sigma)\leq 1\ \forall\sigma\in[0,1]\}.$$

Equivalently if $\Phi(r,\sigma):=(1-r^2)(\mathcal{LP}(r))-r(1+\sigma)$ then we need to determine

$$r_m=\max\{r\in[0,r^*]:\Phi(r,\sigma)\geq 0\ \forall\sigma\in[0,1]\}.$$

Since $\partial\Phi/\partial\sigma=-r<0$, then $\max_{\sigma\in[0,1]}\Phi(r,\sigma)=\Phi(r,0)=:\phi_0(r)$. Further it is evident that, as $\phi_0(0)=1>0$ and $\phi_0(r^*)=-r^*<0$, then there exists $r_m\leq r^*$, a smallest positive root of the equation given in (2.7) such that $\phi_0(r)\geq 0$ for each $r\in[0,r_m]$. This completes the proof. ∎

In 2017, Peng and Zhong [22], introduced the class $\Omega\subset\mathcal{A}$, defined as

$$\Omega=\{f\in\mathcal{A}:|zf'(z)-f(z)|<1/2\}.$$

We conclude this section by determining sharp $\Omega-$radius for the class $\mathcal{F}_{\mathcal{LP}}$.

Theorem 2.21. *Let $f\in\mathcal{F}_{\mathcal{LP}}$, then $f\in\Omega$ in $|z|<r_{\mathcal{L}}\approx 0.522\ldots$ is the smallest positive root of*

$$4f_0(r)(\log((1+\sqrt{r})/(1-\sqrt{r})))^2=\pi^2$$

and

$$g_0(z)=z\left(\exp\int_0^z\frac{\mathcal{P}_0(-t)}{t}dt\right)=z+\frac{8}{\pi^2}z^2-\frac{8}{3\pi^4}(\pi^2-12)z^3+\frac{8}{135\pi^6}(1440-360\pi^2+23\pi^4)z^4-\cdots. \tag{2.11}$$

This is a sharp estimate.

Proof. Since $f\in\mathcal{F}_{\mathcal{LP}}$, then as a consequence of Remark 2.4 for $|z|=r<1$, we have

$$\left|\frac{zf'(z)}{f(z)}-1\right|<|\mathcal{LP}(r)-1|=|\mathcal{P}_0(r)|.$$

Due to the growth theorem as mentioned in [15, Theorem 1] and Theorem 2.2, we observe that $|f(z)|\leq g_0(r)$, where $g_0(r)$ is given by (2.11). Further

$$|zf'(z)-f(z)|=|f(z)|\left|\frac{zf'(z)}{f(z)}-1\right|\leq g_0(r)|\mathcal{P}_0(r)|.$$

Thus $g_0(r)|\mathcal{P}_0(r)|\leq 1/2$ provided $|z|<r_{\mathcal{L}}\approx 0.522864$. Hence the result is established. ∎

2.3. **Sufficient Conditions for the class $\mathcal{F}_{\mathcal{LP}}$.** In this section, we establish some sufficient conditions for the class $\mathcal{F}_{\mathcal{LP}}$. Below, we state a Lemma given by Jack [7], which is utilised in context of the class under consideration.

Lemma 2.22. [7, Lemma 1, p.470] *Let $\nu(z)$ be a non-constant analytic function in $\mathbb{D}$, such that $\nu(0) = 0$. If $|\nu(z)|$ attains its maximum value on the circle $|z|= r$ at a point z_0, then $z_0\nu'(z_0) = k\nu(z_0)$, where k is real and $k \geq 1$.*

Theorem 2.23. *Suppose $0 \leq t \leq 1$ and let $f \in \mathcal{A}$ satisfy the following differential inequality*

$$\left| t\left(1 + \frac{zf''(z)}{f'(z)}\right) + (1-t)\frac{zf'(z)}{f(z)} - 1 \right| < \frac{1}{6}(3+2t), \quad z \in \mathbb{D}, \tag{2.12}$$

then $f \in \mathcal{F}_{\mathcal{LP}}$.

Proof. Consider an analytic function $\nu(z)$ with $\nu(0) = 0$. Assume $f \in \mathcal{A}$ such that

$$\frac{zf'(z)}{f(z)} - 1 = \frac{1}{2}\nu(z).$$

We show that $|\nu(z)|< 1$ in $\mathbb{D}$. Suppose on the contrary $|\nu(z)|\geq 1$, then by an application of Lemma 2.22, there exists $z_0 \in \mathbb{D}$ such that for $k \geq 1$, $|\nu(z_0)|= 1$ and $z_0\nu'(z_0) = k\nu(z_0)$. Substituting $\nu(z_0) = e^{i\mu}$, $-\pi < \mu \leq \pi$ leads to

$$\begin{aligned}
&\left| t\left(1 + \frac{zf''(z)}{f'(z)}\right) + (1-t)\frac{zf'(z)}{f(z)} - 1 \right| \\
&= \left| t\left(1 + \frac{\nu(z_0)}{2} + \frac{k\nu(z_0)}{2+\nu(z_0)}\right) + (1-t)\left(1 + \frac{\nu(z_0)}{2}\right) - 1 \right| \\
&= \left| \frac{tke^{i\mu}}{2+e^{i\mu}} + \frac{e^{i\mu}}{2} \right| \geq \frac{1}{6}(3+2t).
\end{aligned}$$

This is contradiction to the assumption given in (2.12). Thus $|\nu(z)|< 1$, which means that $zf'(z)/f(z)$ lies in the disc $|(zf'(z)/f(z)) - 1|< 1/2$. Hence in view of Lemma 2.5, (with $a = 1$) required result is achieved. ∎

For $t = 1/2$, $t = 0$ and $t = 1$ in Theorem 2.23, we obtain the following corollary,

Corollary 2.24. *Let $f \in \mathcal{A}$ satisfy the following differential inequalities*

(i) $|(zf'(z)/f(z) + zf''(z)/f'(z)) - 1| < 4/3$, *or*
(ii) $|(zf'(z)/f(z)) - 1| < 1/2$, *or*
(iii) $|zf''(z)/f'(z)| < 5/6$,

then $f \in \mathcal{F}_{\mathcal{LP}}$.

Conclusion

In the present investigation, we introduce a class of analytic functions associated with certain parabolic region. In particular, we have considered a case when parabola is lying majorly in the left half plane and symmetric about real axis. The other cases namely, oblique parabolic regions are still open, for similar investigations.

Appendix

The classes we come across in the present investigation are listed below for the ready reference of the reader.

TABLE 1. Special subclasses of Ma-Minda starlike functions for specific choices of $\phi(z)$

Class $\mathcal{S}^*(\phi)$	$\phi(z)$	$\phi(\mathbb{D})$	**References**
$\mathcal{S}^*_{\alpha,e}$	$\alpha+(1-\alpha)e^z$	$\Omega_{\alpha,e}$	[13] Khatter et al.
$\mathcal{SL}^*(\alpha)$	$\alpha+(1-\alpha)\sqrt{1+z}$	$\Omega_{\alpha,L}$	[13] Khatter et al.
$\mathcal{S}^*_{\wp}$	$1+ze^z$	$\Omega_{\wp}$	[14] Kumar et al.
$\mathcal{S}^*_{SG}$	$2/(1+e^{-z})$	Ω_{SG}	[6] Goel et al.
$\mathcal{S}_s$	$1+\sin z$	Ω_s	[3] Cho et al.
$\mathcal{S}^*_{\rho}$	$1+\sinh^{-1}z$	Ω_{ρ}	[1] Arora et al.
$\mathcal{S}^*_{\varrho}$	$\cosh\sqrt{z}$	Ω_{ϱ}	[21] Mundalia et al.
Δ^*	$z+\sqrt{1+z^2}$	Ω_{Δ}	[23] Raina et al.
$\mathcal{S}^*_{\mathcal{L}}$	$\sqrt{1+z}$	Ω_L	[27] Sokół et al.
$\mathcal{S}^*(A,B)$	$(1+Az)/(1+Bz)$	$\Omega_{A,B}$	[8] Janowski
$\mathcal{S}^*(N_e)$	$1+z-z^3/3$	Ω_{N_e}	[29] Wani et al.
$\mathcal{S}^*_p$	$1+(2/\pi^2)(\log((1+\sqrt{z})/(1-\sqrt{z})))^2$	Ω_p	[26] Ronning
$\mathcal{S}^*_{\mathcal{RL}}$	$\sqrt{2}-\left(\sqrt{2}-1\right)\sqrt{(1-z)(1+2\left(\sqrt{2}-1\right)z)}$	Ω_{RL}	[20] Mendiratta et al.

REFERENCES

[1] Arora K., Kumar, S. S.: Starlike functions associated with a petal shaped dimain, Bull. Korean Math Soc. **59**(4), 903-1010 (2022)

[2] Cho, N. E., Kumar, S., Kumar, V., Ravichandran, V.: Differential subordination and radius estimates for starlike functions associated with the Booth lemniscate, Turkish J. Math. **42**, 1380-1399 (2018)

[3] Cho, N. E., Kumar, V., Kumar, S. S., Ravichandran, V.: Radius problems for starlike functions associated with the sine function. Bull. Iranian Math. Soc. **45**(1), 213-232 (2019)

[4] Gangania, K., Kumar, S. S.: On Certain Generalizations of $\mathcal{S}^*(\psi)$ Comput. Methods Funct. Theory. (2021). https://doi.org/10.1007/s40315-021-00386-5

[5] Tang, H., Deng, G.: Majorization problems for some subclasses of starlike functions. J. Math. Res. Appl. **39**, 153-159 (2019)

[6] Goel, P., Kumar, S. S.: Certain class of starlike functions associated with modified sigmoid function. Bull. Malays. Math. Sci. Soc. **43**(1), 957-991 (2020)

[7] Jack, I. S.: Functions starlike and convex of order α. J. London Math. Soc. (Ser. 2)2, 469-474 (1971)

[8] Janowski, W.: Some extremal probles for certain families of analytic functions. I. Ann. Polon. Math. **28**(3), 297-326 (1973)

[9] Kanas, S.: Differential subordination related to conic sections, J. Math. Anal. Appl., **317**, 650-658 (2006)

[10] Kanas, S., Wiśniowska, A.: Rev. Roumaine Math. Pures Appl. **45**, no. 4, 647-657 (2000)

[11] Kanas, S.: Techniques of the differential subordination for domains bounded by conic sections, Int. J. Math. Math. Sci., **38**, 2389-2400 (2003)

[12] Kargar, R., Ebadian, A., Sokół, J.: On Booth lemniscate and starlike functions. Anal. Math. Phys. **9**(1), 143-154 (2019)

[13] Khatter, K., Ravichandran, V., Sivaprasad Kumar, S.: Starlike functions associated with exponential function and the lemniscate of Bernoulli. Rev. R. Acad. Cienc. Exactas Fis. Nat. Ser. A Math. RACSAM **113**(1), 233-253 (2019)

[14] Kumar, S. S., Kamaljeet, G.: A cardioid domain and starlike functions. Anal. Math. Phys. **11**(2), 1-34 (2021)

[15] Kumar, S. S., Gangania, K.: On geometrical properties of certain analytic functions, Iran. J. Sci. Technol. Trans. A Sci. **45**(4), 1437-1445 (2021)

[16] Kumar, S. S., Yadav, P.: Starlike Functions asociated with a Non-Ma-Minda Function. (2022). Preprint arXiv:2208.01245

[17] Ma, W. C., Minda, D.: A unified treatment of some special classes of univalent. functions. In: Proceedings of the Conference on Complex Analysis (Tianjin, 1992), pp. 157-169, Conf. Proc. Lecture Notes Anal., I, Int. Press, Cambridge, MA

[18] Ma, W., Minda, D.: Uniformly convex functions II, Ann. Polon. Math., **58**, 275-285 (1993)

[19] Masih, V. S., Ebadian, A., Yalçin, S.: Some properties associated to a certain class of starlike functions, Math. Slovaca **69**(6), 1329-1340 (2019)
[20] Mendiratta, R., Nagpal, S., Ravichandran, V.: A subclass of starlike functions associated with left-half of the lemniscate of Bernoulli. Int. J. Math. **25**(9), 17 (2014)
[21] Mundalia, M., Kumar, S. S.: On a subfamily of starlike functions related to Hyperbolic Cosine function. (2022) preprint arXiv:2201.05811
[22] Peng, Z., Zhong, G.: Some properties for certain classes of univalent functions defined by differential inequalities, Acta Math. Sci. Ser. B (Engl. Ed.) **37B**, 69-78 (2017)
[23] Raina, R. K., Sokół, J.: Some properties related to a certain class of starlike functions. C. R. Math. Acad. Sci. Paris **353**(11), 973-978 (2015)
[24] Ravichandran, V., Rønning F. Shanmugam, T. N.: Radius of convexity and radius of starlikeness for some classes of analytic functions. Complex Var. Theory Appl. **33**(1-4), 265-280 (1997)
[25] Ravichandran, V., Silverman, H., Hussain Khan, M., Subramanian, K. G.: Radius problems for a class of analytic functions. Demonstratio Math. **39**(1), 67-74 (2006)
[26] Rønning F.: Uniformly convex functions and a corresponding class of starlike functions, Proc. Amer. Math. Soc. **118**(1), 189-196 (1993)
[27] Sokół, J., Stankiewicz, J.: Radius of convexity of some subclasses of strongly starlike functions. Zeszyty Nauk. Politech. Rzeszowskiej Mat. No. **19**, 101-105 (1996)
[28] Uralegaddi, B. A., Ganigi, M. D., Sarangi, S. M.: Univalent functions with positive coefficients. Tamkang J. Math. **25**(3), 225-230 (1994)
[29] Wani, L. A., Swaminathan, A.: Starlike and convex functions associated with a nephroid domain, Bull. Malays, Math. Sci. Soc. **44**(1), 79-104 (2021)

Department of Applied Mathematics, Delhi Technological University, Delhi–110042, India
Email address: mridulamundalia@yahoo.co.in

Department of Applied Mathematics, Delhi Technological University, Delhi–110042, India
Email address: spkumar@dce.ac.in

Optimization and manufacturing of friction drilling tools by using solid works and Ansys

Ebrahim Mohibi[1], Dr. A.K Madan[2]

1, M.Tech student, Department of Mechanical Engineering, Delhi Technological University Delhi.

2, Professor, Department of Mechanical Engineering, Delhi Technological University, Delhi

ABSTRACT

Friction drilling is one of the processes used to produce a hole. Friction drilling is a hole-making method in which a tool rotates around its axis and is used to penetrate into the workpiece and then create a hole in a single step, with no chips formed during the process. This procedure is based on the generation of heat as a result of frictional forces between the workpiece and the friction drilling tool. In this paper, friction drilling tool models with varied tool angles were built using a CAD program (solid works). We started with a 30-angle tool, which was already in existence, and then added three additional tools to increase the overall performance of the item (25 angles, 35angle,50angle). After developing and analyzing each tool assembly under static and thermal stress conditions using the CAD tool (Ansys workbench), sae 1137 material was selected for the tool and Al-5059 material was selected for the plate in each instance.

***Key words**: (solid works, Ansys workbench, al-5059, steel1137, tools with different angles, al-5059plate.)*

Introduction

For the most part, drilling is a basic approach that uses a drill bit to create an irregularly shaped hole in a solid object. Typically, a drill bit is a rotating piecing tool; however, Nemours sometimes uses a multi-point drill bit. After being forced aggressively on the working plate, the bit revolved at speeds of hundreds to thousands of rotations per minute. In order to form a hole, the cutting edge must be hostile to the work place, and the material must be removed like in the case of drilling.

The hole in rock drilling is often not created using a round piecing process; rather, the bit is typically rotated during the procedure. As a replacement, the hole is often formed by pounding a drill bit dip into the hole, which promptly repeated small modifications. Drifter drills were used to make holes for parallel drills, and on rare occasions, other types of bits were used to create holes, including non-circular cross section drill bits that were specifically designed for the special situation. Drifter drills were used to make holes for parallel drills from the outside, or else from the inside.

The size and shape of holes are generally defined by their size and shape after the process is completed; for example, if it has an acute boundary on the entryway, the presence of burrs will be the departure way until they are removed; in the same manner, if it is used from the inside, it could be helical; and this drilling can cause the working plate mechanical properties to create squat lasting stress surrounding the hole opening and a very thin layer of highly stressed and disturbed material on the newly formed surface The workpiece becomes more sensitive to corrosion and fracture propagation at the stressed surface as a result of this increase in susceptibility. It may be necessary to perform a finishing procedure in order to prevent these adverse situations.

The fluting of the drill bit is res ponsible for transporting swarf up and away from the tip of the drill bit during normal operation. As the cutting blades make new chips, the chips migrate outwards from the hole, continuing

the movement of the chips. This works well until the chips pack together too tightly, which may happen when the holes are deeper than typical or when there is inadequate backing off (removing the drill slightly or totally from the hole while drilling). It is occasionally necessary to apply cutting fluid to alleviate this issue and to increase the tool's life by cooling and lubricating the tip and chip flow. When utilizing a gun drill, coolant may be supplied into the drill shank via holes drilled into the shank of the drill. Aluminum is particularly susceptible to gripping the drill bit, thus cutting fluid is used to guarantee a clean and precise hole while also preventing the metal from holding the drill bit during the drilling process. When cutting brass and other soft metals that can grab the drill bit and cause "chatter," a face of approximately 1-2 millimeters can be ground on the cutting edge to create an obtuse angle of 91 to 93 degrees. This is useful when cutting brass and other soft metals that can grab the drill bit and cause "chatter." This avoids "chatter," which is a condition in which the drill rips rather than slices the metal being worked. In this case, however, the drill is pushing the metal away rather than grasping it because of the geometry of the bit cutting edge. High friction and very hot swarf are produced as a result of this.

Oil-hole drills are used in the drill bit for heavy feeds and somewhat deep holes. An emollient fluid is applied to the drill head over a small hole in the bit and unconfined out along the fluting. It is possible to use a standard drill hustle placement in oil-hole drilling; however, it is far more common with instinctive drilling equipment, in which the working plate spins rather than the drill bit.

Peck drilling, also known as interrupted cut drilling, is a procedure performed by a computer numerical control machine that is used to prevent swarf from harming the premises while drilling deep holes (Roughly when the deepest hole is 3 times higher than the drill diameter). Chip breaking is a modified form of this method, and this method is similar to a flash, but it is used when there is a medium range of holes, or else it would generate excessive heat to the drill bit, and it is also applicable when the hole is larger than the diameter of the drill.

It has happened that material could not be transported to the CN machine, in which case a Magnetic Base Drilling Machine was used. Drilling may be done in any direction, including horizontally and even on the ceiling, thanks to the base. In general, cutters are a superior alternative for these types of machines; yet, they may drill considerably more quickly with less speed than other options. These Cutters range in size from 12.5mm to 20cm in diameter and from 3cm to 20cm in depth of cut (depth of cut). These devices were often used in the construction, manufacturing, maritime, and oil and gas sectors, among other fields. Pneumatic magnetic drilling machines were used in the oil and gas sector to prevent explosions and fires, even within specific tube magnetic drilling machines that were attached to pipes of varying diameters.

LITERATURE REVIEW

From the study of Miller S.F Li R., Wang H.& Shih A.J. [1] developed an exploratory investigation and modeling of friction drilling process. Exploratory calculating thrust force and torque under given tool speed were calculated and analyzed. A numerical design was proposed, position on the exposure area between the non-deformed workpiece of AISI 1020 cold-rolled carbon steel and the rigid unworn WC in a Co matrix tool, by applying ANSYS 7 finite element software package. The model estimated thrust force and torque had good relation with exploratory calculations some inconsistency still there and showed the limitation of the rationalize friction model used in this study.

From the papers of Miller SF Blau Pj & Shih AJ. [2] Investigated the having a conical tungsten carbide tool applied for friction drilling of steel015 working piece. In Insertion, the accelerator force and torque while drilling are calculated step by step to calculate the results of tool wear. Exact calculation of tool dimensions showing that the wear was concerted at the tool main region and at the convergence between the conical and cylindrical regions, The tool edge self-sharp while friction drilling, which decreases the accelerator force as tool wear increases. The torque was not display any possible changes at various stages of tool have.

Miller sf.& Shih aj. [3] Calculating 3D defined element modeling for friction drilling of Al6061T6 work-material by A2 air hardening tool steel utilizing ABAQUS/EXPLICIT finite element analysis tool. The model is accepted with differentiated the resulted accelerator force and torque to exploratory standards, the profiles of thrust force finite element modeling were shifted to the left of the experimental profiles due to the noticeable deflection of the sheet Al workpiece at the initial contact stage. This phenomenon was not accurately modeled by finite element modeling and led to the early peak of the thrust force. For torque, the difference of

designing and exploratory also increases at squat tool feed rate. The object is estimated over the torque in the beginning contact stage, but underrate the torque as it reaches its highest. This illustrate the constrain of the Coulomb's friction

Chow HM. Lee SM. &Yang LD. [4] applied Taguchi method to explore how different process parameters such as friction angle (30o, 45o, and 60o), friction contact area ratio (50%, 75%, and 100), feed rate (75, 100, and 125 mm/min), and drilling speed (30, 60, and 90 m/min) would affect the produced hole surface roughness, while drilling AISI 304 stainless steel by sintered carbide drill. The optimal drilling conditions that produced the smallest surface roughness were: 30o friction angle, 50 % friction contact area ratio, 100 mm/min feed rate, and 90 m/min drilling speed.

From the studies of Qu J. and Blau P.J. [5] they were created a recently developed analytical modelling for thermal drilling helpful to know the better friction coefficient and also shear stress values that take place while running the process, originating from exploratory calculated results of the immediate acceleration force and torque of reference [5]. The created profiles of coefficient of friction and shear stress.

Fernández A., Lopez de Lacalle L.N.& Lamikiz A. [6] analyzed, through controlled tests at various rotational speeds and feed rates, the friction drilling of austenitic stainless steel with different thicknesses by tungsten carbide with cobalt matrix tool. Experimentation shows that for higher spindle speeds, higher process temperature is produced. When the temperature of the workpiece increases, the form of the burr is more cylindrical with a greater depth. According to the feed rate, higher feed rate leads to a maximum temperature reduction, because the contact time between tool and workpiece decreases.

Krishna P.V Kishore K.& Satyanarayana V.V. [7] designed an experimental layout and applied Taguchi method to estimate the execution of high speed steel friction drill on Al alloy 6351 workpiece. The selected process parameters are rotational speed (2000 3000 rpm) feed rate (0.1 0.3 mm/rev), and friction angle (45o 90o). Figure 10 shows the variation of torque and thrust force claimed to the selected parameters

K W.L.Hung .L Lee S.M.& Chow H.M. [8] investigated the effects of friction angle (30o, 45o, and 60o), friction contact area ratio (50, 75, and 100 %), feed rate (75, 100, and 125 mm/min), and spindle speed (1200, 2400, and 3600 rpm) while drilling SUS 304 stainless steel by tungsten carbide tool, on the 2 slandered feature surface roughness and bushing length by using Taguchi method.Results showed that friction angle and spindle speed were the significant machining parameters that most intensively affect surface roughness while friction contact area ratio was the only significant parameter for bushing length.

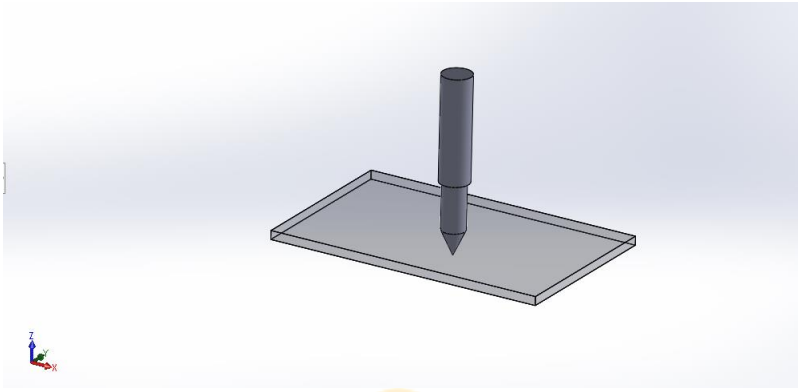
Krasauskas P. [9] performed a statistical 5 variable linear regression analysis in order to evaluate the influence of plate thickness (1.5, 2, and 2.5 mm), rotational speed (2000, 2500, and 3000 rpm), feed rate (60, 100, and 140 mm/min), and materials mechanical properties, by choosing three various materials steel5235, steel4301, and Al5652 alloy, on axial drilling force and torque. The analysis of spindle rotational speed and feed rate influencing on axial force variation showed that minimal spindle speed calls bigger drilling force, while bigger the feed bigger the axial force.

Pantawane P.D. and Ahuja B.B. [10] calculating the Response Surface method to analyzing the consequence of friction drilling input parameters as rotational speed (2.085e3, 2.5e3, 3.5e3, 4.5e3, and 4.914e3 rpm), feed rate (71.36, 90, 135, 180, and 198.64 mm/min), and tool diameter (0.73 and 0.92 cm) on the response dimensional misconception and surface roughness of steel1015 bush by resort to tungsten carbide in a cobalt matrix drilling tools. Optimization using eligibility Function has been used to optimize the machining conditions. The optimum setting which results in maximum desirability found to be as 4500 rpm speed, 71.36 mm/min feed, and 7.3 mm tool diameter. The low desirability M10 tool indicates its poor performance, hence this tool is not recommended

Implementation

To assemble components first need to import all models into assembly window then use coincident option to connect plane to plane or surface to surface to adjust. And some time distance constraint also used while creating it

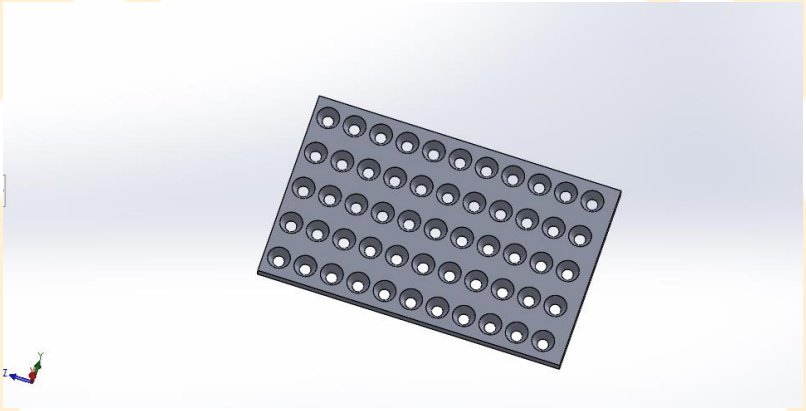
25* drilling tool with plate assembly



Final assembly model of 25* angle drilling tool with plate

To show imaginary holes on the plate, revolve cut option were used with same dimensions which were used while creating in friction drilling tool

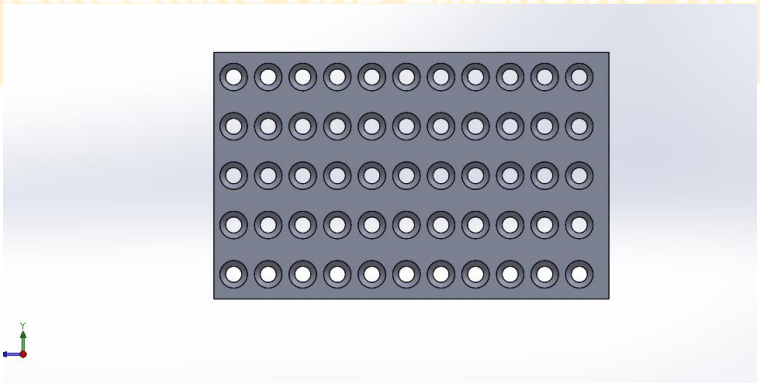
30* angle imaginary drilling holes



Imaginary holes for 30* angle drilling tool

INTERNATIONAL JOURNAL FOR RESEARCH IN ENGINEERING AND EMERGING TRENDS

25* angle imaginary drilling holes



Imaginary holes for 25* angle drilling tool

Steel416 grade chemical Composition

Grade		C	Mn	Si	P	S	Cr	Mo
416	min.	-	-	-	-	0.15	12	-
	max.	0.15	1.25	1	0.06	-	14	0.6

Steel416 Corrosion Resistance properties

This steel416 material is having huge resistant to acids and water and dry air, but they are small corrosion resistant than no free machining steels, and some of grades like austenitic and 430 grades are having chromium percentage nearly 17, due to this high percentage of chromium these steels are very hard to meet high corrosion and smooth surface, steel416 is a free machining grade and it is having huge sulphur content and it not suitable for chloride and marine type of environment

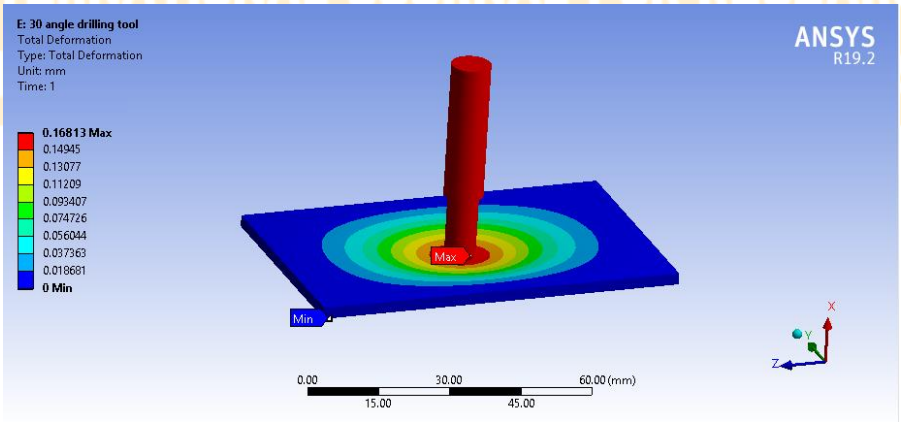
Steel416 material uses

Here below discussing some major important applications of steel416

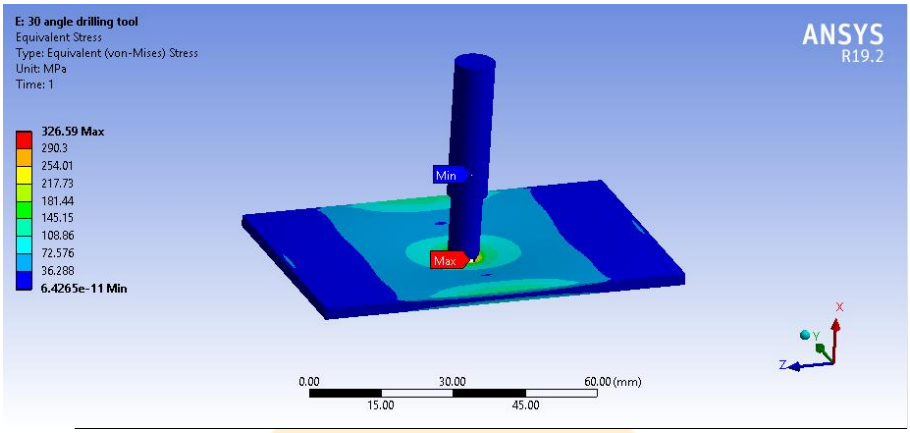
- These materials can be uses in Valves, and pump and their shafts and motor shafts
- washing machines parts sometimes
- It can also use while making Gears, nuts and bolts,
- This steel416 material is used in component of automatic screw machined

Results

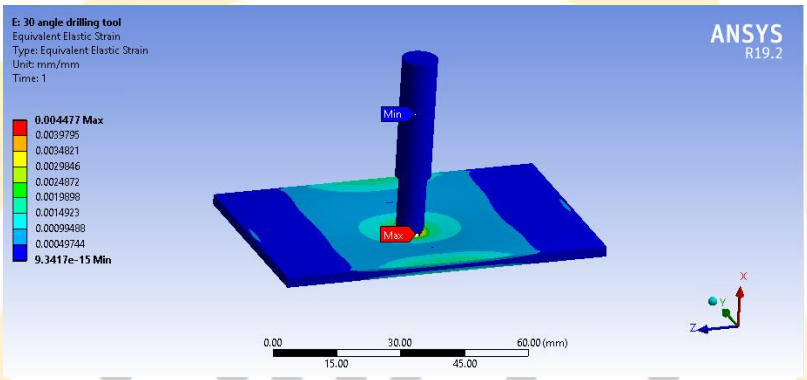
Deformation



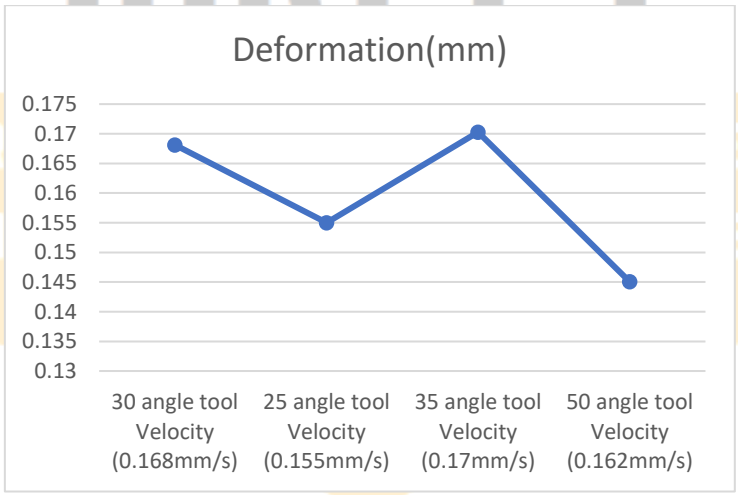
Stress

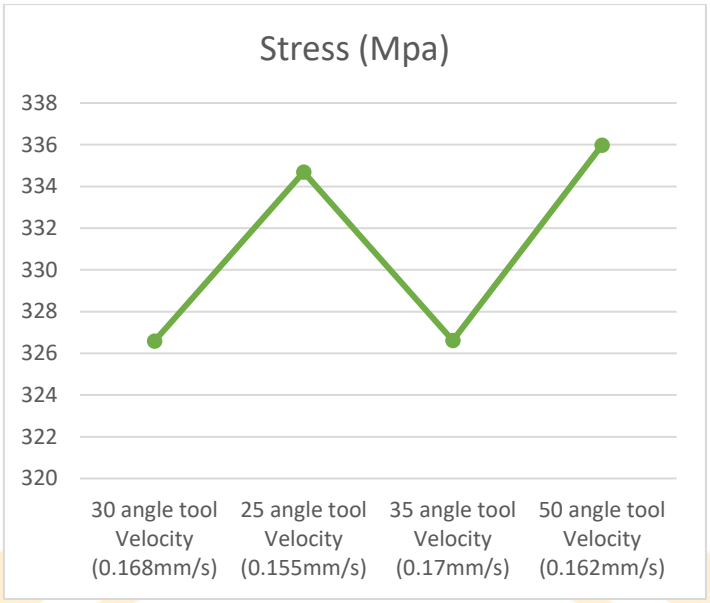


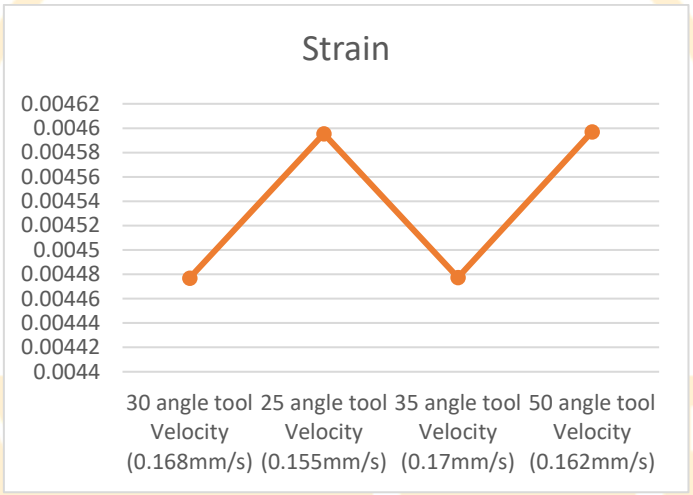
Strain

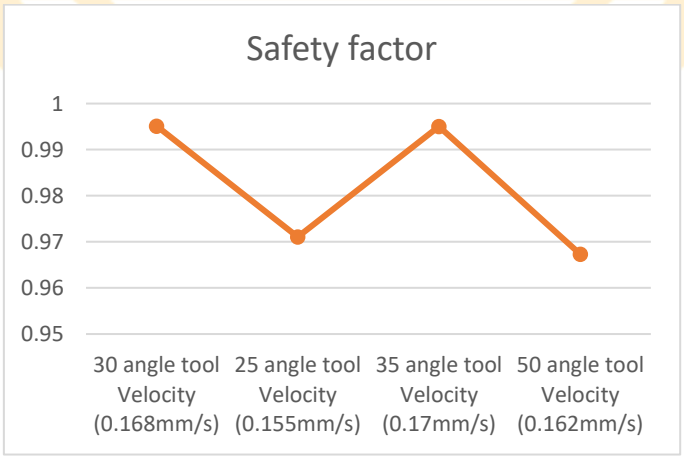


Safety factor









From the given table and graphs it clearly shown 25 and 50 degrees tools are best compare to other tools, but it is not the right way to choose an optimum tool for real time purpose, to get more clear knowledge

about each tool here need to know their thermal behavior when applied thermal boundary conditions on it, and those thermal boundary conditions results were shown in below image and discussed it clearly

CONCLUSION

Optimal friction drilling tool design and analysis are the primary objectives of this paper, with the goal of achieving better real-time performance. Friction drilling tool designs were created using solid works, and then analysed with the help of Ansys workbench. For the friction drilling tool, sae 1137 material was selected, and for the plate material, al-5059 material was selected, in this process.

Using SolidWorks, four different drilling tools (30 degrees, 25 degrees, 35 degrees, and 50 degrees) were created. While the 30* degree drilling tool is considered to be an existing tool, the remaining three tools were considered to be new or modified drilling tools in order to improve the performance of the object.

When compared to current and other planned models, a 25*degree drilling tool requires less axial velocity to complete the drilling operation, as determined by the findings of static analysis. Furthermore, it satisfies thermal boundary criteria; the heat flow value is lower when compared to other tools, allowing the object's performance to be improved in less time while experiencing no temperature losses.

. REFERENCES

[1] Miller S.F. Li R. Wang H. and Shih A.J. “Experimental and Numerical Analysis of the Friction Drilling Process” Journal of Manufacturing Science and Engineering Vol. 128, August 2006, pp. 802-810.

[2] Miller S.F. Blau P.J. and Shih A.J. “Tool Wear in Friction Drilling” International Journal of Machine Tools & Manufacture, Vol. 47, 2007, pp. 1636-1645.

[3] Miller S.F. and Shih A.J. “Thermo-Mechanical Finite Element Modeling of the Friction Drilling Process ” Journal of Manufacturing Science and Engineering Vol. 129 June 2007, pp. 531-538.

[4] Chow H.M. Lee S.M. and Yang L.D. “Machining Characteristic Study of Friction Drilling on AISI 304 Stainless Steel ” Journal of materials processing technology, Vol. 207, 2008, pp. 180-186.

[5] Qu J. and Blau P.J. “A New Model to Calculate Friction Coefficients and Shear Stresses in Thermal Drilling ” Journal of Manufacturing Science and Engineering Vol. 130, February 2008, pp. 014502/1-014502/4.

[6] Lee S.M., Chow H.M., Huang F.Y. and Yan B.H. “Friction Drilling of Austenitic Stainless Steel by Uncoated and PVD AlCrN- and TiAlN-Coated Tungsten Carbide Tools ” International Journal of Machine Tools & Manufacture Vol. 49 2009 pp. 81-88.

[7] Fernández A., Lopez de Lacalle L.N. and Lamikiz A. “Friction Drilling of Stainless Steels Pipes ” International Conference on Advances in Materials and Processing Technologies, 2010, pp. 1187-1192.

[8] Krishna P.V., Kishore K. and Satyanarayana V.V. “Some Investigations in Friction Drilling AA6351 Using High Speed Steel Tools ” ARPN Journal of Engineering and Applied Sciences, Vol. 5, No. 3, March 2010, pp. 11-15.

[9] Ku W.L. Hung C.L. Lee S.M. and Chow H.M. “Optimization in Thermal Friction Drilling for SUS 304 Stainless Steel ” Int J Adv Manuf Technol Vol. 53 2011 pp. 935-944.

[10] Krasauskas P. “Experimental and Statistical Investigation of Thermo-Mechanical Friction Drilling Process ” Mechanika Vol. 17(6) 2011 pp. 681-686.

Arabian Journal of Geosciences (2023) 16:204
https://doi.org/10.1007/s12517-023-11299-z

ORIGINAL PAPER

Performance and emission characteristics analysis of LPG-Karanja biodiesel on CI engine with optimization

Sheel Bhadra[1] · Hansham Dewal[1] · Naveen Kumar[1]

Received: 19 October 2022 / Accepted: 17 February 2023
© Saudi Society for Geosciences and Springer Nature Switzerland AG 2023

Abstract
Due to the increasing growth of industrialization and transportation, fuel consumption of fossil fuel as the main energy source is increasing exponentially which is undoubtedly crucial for world growth; however, exhaust emissions are hampering the environmental health to a large extent. Efforts have been undertaken in recent years to minimize harmful emissions from diesel engines by exploring alternative fuels. This study aims to investigate the performance and emission characteristics of a CI engine in dual fuel mode using Karanja oil biodiesel/diesel in various proportions along with liquefied petroleum gas (LPG) that is supplied as a secondary fuel. The effect of different fractions of biodiesel to LPG on emission and performance of an engine at different loads is studied. The performance characteristics like specific fuel consumption and brake thermal efficiency are analyzed. Various exhaust emissions like of carbon dioxides, carbon monoxides, and nitrogen oxides are calculated at different LPG fractions and blends. Response surface methodology (RSM)–based multi-objective optimization technique is used to optimize the engine performance and emission characteristics and is validated with an experimental data. Optimized values like brake-specific fuel consumption (BSFC), brake thermal efficiency (BTE), emission of nitrogen oxides (NOx), hydrocarbon (HC), and carbon monoxide (CO) were 23.2962%, 0.294243 g/kWh, 384.121 ppm, 22.0585 ppm, and 0.0295123%, respectively. Consequently, this study recommends to keep inputs, load (%), LPG (%R), and biodiesel (%Vol) near to 53.84, 18.99, and 20 respectively to get best and optimized results.

Keywords Biodiesel · Internal combustion engine · Optimization · Physiochemical properties · Alternate fuels

Introduction

In the second half of the twentieth century, the population growth rate was the highest. Due to the suddenly rise in the population and human living standards, energy demand increased exponentially which causes unexpected growth in the industries and transportation sector, as a result, fuel reserves around the world are emptying. Around 84 percent of this energy is derived through the utilization of fossil fuels. A major portion of the toxic and harmful gases in the environment are released by burning of fossil fuels. Out of the various forms of fossil fuels, diesel is one of the main contributors to producing harmful gaseous products. Therefore, it acts as a main emitter of pollutants in the environment (Chow 2001). These problems cause researchers; attraction towards renewable and alternative sources of energy which not only fulfill the escalating energy demands but also play a major role in substituting fossil-based derived fuels. In this study, biodiesel and LPG respectively act as alternative sources of energy.

Biodiesel is prepared by various feedstocks which are both edible and non-edible in nature, such as vegetable oils and animal fats. Due to its easily availability, low cost, desirable composition, and good own life, vegetable oils (edible and non-edible) have high potential to biodiesel (Salvi and Panwar 2012). Pyrolysis, micro-emulsification, and transesterification are three different methods from which transesterification commonly used for converting vegetable oils into biodiesel (Fukuda et al. 2001). Production of biodiesel from edible oil such as palm, coconut, soybean, corn, and rape seed oils or other food crops like sugarcane will lead to severe shortage in food, and its security and high price

Responsible Editor: Santanu Banerjee

✉ Sheel Bhadra
sheelbhadra_2k19me230@dtu.ac.in

[1] Mechanical Engineering Department, Centre for Advanced Studies and Research in Automotive Engineering, Delhi Technological University, New Delhi 110042, India

Published online: 25 February 2023

of edible oil make them not feasible for the production biodiesel (Islam et al. 2018). On the other end, non-edible oil like Jatropha, Karanja, and Mahua are the better alternatives for feed stock of biodiesel. The exhaust gas of Karanja biodiesel has less content of carbon monoxide and hydrocarbon, which focus more attention nowadays (Patel and Sankhavara 2017). Because of the high viscosity of biodiesel, the fuel injectors of the diesel engine are getting choked. Hence, to avoid this situation, blending of biodiesel with some other fuels is generally suggested. Twenty-five percent blend of Karanja biodiesel with neat diesel is most suitable for CI engine (Patel and Sankhavara 2017). Although blending of biodiesel can go up to 30 percent, but there is still 70% diesel consumed. More diesel can be replaced by dual fuel mode in compressed ignition engine. In dual fuel mode engine work as both SI and CI engine. SI engine, because air is mixed with Gaseous fuel as secondary fuel and supplied from Intake Port and CI engine, because conventional diesel or biodiesel blends as a primary fuel is supplied by injector for Ignition. Gaseous fuels produce virtually insignificant SOx and relatively little NOx. SOx and NOx are the main constituents of acid rain. Moreover, gaseous fuels emit substantially less CO_2 which is a key culprit in the greenhouse debate, than the most oil products and coal (Sahoo et al. 2009). Therefore, gaseous fuels can be considered more environment-friendly. There are different researchers who done their studies on gaseous fuels like hydrogen, CNG, natural gas, di-methyl ether, and biogas in dual fuel mode (Saravanan et al. 2008, Liu et al. 2013, Papagiannakis and Hountalas 2003, Arcoumanis et al. 2008, Bari 1996). Raman and Kumar (2022) conducted research and used acetylene as a gaseous fuel in CI engine and reported the significant reductions in emissions at 80% load as compared to only diesel. Blending of biodiesel with diesel or pilot fuels may increase the total percentage of alternative fuel to much more extent. Santhoshkumar et al. (2019) have blended the waste cooking oil biodiesel, blended with neat diesel, and investigated in dual fuel engine with LPG. This study found that smoke has reduced by 46.3% as compared to baseline. The dual fuel engine that utilizes LPG not only produces good power and efficiency but also reduces the emission like CO, CO_2, and NOx (Selim 2004, Luft 2001). There are many input factors on which engine's performance and emissions depend, and in case of dual fuel engine, this increases to much more. There is need to optimize these input factors to get best results. There are a few of studies available on optimization of performance and emission of CI engines with diesel or biodiesel. Yessian and Varthanan (2020), used Taguchi method on biodiesel, Atmanli et al. (2015) used RSM on ternary blend, and Bhatt and Shrivastava (2021) applied ANN on CI.

Although a number of studies have considered dual fuel engine and blended pilot fuel, several things still exist that have to be investigated, including (1) incorporating a higher percentage of alternative fuel in dual fuel engines and (2) optimization of input factors in dual fuel engine. This study focuses on maximize the use of alternative fuels more over improved the engine performance and reduced the emission simultaneously. LPG is supplied as a gaseous fuel. Response surface methodology (RSM) is used in this study for optimization of performance characteristics, brake specific fuel consumption (BSFC), brake thermal efficiency (BTE) and emission characteristics, hydrocarbon (HC), carbon monoxide (CO), and nitrogen oxides (NOx). In this model percentages of load, LPG and biodiesel are considered input factors.

Table 1 Equipments to measure the properties

Properties	Equipments
Density	DMA 4500, Anton Par
Kinematic viscosity at 40 °C	Visco bath, Petrotest
Calorific value	Oxygen bomb calorimeter, parr-6100
Cetane index	-
Flash point	Pensky-Martens (AD0093-600)
Sulfur content	Sandie-7039, sulfur analyzer
Fatty acid compositions	GCMS-QP2010 PLUS
Carbon residue	Alcor Micro Carbon Residue Tester

Table 2 Physiochemical properties of crude Karanja oil

Properties	Values
Density (Kg/m^3)	930
Kinematic viscosity (cSt) at 40 °C	35.5
Calorific value (MJ/Kg)	42
Flash point (°C)	237
FFA	0.9

Material and methods

Karanja oil purchased from certified seller Himani International, India. Oil was extracted from kernel by mechanical extractor which is very common method. The equipment used for measuring the properties of Karanja oil and its biodiesel (KBO) are given in Table 1. Physiochemical properties of crude oil measured and compared with the ASTM standards.

Properties of crude oil and its acid compositions are tested and given in the Tables 2 and 3 respectively. For finding the fatty acids and their percentages present in oil, the GC-MS (Shimadzu, Japan, Model/Serial Number GCMS-QP2010 PLUS) was used.

Table 3 Composition of fatty acids

Fatty acids	Percentages	Structure
Oleic acid	52.14	C (18:1)
Linoleic acid	15.98	C (18:2)
Palmitic acid	12.11	C (16:0)
Stearic acid	8.00	C (18:0)

Table 4 Physiochemical properties of LPG (Atmanli et al. 2015)

Properties	Values
Lower heating value (KJ/kg)	46,000
Kinematic viscosity at 40 °C	-
Lower heating value (KJ/kg)	46,000
Cetane index	< 3
Flash point (°C)	− 102
Density (Kg/m^3)	562

In this study, LPG was considered a secondary fuel which consists of 70% propane and 30% butane. Properties of the LPG are shown in the Table 4.

LPG percentage defines in Eq. (1):

$$R\% = \frac{m_{LPG}}{m_{pilot\ fuel} + m_{LPG}} \times 100\% \quad (1)$$

LPG was supplied from the cylinder through the gas valve which controls the discharge and kept inside the intake manifold by a simple mixer.

Preparation of biodiesel

Karanja oil is heated in order to remove the moisture content which led to formation of soap during transesterification due to high free fatty acids (FFA) (Zulqarnain et al. 2021). Test the FFA of the oil whether it is less than or greater than two; in the former case, directly go for transesterification and in later case need to neutralize by esterification and then go for transesterification. Although there are several methods available for producing biodiesel, transesterification of natural oils and fats is currently the method of choice (Ma and Hanna 1999). Transesterification is the three-step process in which triglycerides react with alcohol in the presence of catalyst and converted it to esters and glycerol. As shown in Fig. 1, esters have very low viscosity and molecular weight than Karanja oil which make suitable to use as diesel (Peterson 1986). After transesterification glycerol separated by gravity separation Method. For separation, it is kept in a separating funnel for around 24 h, and glycerol is settled at the end and removed. Biodiesel is further washed with hot water for removing the mixed glycerol. Due to water washing some water left in biodiesel, this moisture has removed by heating. This available biodiesel has high viscosity, so it needs to blend with diesel. Blends of 20 (%Vol) KBO and 10 (%Vol) were kept in a test tube to check the stability (Fig. 2).

Fig. 1 Transesterification

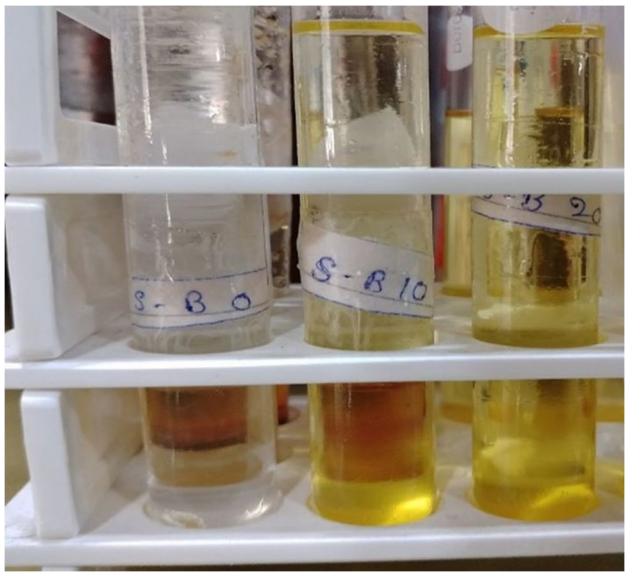


Fig. 2 Stability test

The author also tested important physiochemical properties of all these KBO blends, and the results is shown in Table 5.

Experimental setup

For testing the performance and emissions characteristics, the author has used a 4-stroke single cylinder direct ignition

Table 5 Physiochemical properties of KBO

Properties	100 % biodiesel	10% biodiesel	20% biodiesel	Diesel	Standard
Density (kg/m^3) at 15	880	838	844	832	ASTM D4052
Kinematic viscosity (cSt) at 40 °C	4.5	2.2	2.42	1.9	ASTM D445
Flash point	185	72	79	64	ASTM D93
Calorific value (MJ/kg)	35	42	41.01	42.22	ASTM D240

Table 6 Specification of engine

Parameters	Specifications
Model	KSA-2057-P-1170 (Kirloskar)
No. of cylinder	1
No. of stroke	4
Injection type	Direct injection engine
Cooling system	Air-cooled
Lubrication system	Forced feed
Cubic capacity	0.78 L
Bore diameter	95 mm
Orifice diameter	13.4 mm
Rated power	3.5 kW
Rated speed	1500 rpm
Compression ratio	17.5:1
Fuel injection timing	23°bTDC (before top dead center)
Fuel injection pressure	200 bar

engine. This engine has rated power of 3.5 kW with a constant 1500 rpm. Eddy current dynamometer is used to operate the load on the engine which is also measured the torque and power supply. A high precision strain gauge was fixed for proper load transmission load. Engine load has a range of 0–25 kg with the accuracy and uncertainty of ± 0.1 kg and ± 0.1 % respectively. At the arm of the dynamometer, a magnetic pick-up type of sensor is attached to determine the RPM. LPG discharge and pressure were controlled through the gas flow control valve and pressure regulator separately. The mass flow rate of the LPG was measured by the reduced weight of the cylinder and the corresponding time trough stopwatch. Similarly, for the calculation of pilot fuel's mass flow rate, volume and density were considered in place of mass. To record the exhaust gas temperature, a thermocouple was mounted on the exhaust pipe. A gas analyzer and smoke meter were mounted on exhaust pipe to record the amount of emissions present in the exhaust gas. Specification and details of the setup are described in Table 6 and Fig. 3.

RSM optimization

In 1951, a multi-objective optimization technique, i.e., RSM, was developed by Box and Wilson. This method combines mathematical and statistical methods to model and analyzes various experimental problems. It is one of the essential optimization techniques that can be used to minimize both time and money invested in conducting different academic and practical trials. This method aims to find the correlation between output parameters (y1, y2, y3, etc.) which are affected by input parameters (x1, x2, x3, etc.) and optimize the output parameters accordingly. In this study, with the help of RSM, optimum dosages of the biodiesel, substitution degree of LPG, and other operating conditions can be found out at which test engine delivers the maximum performance with negligible emissions.

Figure 4 depicts the detailed procedure to apply the RSM technique. The preliminary stage involves the selection of both input and output parameters. In this study, the input parameters are substitution degree of LPG (%R), vol % of biodiesel (KBO), and load (%), which were considered input variables, and the parameters which are affected by these input parameters are break thermal efficiency (BTE), break-specific fuel consumptions (BSFC), nitrogen oxide (NOx), carbon monoxide (CO), and unburnt hydrocarbons (HC) which were considered output responses. Afterwards, the range of input variables was identified. On the basis of comprehensive literature survey, a wide range (i.e., difference of upper limit and lower limit of input parameters) were selected, and on these values, some preliminary experiments were carried out. The values of upper limit, lower limit, and middle limit for each input variables are shown in Table 7.

In the third step, a suitable model is selected for designing the experiments. A central composite face centered (CCFD) model is used, which helps to maintain each prediction result's accuracy. In the CCFD model, eight cube points, six points in a cube, and six axial points were present. Furthermore, the presence of a star point on the face of the cubic part in the design is also why this form is called face-centered. For making the RSM model, MINITAB software version 19.1 was used. In Table 7, the specifics of experimental trails carried out with the help of the CCFD model are shown in the form of a design matrix. This design matrix consists of 20 experimental runs

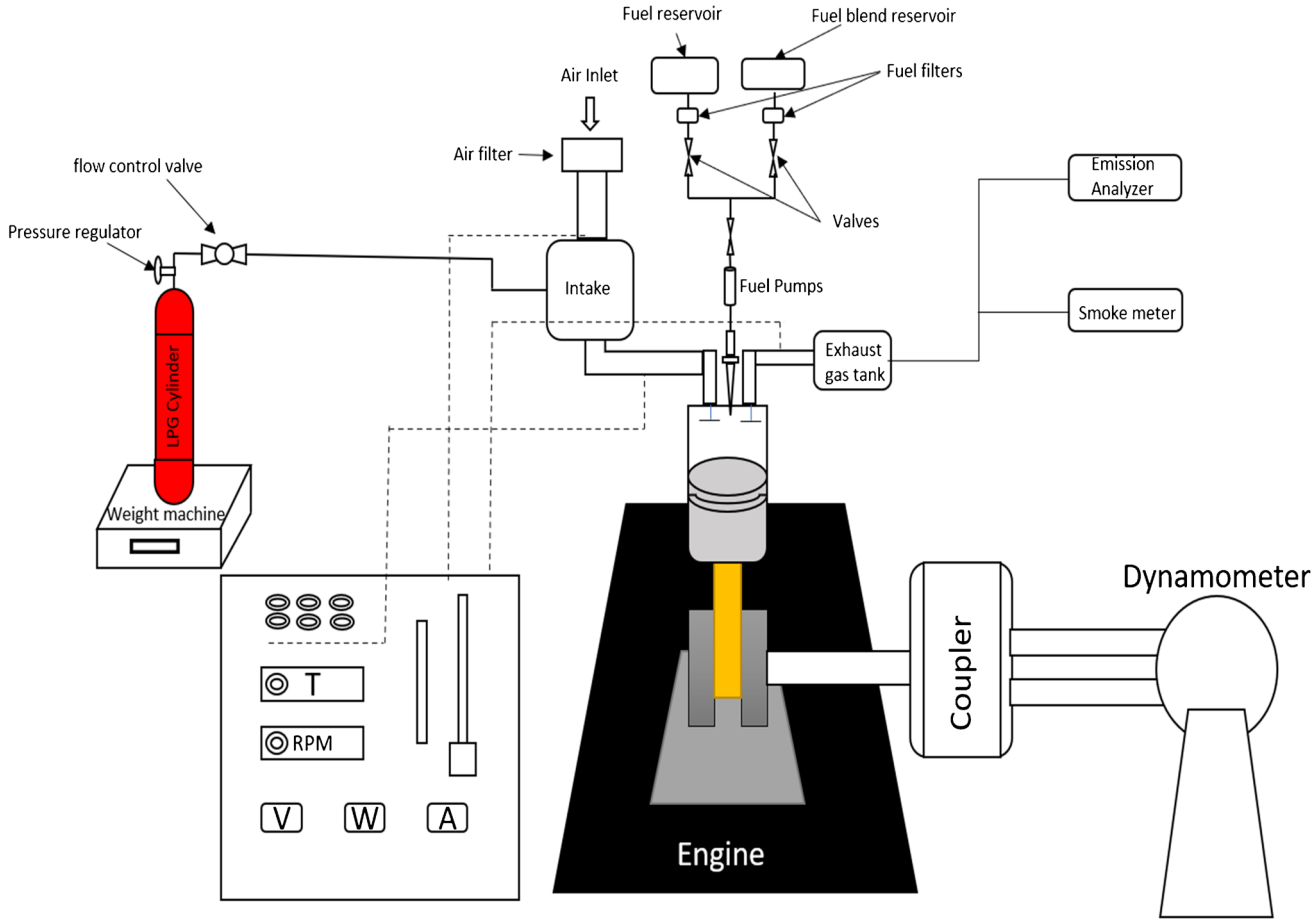


Fig. 3 Experimental setup

for various combinations of input variables. The input factors in the CCFD matrix represented by codes are given in Table 8.

A second order regression model is used to indicate the correlation between input variables and output response and is given below in Eq. (2):

$$Y = \beta_0 + \sum_{i=1}^{k} \beta_i X_i + \sum_{i=1}^{k} \sum_{j\geq 1}^{k} \beta_{ij} X_i X_j + \propto \quad (2)$$

where,
Y = predicted response
β = regression coefficient
k = number of factors
i = linear coefficient
j = quadratic coefficient
α = random error in the response

Result and discussion

Performance parameters

Brake thermal efficiency

Brake thermal efficiency is a vital performance characteristic of an IC engine (Thangavel et al., 2022). Brake thermal efficiency is an evaluation of the engine's potential to convert heat which is generated from fuel into mechanical energy. The BTE of an internal combustion engine is influenced by a number of factors, including the type and design of the engine, the fuel used, and the operating conditions. The surface plots and contour plots in Fig. 5 demonstrates the influence of engine input factors (LPG (g/min), engine

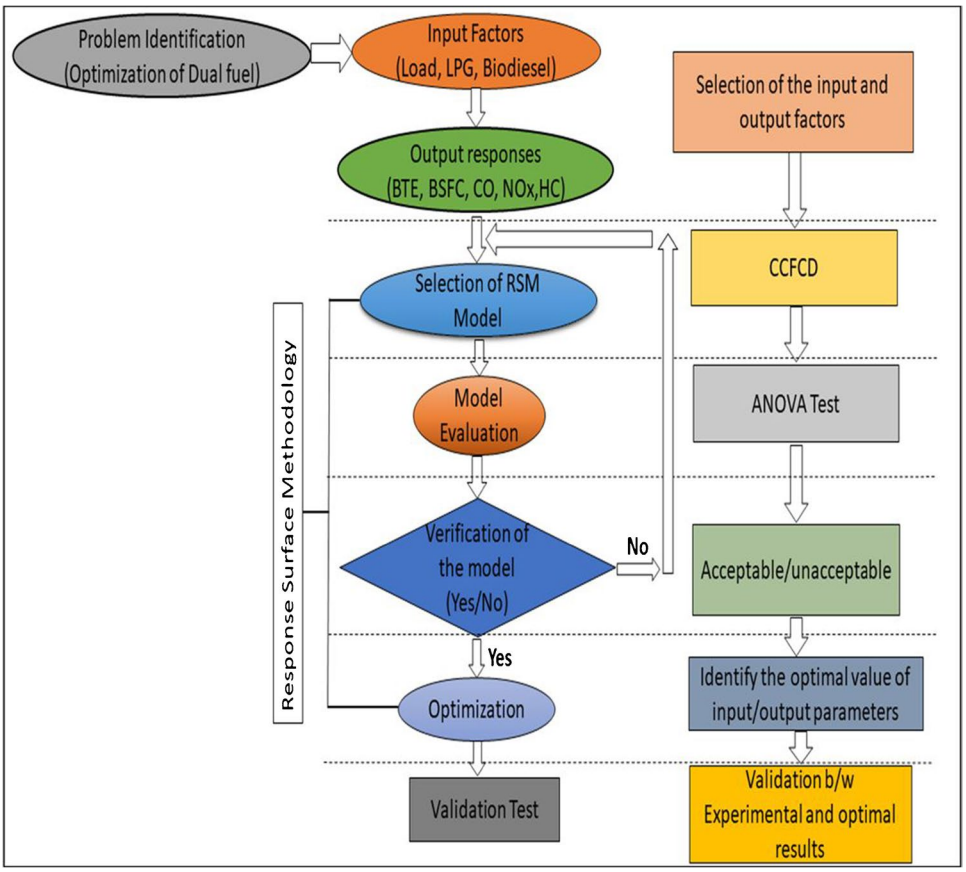


Fig. 4 Response surface methodology procedure (flow chart)

Table 7 CCFD matrix for coded three input factors with experimental output

Run	A	B	C	BTE %	BSFC	NOx (ppm)	CO %	HC (ppm)
1	0	0	1	20	0.35	290	0.03	27
2	0	0	0	18	0.34	230	0.04	30
3	0	− 1	0	21	0.31	450	0.06	20
4	− 1	− 1	1	15	0.62	450	0.05	5
5	1	1	− 1	26	0.32	500	0.13	70
6	1	1	1	25	0.323	540	0.065	65
7	0	0	− 1	22	0.32	250	0.085	32
8	− 1	− 1	− 1	16	0.59	350	0.07	20
9	0	0	0	18	0.34	230	0.04	30
10	0	0	0	18	0.34	230	0.04	30
11	− 1	1	1	12	0.64	340	0.075	50
12	1	0	0	24	0.34	450	0.04	38
13	0	0	0	18	0.34	230	0.04	30
14	0	0	0	18	0.34	230	0.04	30
15	0	1	0	16	0.36	300	0.08	50
16	− 1	0	0	12	0.62	150	0.04	23
17	1	− 1	1	29	0.283	650	0.04	20
18	0	0	0	18	0.34	230	0.04	30
19	− 1	1	− 1	13	0.61	300	0.09	55
20	1	− 1	− 1	30	0.28	550	0.11	35

Table 8 Codes given to input factors in CCFD matrix

	Codes		
	− 1	0	1
Input factors	20	45	70
LPG	0	30	60
Biodiesel	0	10	20

load (%), and Karanja biodiesel (%vol)) on the brake thermal efficiency of engines. This is noticed that on fixing the amount of KBO at 10 (%vol) in the fuel blend, and with increasing the substitution degree of LPG, the BTE values of the engine tends to decrease. This can be attributed due to the less availability of the oxygen in the engine cylinder which tends to decrease the quality of the combustion. On the other side, by keeping the substitution degree of LPG fixed at 30 (R%), and increasing the dosages of KBO up to 10 (%vol) in the fuel blends, it tends to decrease the BTE of the engine due to the less calorific value of the KBO, while on further increasing the amount of KBO in the test fuel blends, BTE tends to increase. This is due to the additional lubricity of the biodiesel which helps to overcome the frictional losses and, hence, the brake thermal efficiency of the engine is increases.

However, for all concentration of KBO (%vol) in the test fuel blend and substitution degree of LPG (R%), on increasing the value of engine load, the BTE of the engine tends to increase. This is due to the fact that as the engine load is increases, the brake power of the engine is increases, which helps to increase the overall efficiency of the engine.

Brake-specific fuel consumption

This parameter reflects the engine's efficiency which produces rotational power at the expense of burning of fuel. The value of BSFC also represents the ability of the engine which transforms the fuel energy into productive power**.** BSFC of a dual fuel engine will vary depending on the operating conditions of the engine and the relative proportions of each fuel type being used. BSFC was measured by obtaining the ratio of total mass of fuel (biodiesel+diesel+LPG) to power output in specific time period. Figure 6 portrays the effect of input factors such as LPG (g/min), engine load (%), and Karanja biodiesel (%vol) on BSFC. It is observed that, on fixing the concentration of KBO at 10 (%vol) in the fuel blend, and with rising the substitution degree of LPG (R%), BSFC increases. This can be attributed due to the unavailability of the oxygen which tends to combust the gaseous fuels partially. On the other hand, while fixing the substitution degree of LPG at 30 (R%), the BSFC tends to increase with increasing the amount of KBO (%vol) in the test fuel blends. This is due to the lower calorific value of the KBO; as a result, a greater amount of fuel is required to produce the same amount power as compared to conventional fuel operation. Moreover, at lower loads, the BSFC is higher which indicates that engine takes more amount of indicated power to overcome the friction. However, at higher loads, the values of BSFC are lower; it is due to the presence of higher cylinder temperature which helps in improving the atomization process; as a result, proper combustion of the test fuel blends has been occured inside the cylinder.

Emission parameters

Carbon monoxide emissions

The formation of hazardous CO emissions depends on three factors: lesser residence time, inadequate supply of oxygen, and lower reaction temperature. Figure 7 implies the effects of listed variable inputs on the formation of CO emissions. On holding the dosage of KBO at 10 (%vol) and with increasing the substitution degree of LPG up to around 25 (R%), the CO emissions tends to decrease, while with further increasing the substitution degree of LPG, CO emissions showed an increasing trend; in a similar trend for the effect of LPG reported by Santhoshkumar et al. (2019), this research studied the LPG with waste cooking oil. This is due to the lower C/H values of LPG which helps to decrease the CO formation. However, the reason behind for increasing the CO emission is due to the non-availability of required oxygen molecules; this results in partial oxidation of CO to CO_2 molecules.

On fixing the substitution degree of LPG at 30 (R%) and increasing the KBO concentration in the fuel blends, CO emissions tend to decrease. It is due to the presence of extra oxygen molecules in biodiesel which helps in enhancing the combustion rate, which tends to improve the oxidation rate; as a result, CO emissions are reduced slightly. Moreover, on increasing the load, the CO emissions tend to increase. This is attributed to the presence of rich mixture and less availability of free oxygen molecules which tends to promote the incomplete combustion of the test fuel blends and hence CO is increases.

Unburnt hydrocarbon emissions

The formation of hydrocarbon emissions occurs due to partially burning or completely unburning of the test fuel blends. Figure 8 depicts the impact of input parameters on the formation of HC emissions. On holding the value of substitution degree of LPG at 30 (R%) and increasing the KBO concentration in the test fuel blends, HC emissions tend to decrease. This is attributed to the lower calorific value of the KBO, which does not facilitate the combustion of the fuel blends; as a result of the lousy combustion process, the formation of HC emissions occurred.

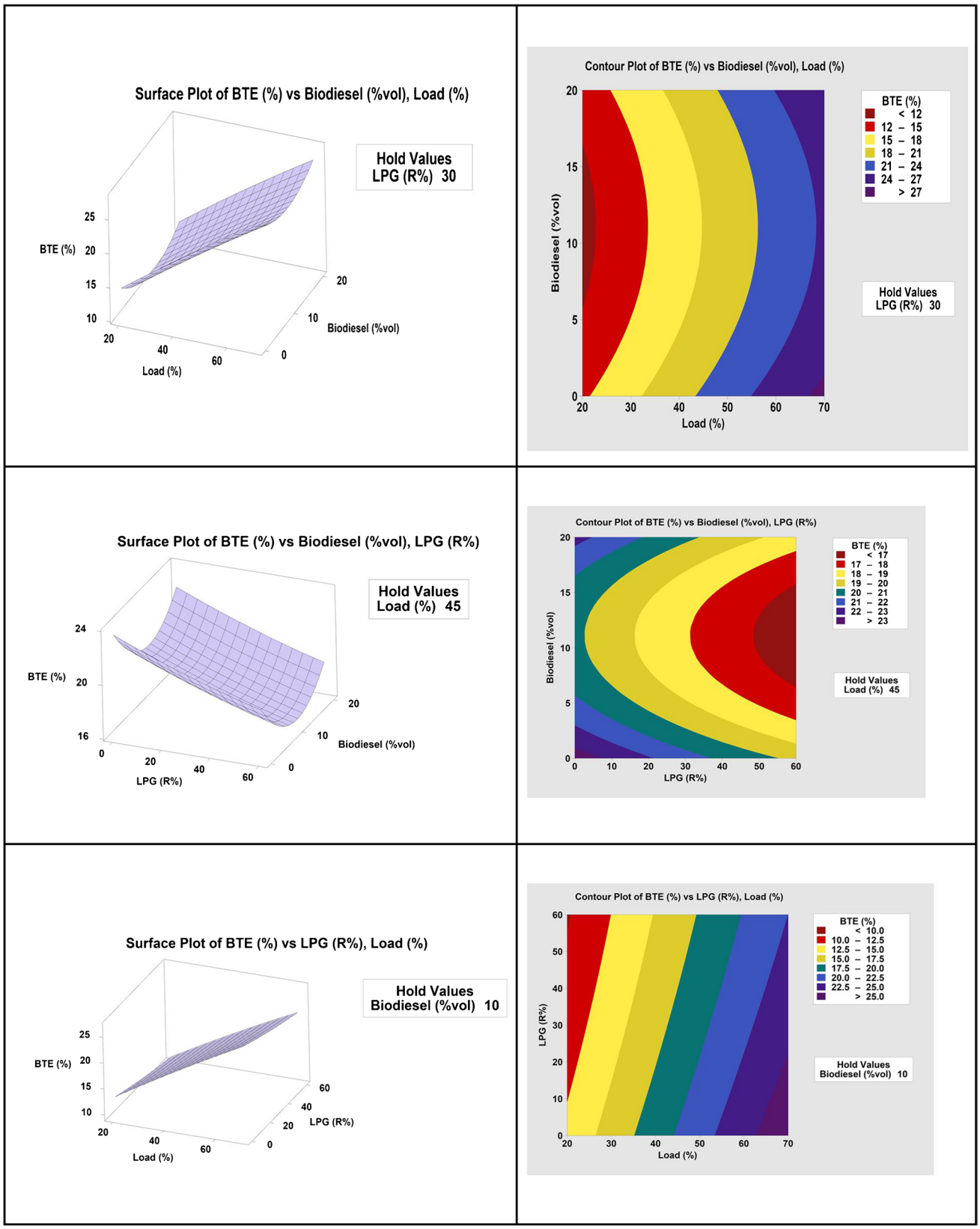


Fig. 5 Aggregate impact of LPG (g/min), load (%), and biodiesel (%vol) on brake thermal efficiency (BTE)

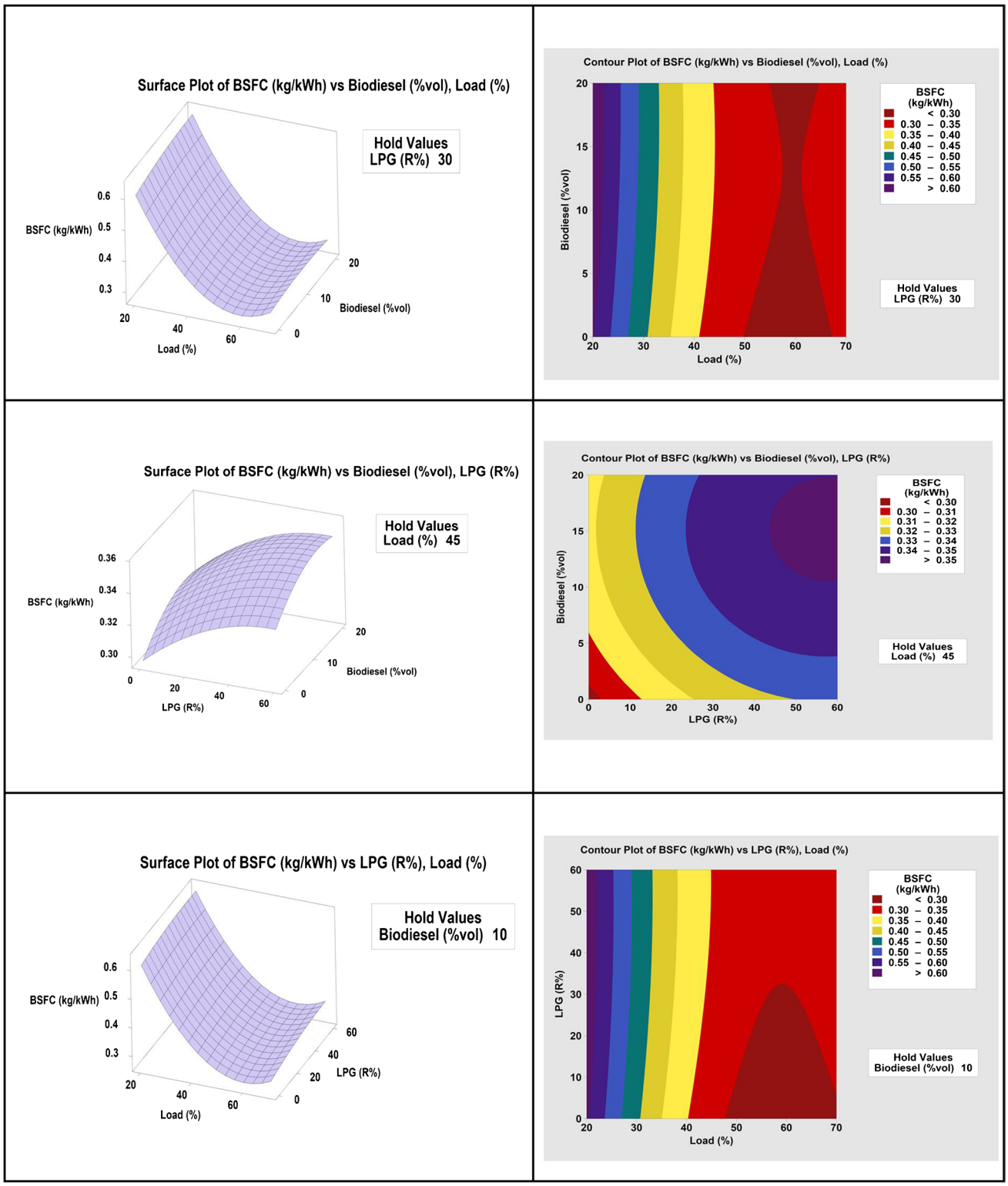


Fig. 6 Aggregate impact of LPG (g/min), load (%), and biodiesel (%vol) on brake-specific fuel consumption (BSFC)

On the other hand, while keeping the value of KBO fixed at 10 (%vol) in the fuel blends, and with increasing the substitution degree of LPG, HC emissions tend to grow up. Guan et al. (2017) studied the LPG-waste cooking oil biodiesel in the compressed ignition engine and reported the similar trend for LPG at different

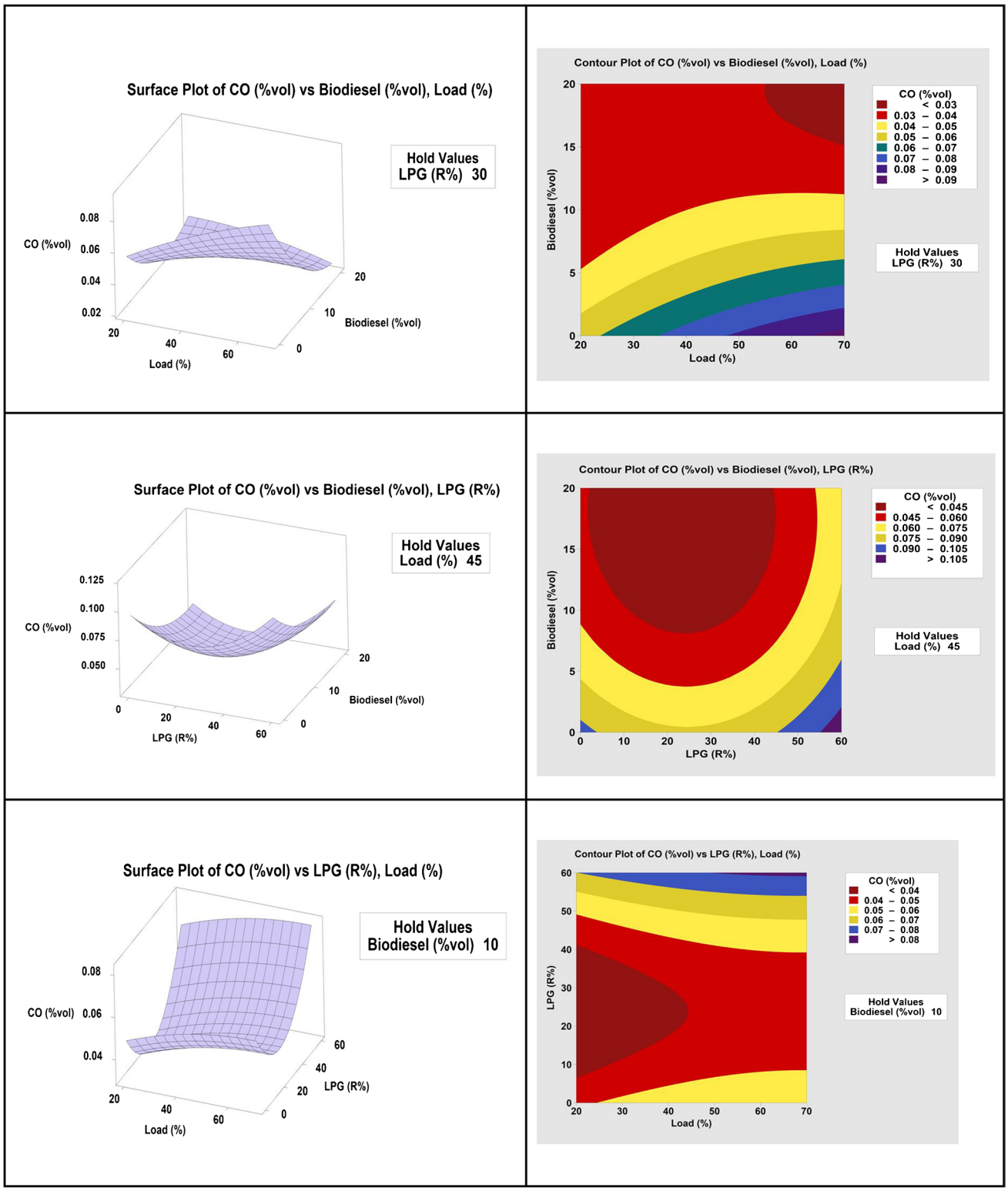


Fig. 7 Aggregate impact of LPG (g/min), load (%), and biodiesel (%vol) on carbon monoxide emissions (CO)

pilot fuels and various load conditions. As the substitution degree of LPG rises in the engine cylinder, the amount of oxygen starts decreasing. Hence, due to the non-availability of oxygen, the combustion of the fuel blends does not occur properly, and therefore HC emissions will be formed.

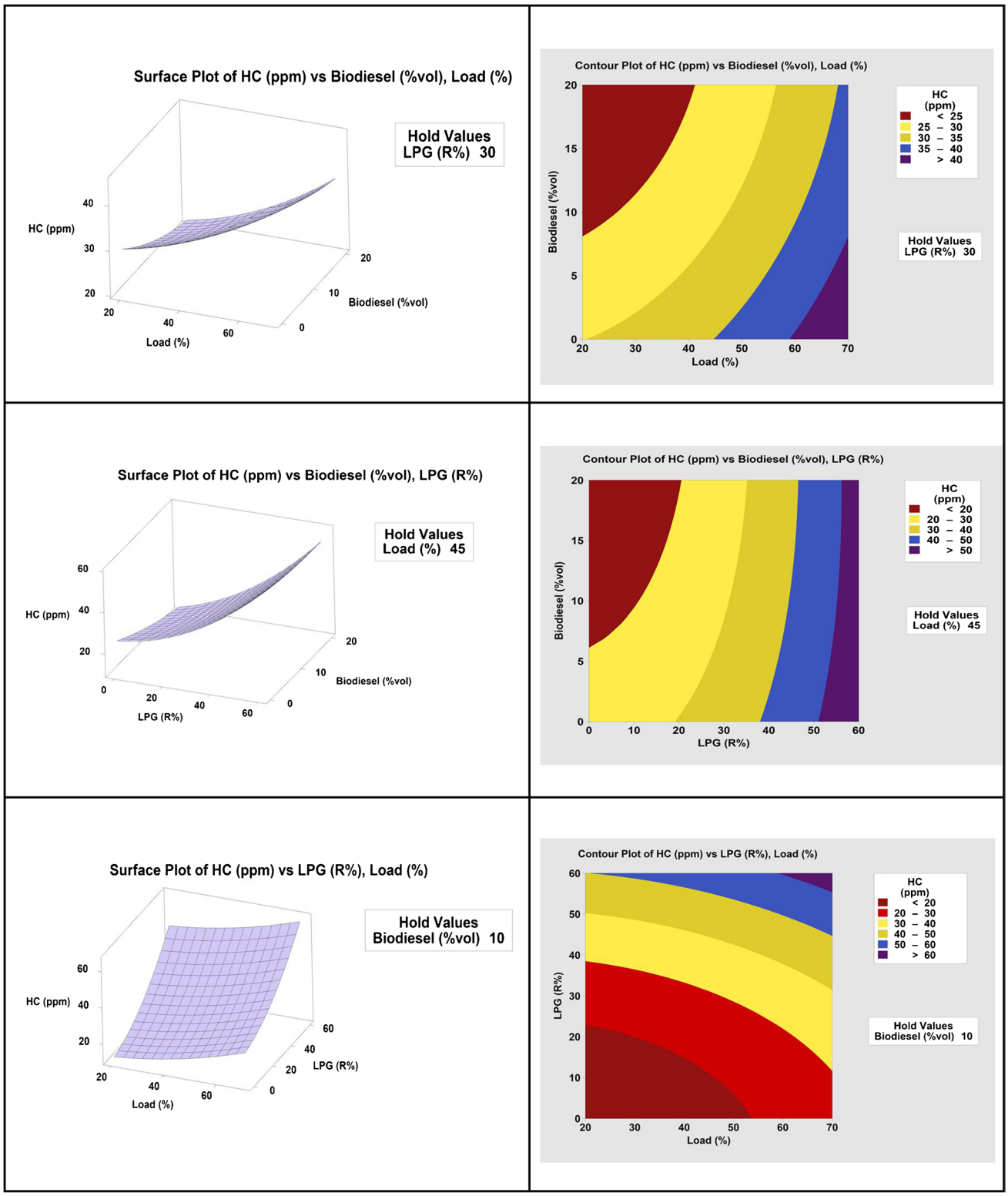


Fig. 8 Aggregate impact of LPG (g/min), load (%), and biodiesel (%vol) on hydrocarbon emissions (HC)

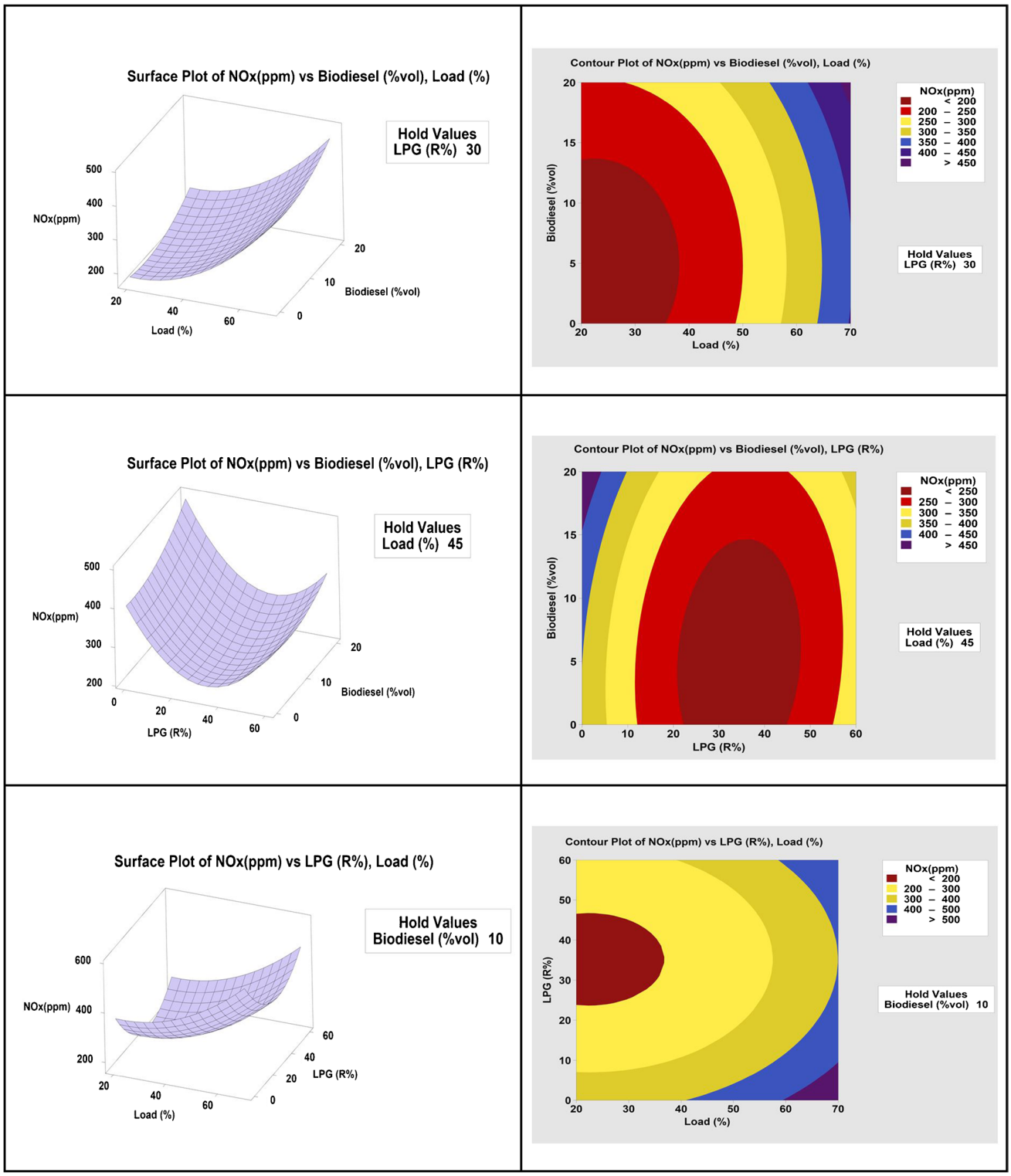


Fig. 9 Aggregate impact of LPG (g/min), load (%), and biodiesel (%vol) on oxides of nitrogen emissions (NOx)

Moreover, on holding the values of both LPG and KBO, and with increasing the engine load, HC emissions tend to rise up. It is attributed to the presence of a rich mixture in the engine cylinder and less availability of oxygen; as a result, incomplete combustion has been taken place, which leads to the formation of HC emissions.

NOx emissions

NOx emissions are one of the most harmful exhaust emissions from conventional internal combustion engine. The production of NOx emissions generally depends upon the engine cylinder temperature. Figure 9 implies the cumulative effect of input factors LPG (R%), engine load (%), and Karanja biodiesel (%vol) on the formation of NOx emissions. The surface plots and contour plots display that keeping the concentration of KBO fixed at 10 (%vol) and on increasing the substitution degree of LPG up to 45 (R%) resulted a significant reduction of the NOx emissions, and with further increasing the substitution degree beyond 45 (R%), NOx emissions are higher. This is due to the non-availability of the free oxygen which can react with nitrogen to form NOx emissions effectively. Anye Ngang and Ngayihi Abbe (2018) investigated experimentally and numerically the effect LPG on CI engine in dual fuel mode and found the same trend for NOx up to 50% of LPG substitution. However, the reason behind the rising nature of NOx emissions is higher self-ignition temperature of LPG. However, the reason behind the rising nature of NOx emissions is higher self-ignition temperature of LPG. Due to this, the ignition delay period of the charge is higher; as a result, the engine cylinder temperature increases when the entire charge in the engine is ignited.

On the other side, keeping the load constant at 45%, on boosting the dosage of KBO in the test fuel blends, NOx emissions show an increasing trend. It is due to the presence of extra oxygen concentration in KBO, which facilitate the formation of NOx emissions. Moreover, on increasing the engine load and remaining the other variable inputs fixed, it showed a higher NOx emission. This can be responsible for the spike in the engine cylinder temperature which helps to the formation of NOx emissions.

Validation results

The analysis of variance (ANOVA) is one of the best approaches used to define the significance of each response variable in a specified range for the various input parameters. Tables 9 and 10 represent the analysis of variance for different engine performance and emissions parameters. Table 10 shows that p values of all the input variables are below the 0.05 for the confidence interval of 95%. This emphasizes the significance of all input variables in the analysis of the BTE of the engine. Similarly, from Table 10, it is noticed that the p values of all variable inputs are also below 0.05, which states that all the input variables have some impact on the evaluation of BSFC. Smaller p values influence the model up to a large extent. In Table 10, the p values of all input variables are less than 0.05, demonstrating their significance in predicting CO, HC, and NOx. However, the squared p values of LPG profoundly affect all the engine exhaust emissions. The R^2 values of BTE, BSFC, CO, HC, and NOx are 99.53%, 99.77%, 98.68%, 98.76, and 98.19, respectively, as shown in Table 9, which signifies that the selected models are highly accurate with the experimental results. R^2 values for BTE depict that the input variables explain 99.53% of the variations in the BTE values effectively.

With the use of second-order response surface models, various regression equations for different engine parameters such as BTE, BSFC, CO, HC, and NOx have been developed. These equations consist of interactions terms, quadratic terms, and linear terms of all the input variables, i.e., load (%), LPG (R%), and KBO (%vol). These regression equations are valid only when the input variables' values are coming down under a certain range, i.e. 20–70% for engine load, 0–60%R for LPG, and 0–20%vol for biodiesel (KBO). These equations are as follows:

$$\begin{aligned} BTE\% =& 10.350 + 0.3133\ load - 0.0635\ LPG \\ & - 0.6055\ Biodiesel - 0.000436\ load * load \\ & + 0.000253\ LPG * LPG + 0.02727\ Biodiesel \\ & * Biodiesel - 0.000333\ load * LPG - 0.000000\ load * Biodiesel \\ & + 0.000000\ LPG * Biodiesel \end{aligned} \tag{3}$$

$$\begin{aligned} BSFC =& 1.0093 - 0.025611\ load + 0.000885\ LPG + 0.00403\ Biodiesel \\ & + 0.000217\ load * load - 0.000010\ LPG * LPG \\ & - 0.000093\ Biodiesel * Biodiesel + 0.000007\ load \\ & * LPG - 0.000027\ load * Biodiesel - 0.000000\ LPG * Biodiesel \end{aligned} \tag{4}$$

$$\begin{aligned} CO\% =& 0.05155 + 0.001198\ load - 0.001463\ LPG - 0.002989\ Biodiesel \\ & - 0.000005\ load * load + 0.000030\ LPG \\ & * LPG + 0.000143\ Biodiesel * Biodiesel \\ & + 0.000000\ load * LPG - 0.000050\ load \\ & * Biodiesel + 0.000004\ LPG * Biodiesel \end{aligned} \tag{5}$$

Table 9 ANOVA table

	BTE model	BSFC Model	NOx model	CO model	HC model
R-square	99.53	99.77	98.19	98.68	98.76
R-square (Adj.)	99.1	99.53	96.55	97.49	97.65
R-square (Pred.)	96.48	98.26	86.19	90.59	91.19

Table 10 Analysis of variance (ANOVA), *F* ratio test, and *p* value test all models

Regression	BTE model				BSFC (Kg/KW.hr) model				NOx (ppm) model				CO % model				HC (ppm) model			
	SS	*F* value	*p* Value	PC (%)	SS	*F* value	*p* Value	PC %	SS	*F* value	*p* Value	PC (%)	SS	*F* value	*p* Value	PC %	SS	*F* value	*p* Value	PC (%)
Model	512.505	232.86	0.000	99.525	0.318074	476.27	0.000	99.7672	356982	60.16	0.000	98.1866	0.014085	83.07	0.000	98.6759	4860.09	88.66	0.000	98.7622
Load (%)	435.6	1781.26	0.000	84.59	0.235316	3171.18	0.000	73.8093	121000	183.54	0.000	33.2806	0.000360	19.11	0.001	2.5220	562.5	92.35	0.000	11.4306
LPG (F %)	36.1	147.62	0.000	7.01	0.00289	38.95	0.000	0.906478	22090	33.51	0.000	6.07577	0.001210	64.22	0.000	8.4769	3610	592.69	0.000	73.359
Biodiesel (Vol %)	3.6	14.72	0.003	0.699	0.000922	12.42	0.006	0.28919	10240	15.53	0.003	2.8164	0.005062	268.7	0.001	35.4630	202.5	33.25	0.000	4.115
Load %*Load%	0.205	0.84	0.382	0.039	0.050660	682.71	0.000	15.89	10202	15.48	0.003	2.8060	0.000028	1.48	0.252	0.1961	14.78	2.43	0.150	0.30034
LPG (F %)*LPG(F %)	0.142	0.58	0.464	0.027	0.00236	3.19	0.105	0.74023	50796	77.05	0.000	13.9712	0.001978	104.98	0.000	13.8573	127.84	20.99	0.001	2.5978
Biodiesel (Vol %)*Biodiesel (Vol %)	20.455	83.64	0.000	3.972	0.00236	3.19	0.105	0.74023	2627	3.99	0.074	0.72254	0.000564	29.92	0.000	3.9512	4.78	0.78	0.397	0.09713
Load(%)*LPG (F%)	0.5	2.04	0.183	0.097	0.0002	2.7	0.132	0.062732	0	0.00	1.000	0	0.000000	0.00	1.000	0.00	0.00	0.00	1.000	0.00
Load (%)*Biodiesel (Vol %)	0.000	0.00	1.00	0.00	0.000365	4.91	0.051	0.1144	0	0.00	1.000	0	0.001250	66.34	0.000	8.7571	0.00	0.00	1.000	0.00
LPG(F%)*Biodiesel (Vol %)	0.000	0.00	1.00	0.00	0.0000	0.00	1.000	0.00	1800	2.73	0.129	0.4950	0.000013	0.66	0.434	0.091	50.00	8.21	0.017	1.01605
Total	514.95				0.318816				363575				0.014274				4921			

$$HC\ (ppm) = 19.42 - 0.034\ load + 0.095\ LPG - 0.964\ Biodiesel + 0.00371\ load * load + 0.00758\ LPG * LPG + 0.0132\ Biodiesel * Biodiesel + 0.00000\ load * LPG - 0.00000\ load * Biodiesel + 0.00833\ LPG * Biodiesel \quad (6)$$

$$NOx\ (ppm) = 399.8 - 4.37\ load - 10.13\ LPG - 1.48\ Biodiesel + 0.0975\ load * load + 0.1510\ LPG * LPG + 0.309\ Biodiesel * Biodiesel - 0.0000\ load * LPG - 0.0000\ load * Biodiesel - 0.0500\ LPG * Biodiesel \quad (7)$$

Equations (3) to (7) were derived from the second order regression model, explained by Eq. (2). These quadratic equations are important for understanding how changes in one input variable affect the response variable. These equations were also used to optimize the process by finding the input variable levels that result in the maximum or minimum response variable output. Additionally, quadratic equations helped to identify the presence of nonlinear relationships between the input and response variables, which assisted in the development of more accurate and robust models.

Optimization results

In this study, an RSM optimizer is used to optimize engine performance and emission parameters. In this optimization process, it is required that BTE achieve its highest value and the values of other emissions parameters such as BSFC, CO, HC, and NOx should be at their lowest level. Figure 4 depicts the details of the optimization principles used in this work. Moreover, Fig. 10 illustrates the findings of the RSM optimizer. The best values of the engine output parameters are 23.29 % for BTE, 0.294 (kg/kWh) for BSFC, 0.029 (%vol) for CO, 22.058 (ppm) for HC, and 384.12 (ppm) for NOx, respectively, and the corresponding optimum functioning parameters of the engine are 20 (%vol) for KBO, 18.987 (R%) for LPG, and 53.838 (%) for the engine load. Response surface methodology (RSM) is able to find the best engine operating parameters between the tested values for input factors. This multi-objective technique decreases the number of experimental trials significantly, saving both time and money.

There is a requirement for validating the results of the RSM with the experimental trials. Hence, in order to achieve this, three consecutive engine trials were conducted on the same combination of input variables which are already obtained from the RSM optimizer tool. The mean values of the three trials were carried out, and compare these values with the predicted values of the RSM. The details of the experimental values, predicted values, and residuals are given in Table 11. The values of the residuals are 5.264 (%) for BTE, 0.243 (%) for BSFC, 1.313 (%) for CO, 5.691 (%) for HC, and 1.111 (%) for NOx respectively, which are within an acceptable limit.

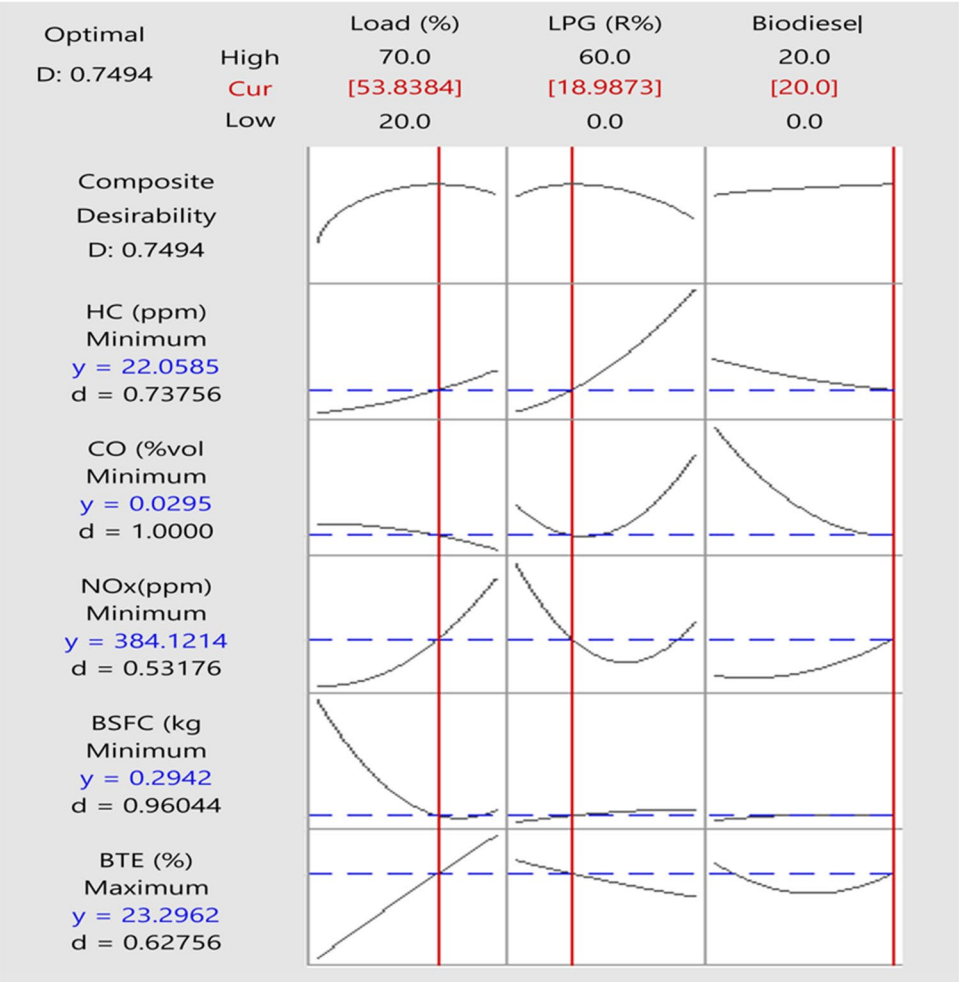


Fig. 10 Response surface methodology optimization curves

Table 11 Validation table

Output factors	Predicted values	Actual values	Errors in percentages
BTE (%)	23.2962	22.07	5.2649
BSFC	0.294243	0.294958	0.243
NOx (ppm)	384.121	379.852	1.111
CO(%Vol)	0.0295123	0.0299	1.31368
HC (ppm)	22.0585	23.314	5.691

Conclusion

In the present study, an effort has been made to find the effect of induction of LPG in a diesel engine with Karanja Biodiesel Oil (KBO) on engine performance and emissions parameters. Compared to the neat diesel operation, an overall reduction in the emissions levels has been observed with the introduction of LPG in the intake manifold at the cost of brake thermal efficiency. However, the values of BSFC tend to fall with increasing the supply of LPG in the engine cylinder. By analyzing both experimental data and optimized data, it is concluded that RSM perfectly optimizes the engine operating parameters by maximizing the BTE and minimizing the harmful exhaust emissions. An ANOVA study is used to demonstrate the impact of engine operating parameters which influences the engine performance and emission characteristics. A validation test is also carried out to confirm the experimental data and predicted data. The value of residuals that were calculated is in the range of acceptable limits. As per the findings of RSM optimizer, it was observed that best response values of BTE, BSFC, CO, HC, and NOx were 23.29%, 0.294 (kg/kWh), 0.029 (%vol), 22.058 (ppm), and 384.12 (ppm) respectively with optimal combination of operating parameters such as 20 (%vol) for KBO, 18.987 (R%) for LPG, and 53.838 (%) for the engine load.

Therefore, it can be concluded that response surface methodology (RSM) is successfully find the engine's optimum working parameters at which engine delivers maximum

performance by minimizing the harmful exhaust emissions. This investigation has a capability to help the engine manufactures and research workers to estimate the best operating factors for internal combustion engines at which engine performs efficiently by conducting less experiment trials.

The following conclusions can be made from this investigation:

1. BTE is increasing on increasing load due to usage of LPG but BTE decreasing on increasing the LPG supply at constant load.
2. BSFC is decreasing on increasing load due to usage of LPG but BSFC rising on increasing the LPG supply at constant load.
3. NOx emission first decreases with increasing LPG, and then NOx emission starts increasing at fixed load. NOx followed similar trend with biodiesel. Also, NOx emissions increased with load at fixed LPG or biodiesel.
4. CO emission decreasing with increasing biodiesel content. Moreover, CO emissions first decreased and then increased with increasing of LPG. CO increased with load at all ratios of fuels.
5. HC emissions raised with LPG and load, but reduced with biodiesel.
6. After considering all conclusions, this study recommends to keep inputs, load (%), LPG (R%), and biodiesel (%Vol), near to 53.84, 18.99, and 20 respectively to get best and optimized results.

Author contribution Sheel Bhadra: conceptualization, formal analysis, methodology, project administration, writing-original draft, writing-review and editing, and validation. Hansham Dewal: investigation, methodology, visualization, writing-original draft, and software. Naveen Kumar: resources, writing-review and editing, and supervision.

Declarations

Conflict of interest The authors declare that they have no competing interests.

References

Anye Ngang E, Ngayihi Abbe CV (2018) Experimental and numerical analysis of the performance of a diesel engine retrofitted to use LPG as secondary fuel. Appl Therm Eng 136:462–474. https://doi.org/10.1016/j.applthermaleng.2018.03.022

Arcoumanis C, Bae C, Crookes R, Kinoshita E (2008) The potential of di-methyl ether (DME) as an alternative fuel for compression-ignition engines: A review. Fuel 87(7):1014–1030. https://doi.org/10.1016/j.fuel.2007.06.007

Atmanli A, Yüksel B, Ileri E, Deniz Karaoglan A (2015) Response surface methodology-based optimization of diesel-n-butanol -cotton oil ternary blend ratios to improve engine performance and exhaust emission characteristics. Energy Convers Manag 90:383–394. https://doi.org/10.1016/j.enconman.2014.11.029

Bari S (1996) Effect of carbon dioxide on the performance of biogas/diesel duel-fuel engine. Renew Energy 9(1–4):1007–1010. https://doi.org/10.1016/0960-1481(96)88450-3

Bhatt AN, Shrivastava N (2021) Application of Artificial Neural Network for Internal Combustion Engines: A State-of-the-Art Review. Archives of Computational Methods in Engineering. Springer Science and Business Media B.V. https://doi.org/10.1007/s11831-021-09596-5

Chow JC (2001) Diesel Engines: Environmental Impact and Control. J Air Waste Manag Assoc 51(9):1258–1270. https://doi.org/10.1080/10473289.2001.10464354

Fukuda H, Kondo A, Noda H (2001) Biodiesel fuel production by transesterification of oils. J Biosci Bioeng 92(5):405–416. https://doi.org/10.1016/S1389-1723(01)80288-7

Guan C, Cheung CS, Ning Z, Wong PK, Huang Z (2017) Comparison on the effect of using diesel fuel and waste cooking oil biodiesel as pilot fuels on the combustion, performance and emissions of a LPG-fumigated compression-ignition engine. Appl Therm Eng 125:1260–1271. https://doi.org/10.1016/j.applthermaleng.2017.07.117

Liu J, Yang F, Wang H, Ouyang M, Hao S (2013) Effects of pilot fuel quantity on the emissions characteristics of a CNG/diesel dual fuel engine with optimized pilot injection timing. Appl Energy 110:201–206. https://doi.org/10.1016/j.apenergy.2013.03.024

Luft S (2001) "The Influence of Regulating Parameters of Dual Fuel Compression Ignition Engine Fuelled with LPG on its Maximum Torque, Overall Efficiency and Emission," SAE Technical Paper 2001-01-3264. https://doi.org/10.4271/2001-01-3264

Ma F, Hanna MA (1999) Biodiesel production: a review. Bioresour Technol 70(1):1–15. https://doi.org/10.1016/S0960-8524(99)00025-5

Papagiannakis RG, Hountalas DT (2003) Experimental investigation concerning the effect of natural gas percentage on performance and emissions of a DI dual fuel diesel engine. Appl Therm Eng 23(3):353–365. https://doi.org/10.1016/S1359-4311(02)00187-4

Patel RL, Sankhavara CD (2017) Biodiesel production from Karanja oil and its use in diesel engine: A review. Renew Sust Energ Rev 71:464–474. https://doi.org/10.1016/j.rser.2016.12.075

Peterson CL (1986) Vegetable Oil as a Diesel Fuel: Status and Research Priorities. Transactions of the ASAE 29(5):1413–1422. https://doi.org/10.13031/2013.30330

Sahoo BB, Sahoo N, Saha UK (2009) Effect of engine parameters and type of gaseous fuel on the performance of dual-fuel gas diesel engines-A critical review. Renew Sust Energ Rev 13(6–7):1151–1184. https://doi.org/10.1016/j.rser.2008.08.003

Salvi BL, Panwar NL (2012) Biodiesel resources and production technologies - A review. Renew Sust Energ Rev 16(6):3680–3689. https://doi.org/10.1016/j.rser.2012.03.050

Santhoshkumar A, Zahir Hussain A, Ramanathan A (2019) An experimental investigation of the effect of liquified petroleum gas addition on dual fuel diesel engine fuelled with pyrolysis waste engine oil. Materials Today: Proceedings 46:9800–9808. https://doi.org/10.1016/j.matpr.2020.10.881

Saravanan N, Nagarajan G, Sanjay G, Dhanasekaran C, Kalaiselvan KM (2008) Combustion analysis on a DI diesel engine with hydrogen in dual fuel mode. Fuel 87(17–18):3591–3599. https://doi.org/10.1016/j.fuel.2008.07.011

Selim MYE (2004) Sensitivity of dual fuel engine combustion and knocking limits to gaseous fuel composition. Energy Convers Manag 45(3):411–425. https://doi.org/10.1016/S0196-8904(03)00150-X

Zulqarnain Ayoub M, Yusoff MH, Nazir MH, Zahid I, Ameen M, Sher F, Floresyona D, Budi Nursanto E (2021) A comprehensive review on oil extraction and biodiesel production technologies.

Sustainability (Switzerland) 13(2):1–28. https://doi.org/10.3390/su13020788

Islam AKMA, Primandari SRP, Yaakob Z (2018) Non-Edible Vegetable Oils as Renewable Resources for Biodiesel Production: South-East Asia Perspective. Advances in Biofuels and Bioenergy. https://doi.org/10.5772/intechopen.73304

Raman R, Kumar N (2022) Comparative assessment of di-ethyl ether/diesel blends on the performance, and emission characteristics in acetylene dual fuel engine. International Journal of Engine Research. https://doi.org/10.1177/14680874221132325

Thangavel V, Subramanian B, Ponnusamy VK (2022) Investigations on the effect of H2 and HHO gas induction on brake thermal efficiency of dual-fuel CI engine. Fuel 126888. https://doi.org/10.1016/J.FUEL.2022.126888

Yessian S, Varthanan PA (2020) Optimization of Performance and Emission Characteristics of Catalytic Coated IC Engine with Biodiesel Using Grey-Taguchi Method. Sci Rep 10(1). https://doi.org/10.1038/s41598-019-57129-9

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.

Food Engineering Reviews
https://doi.org/10.1007/s12393-023-09335-5

Performance Assessment and Modeling Techniques for Domestic Solar Dryers

Shimpy[1] · Mahesh Kumar[1] · Anil Kumar[2,3]

Received: 7 December 2022 / Accepted: 19 January 2023
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

Abstract
Solar dryers have always been criticized for their lower performances. There are numerous ways to define the performance of a solar drying system such as thermal performance, drying kinetics, environmental aspects, economic evaluations, and quality of the dried product. Different modeling techniques have also been developed to design and analyze solar dryers and drying processes. This article presents a systematic, comprehensive, and state-of-the-art overview of various performance indicators and modeling techniques used for the evaluation and analysis of solar dryers, especially domestic and low cost solar dryers. Environmental analysis has severe global implications, and product quality is one of the biggest concerns of consumers. But the environmental impact and product quality assessments for domestic solar dryers are observed to be rarely reported in the literature. The use of modeling techniques in solar drying has changed the way of analyzing any thermal system. Here, an attempt is made to establish an overall assessment criterion for domestic solar dryers and to give a one-stop solution for researchers and users around the world.

Keywords Domestic solar dryers · Performance assessment · Modeling techniques · Environmental analysis · Dried product quality

Introduction

Drying is one of the most attractive techniques used traditionally throughout the globe to preserve food items. It does not just maintain the nutrition of the food but also helps in easy and long storage and transportation due to reduced weight [1]. The ever-increasing prices and shortage of fossil fuels and climate change resulting from the emission of greenhouse gases (GHGs) have forced researchers to explore alternatives to conventional energy demands as drying in itself is an energy-intensive process and contributes to 30% of the total processing cost of fresh produce [2–4].

✉ Mahesh Kumar
mkshandilya1@gmail.com

Shimpy
shimpymehra@gmail.com

Anil Kumar
anilkumar76@dtu.ac.in

[1] Department of Mechanical Engineering, Guru Jambheshwar University of Science & Technology, Hisar, India

[2] Department of Mechanical Engineering, Delhi Technological University, Delhi, India

[3] Centre for Energy and Environment, Delhi Technological University, Delhi, India

It is estimated that 8–10% of total emissions of GHGs are related to the food that is produced but not consumed which is about 17% of the food available at retail, food service, and consumer level. It has also been estimated by the Food and Agriculture Organization (FAO) of the United Nations that 690 million people were hungry in 2019 and this number would be increased by an additional 3 billion after the pandemic of COVID-19 [5].

Solar drying has emerged as one of the most attractive alternatives to replace conventionally powered dryers. Solar dryers not only reduce food loss or wastage but also remain unaffected by the problems of price rise and shortage of fuels. The environmentally friendly nature of solar dryers is one of the major reasons for their popularity among researchers [6–8]. Solar dryers are generally categorized on the basis of modes of flow of the drying fluid inside the dryer and heat transfer to the drying commodity. A new class of solar dryers called "Hybrid solar dryers" has also emerged in the last two decades and got much popularity among researchers [9, 10]. Figure 1 shows the categorization of solar dryers.

The energy required to operate a solar dryer comes from the ultimate source of energy i.e., "Sun". This solar energy can be used in three ways to raise the temperature of the food item. The first is the direct way when the solar radiation

Published online: 31 January 2023

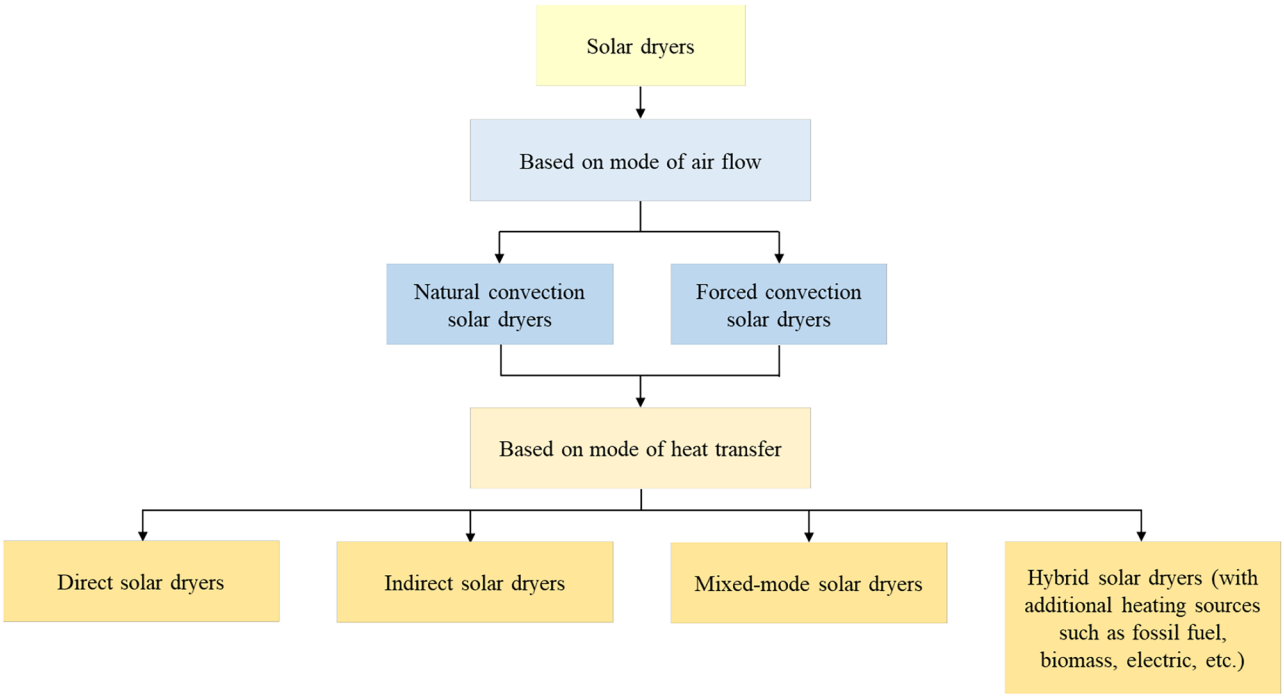


Fig. 1 Categorization of solar dryers [11, 12]

strikes the product directly and gets absorbed, resulting in temperature rise and moisture evaporation. The second way is using solar radiation to raise the temperature of a drying fluid separately and then using that heated fluid to evaporate moisture from the commodity placed in a well-insulated drying chamber. The third way combines both the direct and indirect ways of heat transfer and is called the mixed-mode type of solar drying [13–15]. The evaporated moisture from the drying commodity has to be removed from the drying system, and it can be done either by natural air currents due to the buoyancy effect known as "natural convection" or by some external source of air circulation such as a fan and blower known as "forced convection" [16, 17].

Many researchers have developed various solar dryers since 1976 when Everitt and Collins first introduced the idea of solar dryers. Each dryer has some modifications in terms of design and operation to improve the performance of the dryer and/or the quality of the dried commodity. The performance of a solar dryer was estimated in many investigations and has been shown in terms of various indicators such as thermal efficiency, drying rate, drying time, and energy consumption [18, 19]. For better understanding, control, and higher outputs, solar dryer is being analyzed by using various modeling techniques such as numerical and simulation, and the drying processes are also being modeled and studied by using various mathematical models for highest quality products [20, 21].

Earlier, Shimpy et al. [11] have reviewed various developments in the designs of domestic solar dryers including their performance and economic feasibility. The literature indicates that a range of domestic solar dryers has been developed and tested by researchers for different commodities. The performance evaluation methods and various modeling techniques have been used in different studies to analyze and compare the dryers, drying processes. and dried products. The present article is mainly focused on the present status of domestic solar dryers in terms of performance assessment parameters and modeling techniques employed to understand their potential for the drying of various household commodities.

Methodology

The literature on solar drying is very vast and diversified. In the present study, keywords such as domestic, household, and small-scale solar dryers were searched from the literature in search engines, namely, Web of Science (WOS), Google Scholar (GS), and Dimensions (D). The details of search results have been given in Table 1. Many studies were available on all the search engines. So, the total results for consideration were comparatively lower, and among them, the most relevant studies were scrutinized on the basis of keywords and the relevancy of the content to the present study. A flow chart for the methodology of present study has been shown in Fig. 2.

Table 1 Publications on domestic solar drying

Solar dryer	Search results		
	Web of Science (WOS)	Google Scholar (GS)	Dimensions (D)
Domestic	72	191	135
Household	39	37	29
Small scale	1	131	8

Performance Parameters and Analysis

"How good a solar dryer performs the drying operation" shows its performance. This goodness can be seen from different perspectives such as thermal, drying kinetics, environmental, economic, and quality of the dried products. This section presents an overview of the performance of domestic solar dryers from different perspectives.

Thermal Performance Parameters and Analysis of Domestic Solar Dryers

Pickup Efficiency

It is the ratio of actual moisture evaporated during the drying process to the total moisture removal capacity of the drying air. It is also called the effectiveness of a solar dryer [22].

$$\eta_{pickup} = \frac{m_{ev}}{m_a \times t\left(h_{sat} - h_{in}\right)} \quad (1)$$

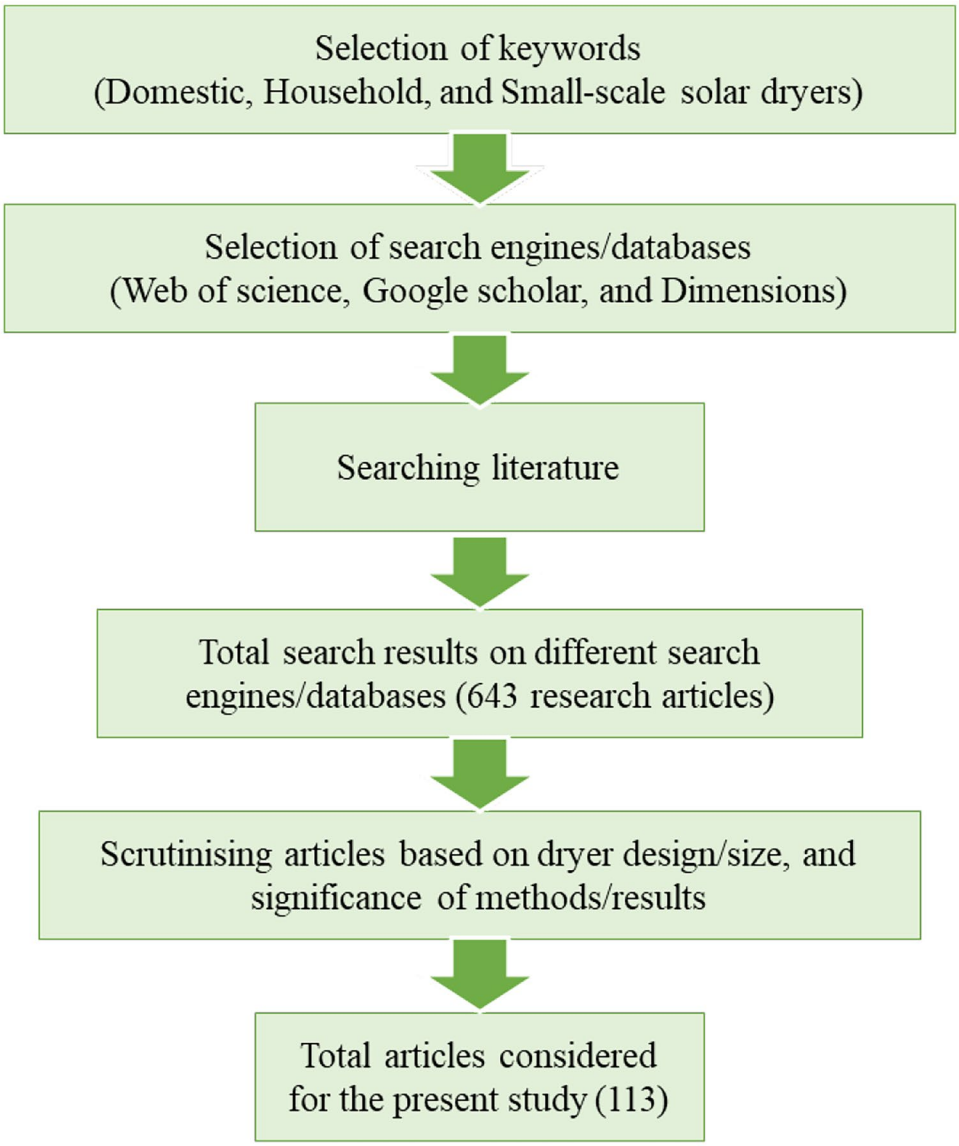


Fig. 2 Flow chart for the methodology of present study

where m_a = Mass flow rate of drying air (kg/s), t = time (s), h_{sat} = adiabatic saturation humidity of the air entering the dryer (kg water/kg dry air), and h_{in} = absolute humidity of the air entering the dryer (kg water/kg dry air).

Thermal Efficiency

Thermal efficiency is one of the most used performance indicators for any solar dryer. It can be calculated as the ratio of energy used in the evaporation of moisture content from the drying commodity (E_{ev}) to the total energy incident on the absorbing surface of the dryer (E_{in}). It is also known as system efficiency or drying efficiency [23, 24].

$$\eta_{th} = \frac{E_{ev}}{E_{in}} \tag{2}$$

Energy Efficiency

Energy efficiency is an indicator of the wellness of energy utilization of a system or in other words of any unaccounted energy losses. It is the ratio of the total energy output (E_{out}) to the total energy input (E_{in}) [25, 26]:

$$\eta_{en} = \frac{E_{out}}{E_{in}} \tag{3}$$

Exergy Efficiency

Exergy is the maximum amount of work that can be produced by a system before attaining equilibrium. The exergy efficiency of a solar dryer can be calculated as the ratio of exergy output (Ex_{out}) to the exergy input (Ex_{in}) [23, 27]:

$$\eta_{ex} = \frac{Ex_{out}}{Ex_{in}} \tag{4}$$

The exergetic performance of a solar dryer can also be described by using different indicators such as exergy loss (E_{loss}), waste exergy ratio (*WER*), improvement potential (*IP*), and sustainability index (*SI*) [28, 29]:

$$E_{loss} = Ex_{in} - Ex_{out} \tag{5}$$

$$\text{WER} = \frac{E_{loss}}{Ex_{in}} \tag{6}$$

$$\text{IP} = \left(1 - \eta_{ex}\right) \times E_{loss} \tag{7}$$

$$\text{SI} = \frac{1}{1 - \eta_{ex}} \tag{8}$$

Specific Energy Consumption

Specific energy consumption (SEC) is an indicator of the amount of energy required by a solar dryer to evaporate a unit mass of moisture content from a drying commodity. It is the ratio of the total energy input to a dryer to the amount of moisture evaporated from the drying commodity [30, 31]:

$$\text{SEC} = \frac{E_{in} \times t}{m_{ev}} \tag{9}$$

Specific Moisture Extraction Rate

The reciprocal of specific energy consumption is considered as the specific moisture extraction rate (*SMER*). It shows the amount of moisture evaporated per unit of energy supplied [22, 32]:

$$\text{SMER} = \frac{m_{ev}}{E_{in} \times t} \tag{10}$$

Heat Utilization Factor

Heat utilization factor (*HUF*) is the ratio of heat utilized during the drying process to the total heat generated inside the dryer during the operation [33–35]:

$$\text{HUF} = \frac{T_f - T_{air}}{T_f - T_{amb}} \tag{11}$$

where T_f = temperature of the floor of the drying chamber, T_{air} = temperature of the air insider drying chamber, and T_{amb} = temperature of ambient air.

Coefficient of Performance

Coefficient of performance (*COP*) is the ratio of available useful heat to the total heat generated inside the solar dryer [34, 36].

$$\text{COP} = \frac{T_{air} - T_{amb}}{T_f - T_{amb}} \tag{12}$$

It can also be given as

$$\text{COP} = 1 - \text{HUF} \tag{13}$$

Convective Heat Transfer Coefficient

Newton's law of cooling states that the rate of convective heat transfer ($\dot{Q}$) is directly proportional to the temperature difference (ΔT) between the body and the surrounding air.

The constant used for the elimination of proportionality sign is known as the convective heat transfer constant (h_c). The value of h_c depends on different variables such as thermophysical properties of drying fluid, type of fluid flow, geometry, and surface roughness of the solid surface [37].

$$\dot{Q} \propto \Delta T \tag{14}$$

$$\dot{Q} = h_c \times \Delta T \tag{15}$$

In the case of solar drying, the value h_c is generally calculated by using the Nusselt number (Nu). It is a non-dimensional number that shows the ratio of the rate of heat transfer between a solid and fluid by convection to conduction.

For natural convection mode [38]:

$$Nu = \frac{h_c L_c}{K_v} = C(Gr\ Pr)^n \tag{16}$$

$$h_c = \frac{K_v}{L_c} C(Gr\ Pr)^n \tag{17}$$

For forced convection mode [39]:

$$Nu = \frac{h_c L_c}{K_v} = C(Re\ Pr)^n \tag{18}$$

$$h_c = \frac{K_v}{L_c} C(Re\ Pr)^n \tag{19}$$

where L_c = characteristic length, K_v = thermal conductivity of the fluid, Gr = Grashof number, Pr = Prandtl number, and C and n = experimental constants.

The rate of heat required to evaporate the moisture from the drying commodity (Q_e) is one of the important parameters that can be given as follows [38, 40]:

$$Q_e = 0.016\ h_c \left[P(T_v) - \gamma P(T_e)\right] \tag{20}$$

where $P(T)$ = partial pressure at the temperature (T), γ = relative humidity, T_v = temperature of the drying commodity, and T_e = temperature of surrounding air of the commodity.

Evaporative Heat Transfer Coefficient

The rate of heat transfer from a drying commodity to the surroundings due to the evaporated moisture is the evaporative heat loss that is majorly governed by a parameter called evaporative heat transfer coefficient (h_e) that can be calculated by the following expression [25, 27, 41]:

$$h_e = 0.016\ h_c \left[\frac{P(T_v) - \gamma P(T_e)}{T_v - T_e}\right] \tag{21}$$

In literature, the thermal performance of domestic solar dryers has been reported in many investigations. Sharma et al. [42] evaluated the energy and area required for the moisture evaporation from peas, grapes, and potatoes using a cabinet type natural convection direct solar dryer (NCDSD) as 5.15, 5.48, and 2.56 kWh/kg and 0.91, 0.96, and 0.45 m^2, respectively. Figure 3 shows the schematics of the cabinet type natural convection solar dryer.

Singh et al. [43] calculated the values of solar energy required/kg of moisture for fenugreek leaves and reported that the energy requirement increased with the drying time, i.e., 2.47, 3.30, and 13.44 kWh/kg for the first, second, and third day of drying. Saleh and Badran [44] tested a domestic type NCDSD for the drying of Jew's mallow and reported that the values of average specific energy consumption (SEC) under fixed and solar tracking modes were 17.78 and 9.17 kWh/kg, respectively. It was observed that the values of SEC increased with the drying time. Haque et al. [45] evaluated the value of exergy efficiency and improvement

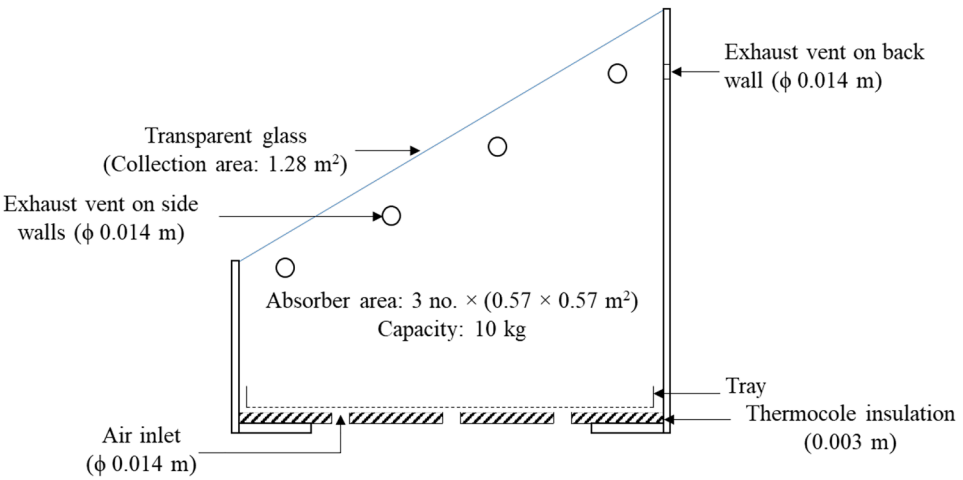


Fig. 3 Cabinet type natural convection direct solar dryer [42]

potential of a portable domestic type NCDSD in the range of 17–44% and 0.25–184 W, respectively. Jain et al. [35] studied a domestic type NCDSD and observed that the convective heat transfer coefficient (h_c), heat utilization factor (*HUF*), and coefficient of performance (*COP*) ranges between 2.44 and 2.81 W/m^2 ℃, 0.54 and 0.69, and 0.31 and 0.46, respectively. Nabnean and Nimnuan [46] calculated the value of SEC of a domestic type forced convection direct solar dryer (FCDSD) for the drying of banana slices as 5.88 kWh/kg. Tiwari [47] developed an FCDSD with a semi-transparent photovoltaic module as the glazing for drying bitter gourd flakes. The values of h_c varied between 0.69 and 14.45 W/m^2 K with an overall thermal gain of 5.41 kWh/m^2. Moghimi et al. [48] reported that the SEC for tomato drying under forced convection indirect solar dryer (FCISD) (Fig. 4) was 4.24 kWh/kg of which electricity consumption was only 10% (0.424 kWh/kg).

Sharma et al. [49] tested an FCISD for the drying of tomato slices and evaluated the values of exergy efficiency, improvement potential, waste exergy ratio, and sustainability index in the range of 32.86–58.26%, 0.006966–0.065984, 0.41–0.67, and 1.55–2.39, respectively. The average exergy loss was estimated to be 56.56 W. The efficiency of domestic solar dryers used to dry various products is given in Table 2.

Drying Kinetics and Analysis of Domestic Solar Dryers

Moisture Content

The amount of moisture present in a commodity can be expressed in two ways, i.e., wet basis (wb) or dry basis (db) as given below [34]:

$$MC\ (wb) = \frac{m_i - m_f}{m_i} \times 100 \tag{22}$$

$$MC\ (db) = \frac{m_i - m_f}{m_f} \times 100 \tag{23}$$

where m_i = initial mass and m_f = final mass.

Moisture Ratio

Moisture ratio (*MR*) is one of the most significant parameters used to understand the drying characteristics of a drying commodity. The values of moisture ratio can be calculated as follows [66]:

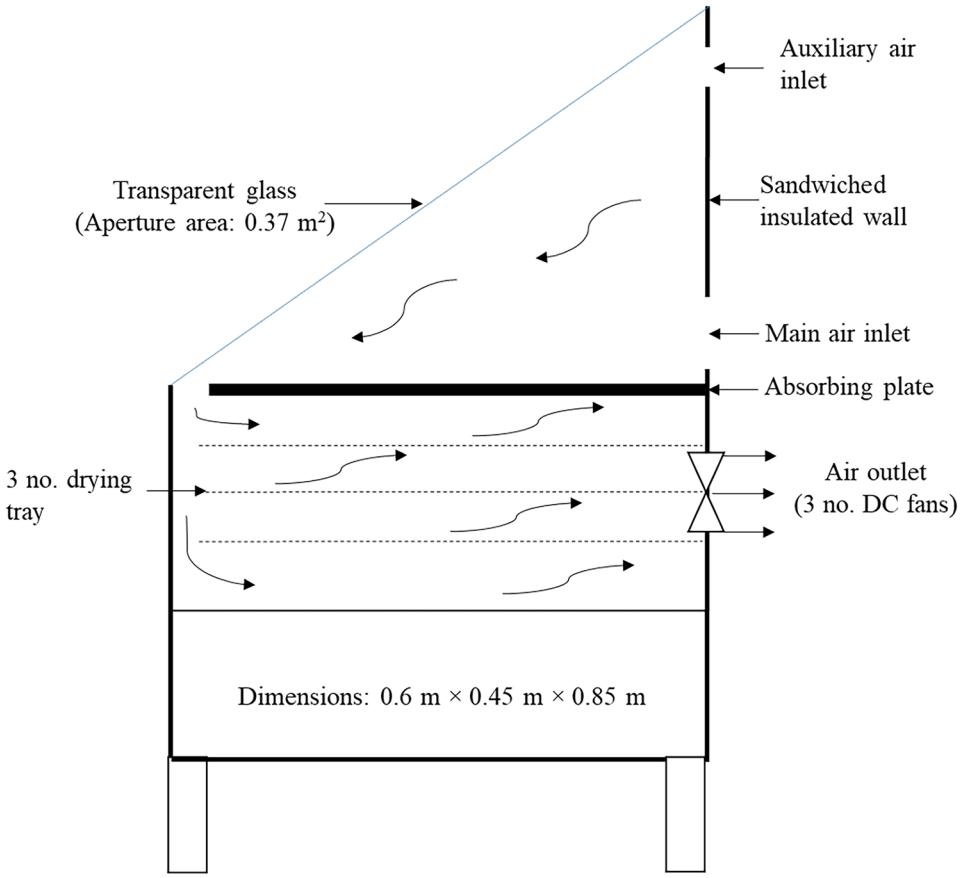


Fig. 4 Forced convection indirect solar dryer [48]

Table 2 Thermal efficiency of domestic solar dryers for different commodities

Researcher/s	Dryer	Commodity	Thermal efficiency (%)
Singh et al. [43]	NCDSD	Fenugreek leaves	Batch mode: 28.05 Semi-continuous mode: 25.6
Saleh and Badran [44]	NCDSD	Jew's mallow	Fixed mode: 3.74 Tracking mode: 7.08
Haque et al. [45]	NCDSD	Bitter gourd, Okra, Hindra Raw mango	10.73
Tiwari [47]	FCDSD	Bitter gourd flakes	49–70
Moghimi et al. [48]	FCISD	Tomato slices	16.4
Sharma et al. [49]	FCISD	Tomato slices	4.04–68.78
Hallak et al. [50]	Natural convection indirect solar dryer (NCISD)	Mango slices	26–65
Thanvi and Pande [51]	NCDSD	Chilies	37.9
Bolaji [52]	NCISD	No-load	60.5
Ezekoye and Enebe [53]	NCDSD	Pepper and Groundnuts	22
Eke [54]	NCDSD	Tomato, carrot, and okra	Tomato: 21.8 Carrot: 19.96 Okra: 24.95
Eke [55]	NCDSD with different collector materials	Tomato	Wood: 19.56 Cement: 20.25 Mud: 20.91 Metal: 27.24
Navale et al. [56]	NCDSD	Fenugreek leaves	34.5
Borah et al. [57]	NCDSD	Turmeric	Whole: 55.36 Sliced: 55.6
Subedi and Bhattarai [58]	NCDSD	Ginger	*Slice thickness* 3–5 mm: 64.36 6–8 mm: 58.86 9–11 mm: 50.97 Whole: 32.84
Chaudhari and Bhavsar [59]	NCDSD cum cooker with and without electrical heating	Green chilies	With electrical heating: 19.50 Without electrical heating: 19.12
Modi et al. [60]	FCDSD	Tomato slices	34.87
Chaudhari et al. [61]	NCDSD	Ginger	19.71
Poonia et al. [62]	FCDSD	Ber (*Zizyphus mauritiana*)	16.7
Poonia et al. [63]	NCDSD and FCDSD	Ber (*Zizyphus mauritiana*)	NCDSD: 15.6 FCDSD: 16.7
Nimnuan and Nabnean [64]	FCDSD	Alpinia galangal	32
Safri et al. [65]	FCISD	No-load	66.34

$$MR = \frac{MC_t - MC_e}{MC_i - MC_e} \tag{24}$$

where MC_t = moisture content at time t, MC_e = equilibrium moisture content, and MC_i = initial moisture content.

The values of MC_e are generally very small as compared to the initial moisture content of food items and hence can be neglected, and the value of moisture ratio can be given as the ratio of moisture content in a commodity at a particular time to the initial moisture content. This parameter is of very high significance as it is widely used for establishing mathematical models to describe the drying behavior of various commodities [67, 68].

$$MR = \frac{MC_t}{MC_i} \tag{25}$$

Drying Rate

The amount of moisture removed from a commodity in a unit time is considered as the drying rate (*DR*). It is generally used to specify the drying capacity of a dryer per unit time and may vary depending upon the mode of operation, i.e., natural or forced, temperature of drying air and type, shape, and size of the drying commodity [49]:

$$DR = \frac{dMC_{t+dt} - MC_t}{dt} \tag{26}$$

where dMC_{t+dt} = moisture content at time t + dt.

Effective Moisture Diffusivity

The entire drying process can be divided in two phases, i.e., the constant and falling rate phases. The drying of most of the agricultural and other food items generally lies under the falling rate period. The period in which the moisture removal is governed by the rate of internal moisture transportation phenomenon called as "moisture diffusivity." This internal moisture diffusion can be a result of different mechanisms viz. capillary action, molecular diffusion, liquid and vapor diffusion through solid and air-filled pores, respectively, vaporization–condensation sequence and hydrodynamic flow; and change in shape, size and texture of the material. The collective action of all the moisture transportation mechanisms is termed as "effective moisture diffusivity" [69].

The solution of Fick's second law in the form of a reduced exponential model was used for the determination of effective moisture diffusivity [70–72].

For sphere:

$$MR = \frac{6}{\pi^2} \times \exp\left(-\frac{\pi^2 D_{ef}}{r^2} t\right) \tag{27a}$$

For cylinder:

$$MR = \frac{4}{b_n^2} \times \exp\left(-\frac{b_n^2 D_{ef}}{r^2} t\right) \tag{27b}$$

For slab:

$$MR = \frac{8}{\pi^2} \times \exp\left(-\frac{\pi^2 D_{ef}}{4L^2} t\right) \tag{27c}$$

where D_{ef} = effective diffusion coefficient (m^2/s), r = half of the thickness of the sample (mm), t = drying time in (s), and b_n^2 = characteristic root of first kind and zero order Bessel functions ($b_1 = 2.4048$).

In the case of longer drying operations, equation can be expressed in logarithmic form as given below:

For sphere:

$$ln(MR) = ln\left(\frac{6}{\pi^2}\right) - \left(\frac{\pi^2 D_{ef}}{r^2} t\right) \tag{28a}$$

For cylinder:

$$ln(MR) = ln\left(\frac{4}{b_n^2}\right) - \left(\frac{b_n^2 D_{ef}}{r^2} t\right) \tag{28b}$$

For slab:

$$ln(MR) = ln\left(\frac{8}{\pi^2}\right) - \left(\frac{\pi^2 D_{ef}}{4L^2} t\right) \tag{28c}$$

The value of D_{ef} can be obtained from the slope of the line obtained from the plot of ln(MR) with respect to the drying time (t).

For sphere:

$$Slope = \frac{\pi^2 D_{ef}}{r^2} \tag{29a}$$

For cylinder:

$$Slope = \frac{b_n^2 D_{ef}}{r^2} \tag{29b}$$

For slab:

$$Slope = \frac{\pi^2 D_{ef}}{4L^2} \tag{29c}$$

Activation Energy

Activation energy is the energy required to initiate the diffusion of the moisture within the drying commodity during the process of drying. It can be a significant criteria to design a solar dryer for a particular product or a class of products. Arrhenius equation correlates the effective moisture diffusivity with the absolute drying air temperature [73, 74]:

$$D_{ef} = D_0 \times exp\left(-\frac{E_A}{\mathrm{RT}}\right) \tag{30}$$

where E_A = activation energy (J/mole), D_0 = pre-exponential factor of the Arrhenius equation (m^2/s), R = ideal gas constant (8.314 J/mole K), and T = drying air temperature in (K).

The natural log of Eq. (30) can be given as

$$ln\left(D_{ef}\right) = ln\left(D_0\right) - \left(\frac{E_A}{\mathrm{RT}}\right) \tag{31}$$

The slope of line obtained from the plot of $ln(D_{ef})$ against the inverse of absolute temperature of the drying air (1/T).

$$Slope = \frac{E_A}{R} \tag{32}$$

Various researchers have used drying kinetics for domestic solar drying. Ezekoye and Enebe [53] reported the drying rate for groundnuts under NCDSD and open sun drying as 0.198 and 0.1 g/day, respectively. Alonge and Adeboye [75] evaluated the drying rate for pepper, okra, and vegetables under NCDSD as 3.94, 17.65, and 13.33 kg/hour, respectively. Eke [55] reported that an NCDSD with metal, wood, cement, and mud as absorbing materials took 76, 96, 96, and 94 h, respectively, for the drying of tomatoes. Borah et al. [57] evaluated the values of effective moisture diffusivity for whole and sliced turmeric samples under NCDSD as 1.456×10^{-10} and 1.852×10^{-10} m^2/s, respectively. Poonia et al. [62] reported that the moisture diffusivity of ber (*Zizyphus mauritiana*) was 3.34×10^{-7} m^2/s under an FCDSD (Fig. 5). Islam et al. [76] developed an NCDSD having three drying chambers with thin tube chimney, attic, and simple-type ventilation arrangements and tested for pineapple, apple, banana, and guava fruits drying. The simply ventilated chamber showed highest moisture removal rate (58.9%) as compared to the chimney type (44.5%) and attic type (33.3%). A comparison of the drying kinetic performance parameters of different domestic solar dryers has been presented in Table 3.

Environmental Performance Parameters and Analysis of Domestic Solar Dryers

Embodied Energy (EE)

Energy invested in the development of a product is considered as the embodied energy. The coefficient of embodied energy

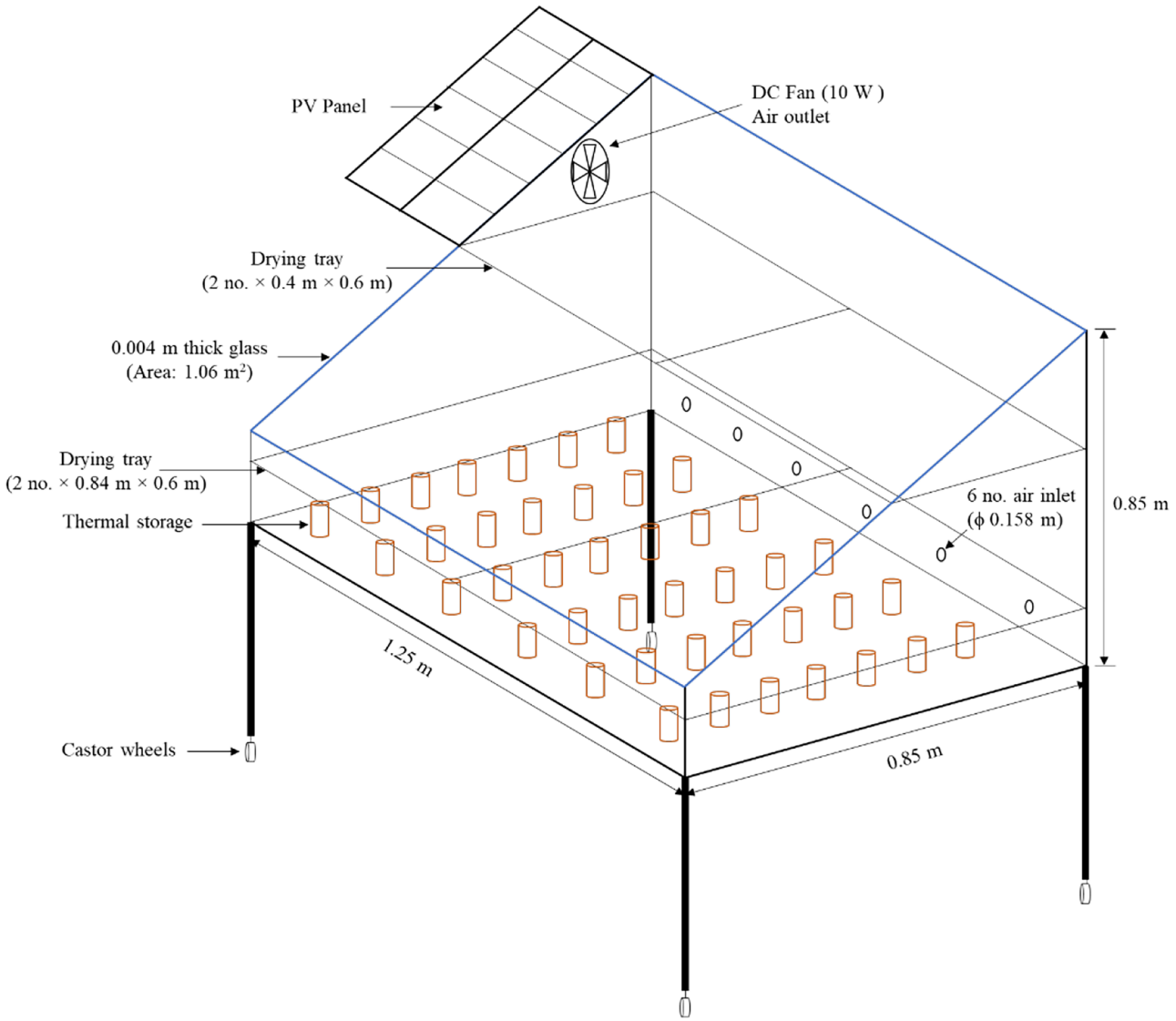


Fig. 5 PVT hybrid solar dryer [62]

Table 3 Drying kinetics parameters of different domestic solar dryers

Researcher (Year)	Dryer	Commodity	Moisture content (% wb)		Drying time (hours)
			Initial	Final	
Sharma et al. [42]	NCDSD	Peas	70	11.2	3 days
		Grapes	75	15	
		Potatoes	40	10.6	
Singh et al. [43]	NCDSD	Chilies	-	-	18.75
		Fenugreek leaves	88.5	7.3	
Haque et al. [45]	NCDSD	Bitter gourd	89	7.6	8
		Okra	88.4	5	9
		Raw mango	88.8	5.7	9
		Hindra	68.6	10	6
Nabnean and Nimnuan [46]	FCDSD	Banana slices	72	28	4 days
Sharma et al. [49]	FCISD	Tomato slices	95	9	10
Thanvi and Pande [51]	NCDSD	Chilies	86.3	4.1	9 days
Eke [54]	NCDSD	Tomato	93	4	34
		Carrot	88	5	35
		Okra	88	4	28
Eke [55]	NCDSD	Tomato	90	4	76–94
Borah et al. [57]	NCDSD	Turmeric	78.65	5.5–6.36	12
Chaudhari and Bhavsar [59]	Hybrid NCDSD	Green chilies	80.58	6.89	9
Modi et al. [60]	FCDSD	Tomato slices	90.48	7.73	5.5
Chaudhari et al. [61]	NCDSD	Ginger	79.31	6.73	9
Poonia et al. [62]	FCDSD	Ber (*Zizyphus mauritiana*)	80	20	240
Nimnuan and Nabnean [64]	FCDSD	Alpinia galangal	89	12	1 day
Alonge and Adeboye [75]	NCDSD	Pepper	78.9	24	33
		Okra	92	20	51
		Vegetables	90	20	30
Sreekumar et al. [77]	FCISD	Bitter gourd	95	5	6
Rawat et al. [78]	NCDSD	Chilies	74.4	4.4	8
Alonge and Omoniwa [79]	NCDSD	Cassava chips	53.7	11	32
Seveda and Jhajharia [80]	FCDSD	Large cardamom	75.6	10.1	24
Daud and Simate [81]	NCDSD	Pineapple slices	85.12	12.23	8
Chavan et al. [82]	NCDSD	Potato slices	80	5	-
Dubey et al. [83]	FCMMSD	Grapes	80	10	1 day

is calculated by considering all the energy inputs during the development of the product. The value of EE can be evaluated by multiplying the coefficient of embodied energy to the weight of the product as follows [28]:

$$\text{EE} = \text{Coefficeint of embodied energy} \times \text{product weight} \tag{33}$$

Energy Payback Time (EPBT)

It is the time required by the product to deliver the amount of required work equivalent to the energy spent during the development of that product and can be calculated as follows [35]:

$$EPBT = \frac{EE}{Annual\ energy\ output} \tag{34}$$

where annual energy output is given as follows [24]:

$$Annual\ energy\ output = Daily\ energy\ output \times Operating\ days/year \tag{35}$$

The daily energy output can be calculated as follows [84]:

$$\text{Daily energy output} = \frac{m_{ev} \times \lambda}{3.6 \times 10^6} \tag{36}$$

where λ = latent heat of vaporization.

CO_2 Emissions

The amount of annual CO_2 emissions by a solar dryer can be given by considering the electricity production from coal. Generally, taken as 0.98 kg/kWh of CO_2 [85]:

$$CO_2 \text{ emission per year} = \frac{EE \times 0.98}{Lifetime} \tag{37a}$$

Considering the losses associated with the electricity such as transmission and distribution losses (L_{td}) and domestic appliance losses (L_{da}), the Eq. (37a) can be given as follows:

$$CO_2 \text{ emission per year} = \frac{EE \times 0.98}{Lifetime} \times \frac{1}{1 - L_{td}} \times \frac{1}{1 - L_{da}} \tag{37b}$$

The values of L_{td} and L_{da} are generally considered as 0.4 and 0.2, respectively. So Eq. (37b) can be written as follows:

$$CO_2 \text{ emission per year} = \frac{EE}{Lifetime} \times 2.042\ kg \tag{37c}$$

CO_2 Mitigation Potential

The CO_2 mitigation potential of the dryer for its entire life can be given as follows [24]:

$$CO_2\ mitigation = (Annual\ energy\ output \times Life - EE) \times 2.042\ kg \tag{38}$$

Carbon Credit Earned (CCE)

For the international trading of energy systems specifically renewable energy systems, carbon credit earned is an important environmental sustainability indicator that can be given as follows [86]:

$$CCE = CO_2 \text{ mitigation} \times Cost\ of\ one\ carbon\ credit \tag{39}$$

where the cost of carbon credit varies from USD 5–20/tones of CO_2 mitigation.

Environmental performance parameters indicate the contribution of the developed solar drying system to the global climate change and environment. Rawat et al. [78] reported that for a working life of 10 years, an NCDSD can save up to 2486.40 kg of fuel wood, 1776 L of light diesel oil, 2072 kg of coal, and 1554 kg of natural gas, respectively. Table 4 presents a quick overview of the environmental performance of different domestic solar dryers.

Economic Performance Parameters and Analysis of Domestic Solar Dryers

Life Cycle Cost

Life cycle cost (LCC) is the total cost involved during the entire life of the dryer and is given as follows [62]:

$$LCC = C_{ic} + C_{lom} - \text{SV} \tag{40}$$

where C_{ic} = initial cost; C_{lom} = cost of labor, operation, and maintenance; and SV = salvage value of the dryer at the end of its life.

Life Cycle Benefit (LCB)

It is the total benefit that can be achieved during the entire life of the solar dryer [62]:

$$LCB = R \times \frac{X(1 - X^j)}{(1 - X^j)} \tag{41}$$

where R = annual benefit, $X = \frac{1 + annual\ escalation}{1 + interst\ rate}$, and j = Lifetime.

Payback Period (PBP)

The payback period (PBP) is the time period required by the developed system to recover the amount equivalent to that is spent for its fabrication. Following expression can be used for the evaluation of PBP for any drying system [47, 88]:

$$PBP = \frac{\ln\left[1 - \frac{C_{cc}}{S_1}(i - d)\right]}{\ln\left(\frac{1+d}{1+i}\right)} \tag{42}$$

Table 4 Environmental performance of various domestic solar dryers

Study	Drying method	Product	EE (kWh)	EPBT (years)	Annual CO_2 emission (kg)	CO_2 mitigation (kg)	CCE ($)
Jain et al. [35]	NCDSD	-	339.015	7.57	16.62	1553.73	30.34
Rawat et al. [78]	NCDSD	Chillies	762.4	1–2	-	-	-
Rawat et al. [87]	NCDSD	Amla (*Phyllanthus emblica* L.)	762.4	3–5	-	8812.4–34,368.3	-

where C_{cc} = capital cost, S_1 = savings after 1 year, i = interest rate, and d = inflation rate.

The various parameters required for calculating PBP has been given in Appendix A.

Economics is one of the most important considerations for domestic and small-scale users. Many researchers have focused on the economic assessment of domestic solar dryers. Singh et al. [43] reported the present cumulative worth of a domestic type NCDSD for fenugreek leaves as $236.26. Sreekumar et al. [77] calculated the annualized cost and net present worth of a cabinet type FCISD for the drying of bitter gourd as $11.86 and $408.23, respectively. Mustapha et al. [89] used plastic, mosquito net, glass, aluminum, and glass with black pebbles in five different solar dryers for the drying of fish. The value of cost to benefit ratio varied from 2.5:1 to 4.5:1 for different dryers. Modi et al. [60] obtained a net profit of $4.7/kg by using a low cost cabinet type FCDSD for the drying of tomatoes. Chaudhari et al. [61] evaluated the values of net present worth and benefit to cost ratio for the drying of ginger under NCDSD as $266.92/year and 2.3, respectively. Haque et al. [45] calculated the values of annualized cost and cumulative present worth for a portable domestic type NCDSD (Fig. 6) as $12.12 and $1216.55, respectively.

Poonia et al. [62] reported that the life cycle cost and benefit of a PV/T enabled hybrid FCDSD were $556.44 and $1033.77, respectively. The values of benefit to cost ratio, net present worth, annuity, and internal rate of return were evaluated as 1.86, $477.33, $64.33, and 54.5%, respectively, for the drying of ber. Sandali et al. [90] appraise the life cycle cost and benefit of NCDSD with various heat supplying techniques as $28,843.37 and $37,017.44, respectively. A quick overview of the initial cost and the payback period of different solar dryers for different products has been presented in Table 5.

Product Quality Parameters and Analysis of Domestic Solar Dryers

Sensory Evaluation

Sensory evaluation is one of the most widely used quality assessment methods used for solar dried food products. It involves the assessment of change in color, texture, taste, and flavor of the food products after drying [20, 46].

Color Deviation

Color deviation (C_{dev}) is the variation in the color of the solar dried samples with reference to the color of the fresh samples. The two most commonly used methods to analyze the color deviation of food items are L^* a^* b^* and L^* C^* h^* methods. The value of C_{dev} can be calculated as follows [45]:

$$C_{dev} = \sqrt{\left(\left(L^* - L_r^*\right)^2 + \left(a^* - a_r^*\right)^2 + \left(b^* - b_r^*\right)^2\right)} \quad (43)$$

where L^* = lightness, a^* = red/green coordinate, b^* = yellow/blue coordinate, C^* = chroma, h^* = hue, and subscript (r) shows the respective reference values.

The values of C^* and h^* can be calculated as given below [46]:

$$C^* = \sqrt{a^{*2} + b^{*2}}$$

$$h^* = \tan^{-1}\left(\frac{b^*}{a^*}\right) \; if \; a^* > 0$$

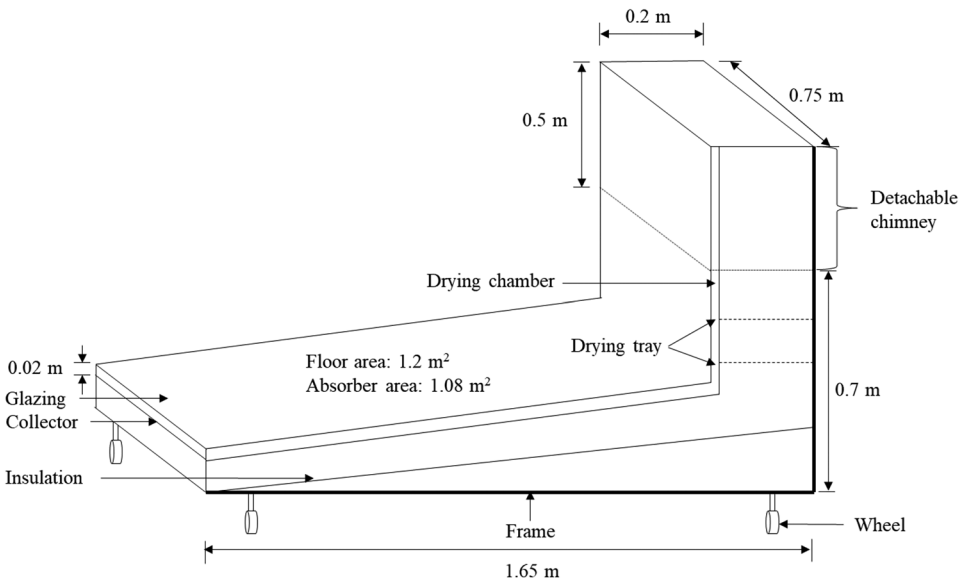


Fig. 6 Schematics of a portable domestic solar dryer [45]

Table 5 Initial cost and payback period of various domestic solar dryers

Study	Dryer	Product	Initial cost ($)	Payback period (years)
Singh et al. [43]	NCDSD	Fenugreek leaves	35.55	1.36
		Chilies		0.57
Haque et al. [45]	NCDSD	Bitter gourd, okra, hirda, and raw mango	77.38	0.56
Nabnean and Nimnuan [46]	FCDSD	Banana slices	390	1.1
Moghimi et al. [48]	FCISD	Tomato	306.6	-
Thanvi and Pande [51]	NCDSD	Chilies	3.61	-
Eke [55]	NCDSD	Tomato slices	12.05–36.14	-
Sreekumar et al. [77]	FCISD	Bitter gourd	83.81	3.26
Alonge and Omoniwa [79]	NCDSD	Cassava chips	500	-
Mustapha et al. [89]	NCDSD	Fish	4.34–8.19	0.25
Sandali et al. [90]	NCDSD	No-load	-	0.8–5.4
Tefera et al. [91]	NCDSD	Potato slices	For box type: 39.92	-
			For pyramid type: 10.94	
Poonia et al. [92]	NCDSD	Ber (*Ziziphus*)	10.32	-
Chaudhari et al. [61]	NCDSD	Ginger	116.05	0.5
Poonia et al. [62]	FCDSD	Ber (*Ziziphus mauritiana*)	180.52	2.26
Nimnuan and Nabnean [64]	FCDSD	Galangal	375	0.9

$$h^* = 180 + \tan^{-1}\left(\frac{b^*}{a^*}\right) \; if \; a^* < 0$$

Ash Content

It is calculated in terms of percentage as the ratio of the total weight of ash content remained after complete combustion of the sample at or above 500 ℃ to the initial weight of the sample [83].

$$Ash\ content\ (\%) = \frac{Weight\ of\ ash}{Initial\ weight\ of\ sample} \times 100 \quad (44)$$

Rehydration Ratio

A measure to estimate the loss of tissues during the drying process is the rehydration ratio which can be given as the ratio of the weight of the sample after rehydration and the weight of the dried sample. It is also known as the rehydration capacity or hydration coefficient [93].

$$Rehydration\ ratio = \frac{Weight\ of\ rehydrated\ sample}{Weight\ of\ the\ dried\ sample} \quad (45)$$

Shrinkage

The percentage of change in dimension after solar drying to the initial dimension of a product is called as the shrinkage [94, 95]:

$$Shrinkage\ (\%) = \frac{D_i - D_f}{D_i} \times 100 \quad (46)$$

where D_i = initial dimension and D_f = final dimension.

Nutritional Analysis

As the drying process can affect various nutritive properties of products, the evaluation of various nutrients (such as carbohydrates, fat, sugar, and vitamins) and minerals (calcium, iron, potassium, magnesium, etc.) can also be considered

Fig. 7 Fresh (**A**), solar dried (**B**), and open sun dried (**C**) bitter gourd samples (with permission from [96])

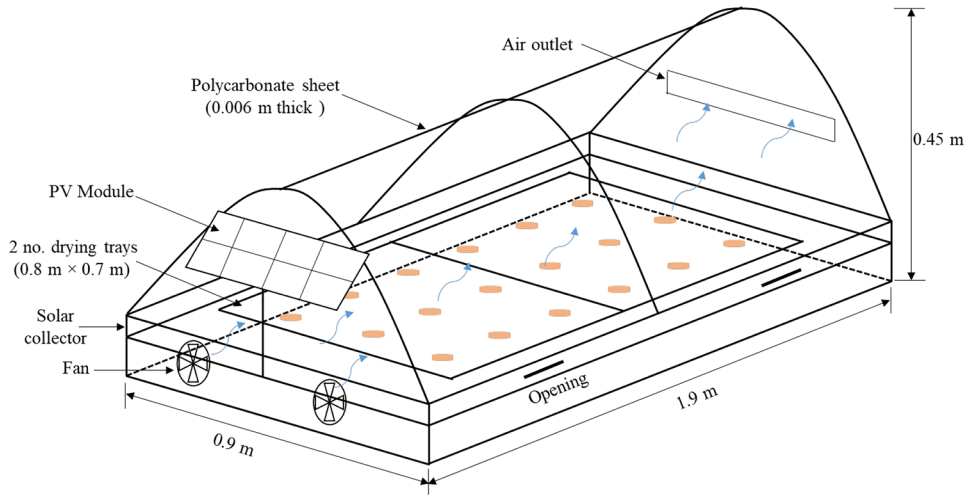


Fig. 8 Parabolic-shaped FCDSD [46]

one of the significant criteria for the solar dried product quality assessment. Various chemical tests can be conducted to evaluate the values of solar dried products so that the best solar drying system can be designed as per the product requirements [20, 83].

Quality of the dried product has also been analyzed and reported in some investigations on domestic solar dryers. Haque et al. [45] recommended the solar dried products over open sun dried products on the basis of color testing parameters. The value of C_{dev} for solar dried bitter gourd and okra were found to be 9.71 and 14.17, respectively. Vijayan et al. [96] presented the effect of solar drying on the quality of the bitter gourd. Higher color retention was observed in solar dried sample as compared to open sun dried sample (Fig. 7). Nabnean and Nimnuan [46] analyzed the quality of banana slices dried in a parabolic-shaped FCDSD (Fig. 8) on the basis of appearance, color, texture, flavor, taste, and overall acceptance. Solar dried samples were found superior in comparison to open sun dried bananas. Table 6 shows the nutrition facts of solar dried banana samples.

Gyawali et al. [97] used biochemical analysis and suggested that the retention of essential oil in solar dried ginger and turmeric samples was comparatively higher than that of under open sun drying. Oleoresin content was reported to be almost similar under both the drying conditions. Sharma et al. [49] observed that the quality of solar dried tomatoes was superior with an overall acceptability score of 4.2 as compared or open sun dried samples on the basis of five quality attributes, namely, color, flavor, mouthfeel, taste, and appearance. Dubey et al. [83] developed a domestic type forced convection mixed mode solar dryer (FCMMSD) and compared the effect of solar and open sun drying on the nutritional properties of grapes such as total sugars, proteins, and lipids. The percentage ash content of solar and open sun dried grapes was observed to be 2.71 and 1.95%, respectively. Solar dried samples were having higher percentages of macronutrients such as calcium, potassium, magnesium, and sodium and micronutrients iron, molybdenum, and zinc (Table 7). A brief summary of various product quality indicators used for solar dried products is presented in Table 8.

Modeling Techniques Used for Domestic Solar Dryers

The testing of solar drying systems can be challenging in both physical and financial ways. The development of any system needs money and testing requires labor. Modeling

Table 6 Nutrition facts of solar dried banana samples [46]

Parameter	Value (per 100 g)
Vitamin B1	30 µg
Calories	286.26 kcal
Carbohydrate	69.29 g
Ash content	1.95 g
Protein	1.69 g
Fat	0.26 g

Table 7 Nutrients retention in solar and open sun dried raisins [83]

Nutrients	Solar drying	Open sun drying
Ca (%)	0.065	0.056
K (%)	0.647	0.477
Mg (%)	0.380	0.030
Na (%)	0.996	0.494
Fe (ppm)	130	90
Mo (ppm)	20	10
Zn (ppm)	40	30

Table 8 Product quality analysis for domestic solar dryers

Study	Drying method	Product	Parameters studied	Remarks
Nabnean and Nimnuan [46]	FCDSD	Banana	• Color deviation • Sensory	The values of color deviation for solar and open sun dried samples were 30.06 and 44.15.
Nimnuan and Nabnean [64]	FCDSD	Galangal	• Color deviation	Color deviations for open sun dried and solar samples were 137.68 and 139.2, respectively.
Dubey et al. [83]	FCMMSD	Grapes	• Ash content • Protein amino acid and sugar • Lipid and fatty acid • Mineral composition • Bacterial load	Solar dried samples showed negligible bacterial growth as compared to open sun dried samples.
Kondareddy et al. [98]	FCMMSD	Black turmeric	• Color analysis • Antioxidant activity • Total phenolic content • Total flavonoid content	The dryer resulted in higher color deviation as compared to open sun drying.
Watson et al. [99]	Forced convection solar bed dryer	Red chili	• Microbial growth	The solar bed dryer reduced the drying time and hence the chances of microbial growth.

can be a very appropriate solution to overcome these problems. A number of modeling techniques have been used by researchers for analyzing various solar drying technologies. Some of the modeling techniques used for analyzing domestic solar dryers are given in Fig. 9.

Computational fluid dynamics (CFD) is one of the most widely used computer assisted process engineering tools to analyze and investigate solar drying systems. It can produce quantitative predictions about the behavior of the fluid flow inside the drying chamber based on the laws of conservation of mass, momentum, and energy given in Appendix C [100–103]. A flow chart of the process of CFD simulation has been shown in Fig. 10. Kam et al. [104] used COMSOL multiphysics software for CFD analysis to predict the velocity, temperature, and pressure distributions inside a natural convection greenhouse type solar dryer. Gyawali et al. [97] used CFD analysis in ANSYS Fluent software for the prediction of the temperature and behavior of airflow inside a forced convection mixed mode solar dryer. Andharia et al. [105] studied the effect of sensible and latent heat storage systems on the airflow and temperature distribution in a small-scale mixed mode solar dryer tested for Indian gooseberry. Dhalsamant [112] estimated the values of temperature and moisture ratio of potato cylinders in a mixed-mode solar dryer by using CFD analysis in COMSOL Multiphysics. The outcomes of CFD analysis were compared and found to be much accurate as compared to the results of an artificial neural network based model.

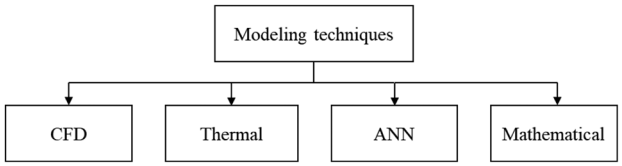


Fig. 9 Modeling techniques for analyzing domestic solar dryers

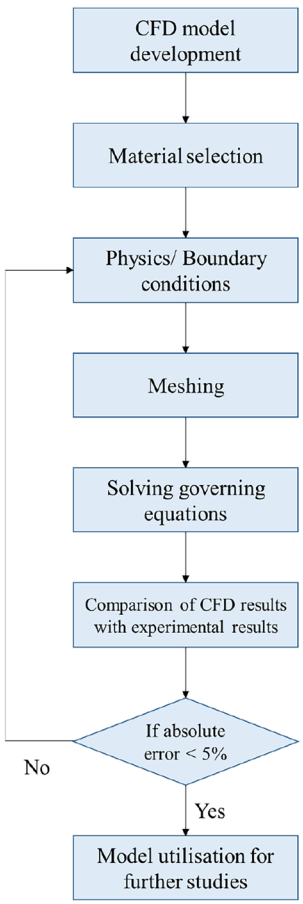


Fig. 10 Flow chart for CFD simulation

Thermal modeling has been used by many researchers for investigating various solar drying systems. It considers the inflow and outflow of energy through the solar dryer that follows the principle of energy balance. This technique has been widely employed to estimate various temperatures associated with a particular drying system such as product temperature and drying air temperature at inlet, outlet, and inside the drying chamber [106–108]. Spall and Sethi [109] have applied thermal modeling for the analysis of a cabinet type forced convection solar dryer having north wall reflector. The model was also validated with the experimental data using root mean square error given in Appendix C.

Artificial neural network (ANN) is a mathematical data processing technique inspired from the human neurons system. It consists a several nodes also called as neurons connected with each other in at least three layers. The first and last layers are called as input and output layers having neurons equal to the number of input and output variables, respectively. All the middle layer are called as hidden layers. The application procedure of ANN technique mainly consists of three steps, namely, training, validation, and testing. The accuracy of ANN model is generally tested by calculating the values of coefficient of correlation (R), mean absolute error (MAE), root mean square error (RMSE), and standard error (SE) given in Appendix C [102, 110–112]. Sadadou et al. [111] studied and recommended the application of ANN modeling for the prediction of the drying behavior (moisture ratio and drying rate) of fruits in the open sun and direct solar drying. Dhalsamant [112] used ANN modeling for the estimation of the temperature and moisture ratio of potato cylinders during solar drying. A multilayer feed-forward neural network was formed in neural network toolbox of MATLAB 2015b. Levenberg–Marquardt algorithm was used for the training of the program. Experiments were performed to get the data for training and testing sets. Three measurable parameters, namely, ambient temperature, solar insolation, and drying time, were considered as the three inputs nodes to the input layer of the model. The values of temperature and moisture ratio of the potato cylinders were obtained from the two nodes of the output layer. A total 197, 300, and 213 number of experimental data points were supplied for each input parameter for training of the model. For the testing of the model 14, 18, and 21, output data points were used. A representation of the developed ANN model is shown in Fig. 11. The ANN model having 9, 8, and 6 neurons with Levenberg–Marquardt algorithm and tansig transfer function was observed to be the best for predicting various parameters during the drying process for different diameters potato samples.

Mathematical modeling has been widely employed to estimate the drying characteristics of various drying commodities using different solar dryers. The moisture ratio of the drying commodity is calculated using experimental data and then fitted to different mathematical drying models. The accuracy of the drying models can also be tested by calculating the values of parameters given in Appendix C [44, 49, 57, 98, 113]. Daud and Simate [81] tested twelve mathematical thin layer drying models and purposed the suitability of Middilli model to estimate the moisture ratio of pineapple slices dried in an NCDSD. Table 9 shows various mathematical models recommended for different domestic solar dryers and drying commodities.

Developments in modeling techniques for domestic solar dryers have improved the span of the analysis and accuracy of the results. From the literature, it is observed that finite element–based computer software such as ANSYS and COMSOL are comparatively better than ANN models in terms of accuracy of the results. However, there is a scope of developing a dedicated model for a particular drying system considering real time conditions with least assumptions for the most accurate outcomes. Table 10 summarizes the modeling technique used for the analysis of domestic solar dryers.

Discussion, Implications, and Recommendations

Present investigation indicates that the performance of a solar drying system can be evaluated on the grounds of thermal, drying kinetic, environmental, economic analyses and quality aspects of dried products. It is observed that there is no study available in the literature showing a complete performance assessment of a domestic solar dryer. Thermal efficiency, moisture content, drying time, cost, and payback period are some of the most reported performance parameters. The values of thermal efficiency of different domestic solar dryers were observed to be varying in the range of 3.74–67.78% that can be further improved by using various design modifications (such as incorporation of solar collectors for higher energy collection, heat storage arrangements for continuous and fluctuation free operation, better insulations for reduced heat losses, and adequate ventilations for easy moisture removal and least energy losses) and process (including mass of the drying

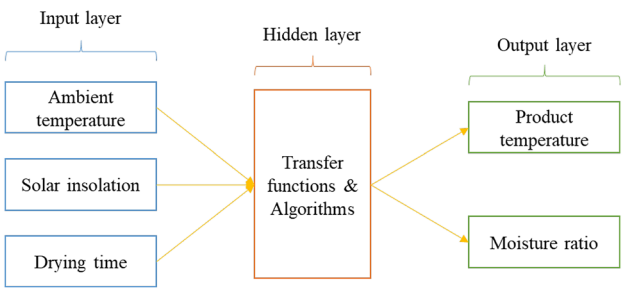


Fig. 11 ANN model [112]

Table 9 Mathematical models recommended for different domestic solar dryers

Sr No.	Study	Dryer	Drying commodity	Preferred/recommended model
1.	Saleh and Badran [44]	NCDSD	Jew's mallow	$MR = \exp(-kt)$
2.	Sharma et al. [49]	FCISD	Tomato slices	$MR = at^3 + bt^2 + ct + d$
3.	Borah et al. [57]	NCDSD	Turmeric	$MR = \exp(-kt^n)$
4.	Daud and Simate [81]	NCDSD	Pineapple slices	$MR = a \times \exp(-kt^n) + bt$
5.	Kondareddy et al. [98]	FCMMSD	Black turmeric	$MR = \exp\left[-(kt)^n\right]$

commodity, rate of flow of the drying air, and temperature distribution inside the drying chamber) optimization techniques such as CFD simulation using ANSYS or COMSOL. The drying time for different commodities under various solar dryers was found to be varying in the range of 5.5–240 h. Some of the dryers took quite higher drying time, which generally depends on the type of drying commodity, its initial and final moisture content values, and ambient conditions. Hybridization of solar dryers seems a solution to reduce the drying time in places where solar insolation is not so significant. Environmental impact assessment has been widely used by researchers for other solar technologies and has great significance in today's time when global climate change has been considered

Table 10 Modeling techniques for domestic solar dryers

Study	Drying method	Product	Modeling technique	Parameters studied	Remarks
Jain et al. [35]	NCDSD	-	CFD numerical simulation	• Static pressure • Temperature • Heat radiation flux • Absorbed visible solar flux	ANSYS Fluent 14.0 simulation software was used.
Saleh and Badran [44]	NCDSD	Jew's mallow	Mathematical modeling	• Drying kinetics	Exponential model was recommended.
Moghimi et al. [48]	FCISD	Tomato	CFD numerical simulation	• Temperature contour • Air streamlines	Fluent simulation was used.
Sharma et al. [49]	FCISD	Tomato slices	Mathematical modeling	• Drying kinetics	Prakash and Kumar model was recommended.
Borah et al. [57]	NCDSD	Turmeric	Mathematical modeling	• Drying kinetics	Page model was found as the most suitable.
Poonia et al. [63]	NCDSD and FCDSD	Ber (*Zizyphus mauritiana*)	Mathematical modeling	• Drying kinetics	Recommended logarithmic model.
Daud and Simate [81]	NCDSD	Pineapple slices	Mathematical modeling	• Drying kinetics	Middilli model showed highest congruence with experimental results.
Chavan et al. [82]	NCDSD	-	CFD modeling	• Air velocity	Employed ANSYS Fluent software.
Kondareddy et al. [98]	FCMMSD	Black turmeric	Mathematical modeling	• Drying kinetics	Modified page model was found more suitable.
Dhalsamant [112]	NCMMSD	Potato cylinders	Finite and ANN modeling	• Temperature • Moisture ratio	Finite modeling in COMSOL software was found better.
Terres et al. [114]	NCDSD	Lemon slices	CFD numerical simulation	• Temperature • Density • Air currents	ANSYS Fluent was used.
Chavan et al. [115]	NCDSD	Potato slices	CFD modeling	• Air velocity	Used ANSYS Fluent software.
Sandali et al. [116]	NCDSD	-	CFD modeling	• Temperature evolution	Employed Fluent CFD software.
Alonge and Obayopo [117]	FCDSD	Fish	CFD numerical simulation	• Temperature profile • Velocity contour	Used ANSYS Fluent simulation.

the biggest threat to life on the earth. The environmental impact assessment of domestic solar dryers has been reported very less in the literature. The initial cost and payback period for various domestic solar dryers were observed in the range of $3.61–500 and 0.25–3.26 years, respectively. The cost of a domestic solar dryer can be controlled by using local materials and the payback period can be reduced by operating at optimum conditions. The

Table 11 Performance assessment index (PAI) for a domestic solar dryer

Type of dryer:
Drying commodity:
Location:
Month:

Sr. No.	Parameter	Dryer 1	Dryer 2	Preferred value
Thermal indicators				
1	Pickup efficiency			Higher
2	Drying efficiency			Higher
3	Energy efficiency			Higher
4	Exergy efficiency			Higher
5	Exergetic indicators:			
	• Exergy loss			Lower
	• Waste exergy ratio			Lower
	• Improvement potential			Lower
	• Sustainability index			Higher
6	Specific energy consumption			Lower
7	Specific moisture extraction rate			Higher
8	Heat utilization factor			Higher
9	Coefficient of performance			Higher
10	Convective and evaporative heat transfer coefficients			Higher
Drying kinetics indicators				
11	Final moisture content			Should be in acceptable range
12	Moisture ratio			Lower
13	Drying rate			Higher
14	Effective moisture diffusivity			Higher
15	Activation energy			Lower
Environmental indicators				
16	Embodied energy			Lower
17	Energy payback time			Lower
18	CO_2 emissions			Lower
19	CO_2 mitigation potential			Higher
20	Carbon credit earned			Higher
Economic indicators				
21	Life cycle cost			Lower
22	Life cycle benefit			Higher
23	Payback period			Lower
Product quality indicators				
24	Sensory: color, texture, taste, and flavor			Should be as per the requirements
25	Ash content			Higher
26	Rehydration ratio			Higher
27	Shrinkage			Should be as per the requirements
28	Nutritional values			Higher

In addition to above mentioned indicators, the following points should be considered:

- Ease of operation and maintenance
- Safety and size
- Durability and portability

quality of the dried products under domestic solar dryers, which is one of the significant performance indicators, has also been reported rarely in the literature. The change in color of the dried product compared to the fresh product is the most reported product quality indicator. Solar drying has several effects on dried product quality, which, if controlled, can result in high quality products with higher market value. It would improve the earnings of the local farmers and their participation in the dried food market contributing to SDG-12 (i.e., responsible consumption and production). Higher quality products would also result in higher health benefits for consumers which is a contribution to SDG-3 (i.e., good health and well-being). Computational fluid dynamics (CFD) and thermal modeling have been widely used for analyzing and predicting temperature and air velocity inside the drying chamber of domestic solar dryers. Drying commodity is an integral part of the drying process, and hence, modeling and simulation of domestic solar dryers considering the drying commodity is a need for better observations. However, mathematical modeling has been widely used only for predicting the drying behavior of various drying commodities.

In the present era, saving of fossil fuels and environment during drying of various commodities using conventional methods is the major implication for renewable energy utilization. The domestic solar dryers could be highly beneficial along with fossil fuel conservation and pollution control by efficiently utilizing solar energy. This study would be of great significance to researchers in designing and developing a better domestic solar dryer and analyzing its performance. It would also motivate readers toward responsible consumption and production. Moreover, it would increase the awareness among household people around the globe who are keenly interested in combating hunger, climate change and pollution.

There are several criteria for the performance assessment of a domestic solar dryer. After careful and in-depth evaluation of the literature in the present study, a performance assessment index (PAI) for domestic solar dryers has been developed (Table 11). This PAI is recommended for the comparison of the performances of different domestic solar dryers which could be quite useful in standardization and certification of a solar dryer.

Summary

The performance of any system is one of the most responsible factors that lead to new developments. There have been significant advancements in the field of solar dryers in the last few decades and so in their performance evaluation techniques. This manuscript presents a comprehensive review of the methods used for the evaluation and analysis of solar drying systems with a particular emphasis on domestic solar dryers. For the assessment of domestic solar dryers, thermal, drying kinetic, and economic analyses have been performed by the researchers in various studies. However, the continuous environmental degradation is forcing researchers to reduce the carbon footprints (greenhouse gases) in new developments; thus, environmental impact assessment becomes quite essential. Quality of the dried product should always be considered in the development of a domestic solar dryer. Both the environmental and quality assessments are rarely reported in the literature based on domestic solar dryers. This work would be a one-stop solution for designing an efficient solar dryer and assessing its performance. A performance assessment index (PAI) for domestic solar dryers has been suggested in this regard.

Appendix

Appendix A

The value of savings (S) for the life of j number of years can be given as

$$S_j = S_d \times D(1+i)^{j-1} \tag{A1}$$

where S_d = savings per day and D = number of operating days/year.

The value of S_d can be given as

$$S_d = \frac{S_b}{D_b} \tag{A2}$$

where S_b = saving per batch and D_b = days/batch.

The value of S_b can be given as

$$S_b = S_{kg} \times M_d \tag{A3}$$

where S_{kg} = savings/kg and M_d = mass of dried product/batch.

The value of S_{kg} can be given as follows [118]:

$$S_{kg} = C_b - C_{sd} \tag{A4}$$

where C_b = market cost of the branded dried product and C_{sd} = cost of one kg of dried product.

The value of C_{sd} can be given as

$$C_{sd} = C_{dp} + C_s \tag{A5}$$

where C_{dp} = cost of fresh product per kg of dried product and C_s = drying cost/kg of dried product, whose values can be calculated as

$$C_{dp} = C_{fp}\frac{M_i}{M_d} \tag{A6}$$

$$C_s = \frac{C_a}{M_y} \tag{A7}$$

where C_{fp} = cost of fresh product, M_i = initial mass of product/batch, M_d = mass of dried product/batch, C_a = annualised cost, and M_y = mass of product dried/year.

The values of C_a and M_y can be given as follows [118]:

$$C_a = C_{ac} + C_m - S_a + C_f \tag{A8}$$

$$M_y = \frac{M_d \times D}{D_b} \tag{A9}$$

where C_{ac} = annualised capital cost, C_m = annual maintenance cost (10% of C_{cc}), S_a = annualised salvage value, and C_f = operational cost in case of fan.

The values of C_{ac}, S_a and C_f can be given as follows [119]:

$$C_{ac} = C_{cc} \times F_c \tag{A10}$$

$$S_a = S_v \times F_s \tag{A11}$$

$$C_f = N_f \times P_f \times C_{eu} \tag{A12}$$

where F_c = capital recovery factor, S_v = salvage value (3% of C_{cc}), F_s = salvage fund factor, N_f = number of annual operating hours of fan, P_f = rated power consumed by the fan during operation, and C_{eu} = cost of electricity/unit.

The values of F_c and F_s can be given as follows [120]:

$$F_c = \frac{i(1+i)^j}{(1+i)^j - 1} \tag{A13}$$

$$F_s = \frac{i}{(1+i)^j - 1} \tag{A14}$$

Appendix B

Continuity equation

$$\frac{\partial \rho}{\partial t} + \nabla(\rho \vec{v}) = 0 \tag{B1}$$

Momentum equation

$$\frac{\partial}{\partial t}\left(\rho \vec{v}\right) + \nabla\left(\rho \vec{v} \otimes \vec{v}\right) = -\nabla \mathrm{P} + \nabla \bar{\bar{\tau}} + \rho \vec{g} + S_m \tag{B2}$$

Energy equation

$$\frac{\partial}{\partial t}(\rho E) + \nabla\left\{\vec{v}(\rho E + P)\right\} = \nabla\left(-\vec{q} + \bar{\bar{\tau}}.\vec{v}\right) + S_n \tag{B3}$$

where $\nabla = \frac{\partial}{\partial x}\vec{i} + \frac{\partial}{\partial y}\vec{j} + \frac{\partial}{\partial z}\vec{k}$, ρ = density of the fluid, $\vec{v}$ = velocity of fluid, P = pressure, $\bar{\bar{\tau}}$ = stress tensor, S_m = momentum source term, E = total energy, $\vec{q}$ = flux vector (positive inside), and S_n = energy source term.

Appendix C

$$\mathrm{R} = \frac{N\sum_{i=1}^{N} \mathrm{MR}_{exp,i}\mathrm{MR}_{pre,i} - \left(\sum_{i=1}^{N} \mathrm{MR}_{exp,i}\right)\left(\sum_{i=1}^{N} \mathrm{MR}_{pre,i}\right)}{\sqrt{N\sum_{i=1}^{N} \mathrm{MR}_{exp,i}^{\ 2} - \left(\sum_{i=1}^{N} \mathrm{MR}_{exp,i}\right)^2}\sqrt{N\sum_{i=1}^{N} \mathrm{MR}_{pre,i}^{\ 2} - \left(\sum_{i=1}^{N} \mathrm{MR}_{pre,i}\right)^2}} \tag{C1}$$

$$\mathrm{MAE} = \frac{\sum_{i=1}^{N}(\mathrm{MR}_{exp,i} - \mathrm{MR}_{pre,i})}{N} \tag{C2}$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}\left(\mathrm{MR}_{exp,i} - \mathrm{MR}_{pre,i}\right)^2}{N}} \tag{C3}$$

$$\mathrm{SE} = \frac{\sqrt{\sum_{i=1}^{N}\left(\mathrm{MR}_{exp,i} - \mathrm{MR}_{pre,i}\right)^2}}{N-1} \tag{C4}$$

where MR_{exp} = experimental value of MR and MR_{pre} = predicted value of MR.

Acknowledgements The authors are highly grateful to Guru Jambheshwar University of Science & Technology, Hisar, India, and Centre for Energy and Environment, Delhi Technological University, Delhi, India, for providing the necessary facilities to compile this work.

Declarations

Competing Interests The authors declare no competing interests.

References

1. Janjai S, Bala BK (2012) Solar drying technology. Food Eng Rev 4:16–54. https://doi.org/10.1007/s12393-011-9044-6
2. Tiris C, Tiris M, Dincer I (1996) Experiments on a new small-scale solar dryer. Appl Therm Eng 16:183–187
3. Kamarulzaman A, Hasanuzzaman M, Rahim NA (2021) Global advancement of solar drying technologies and its future prospects: a review. Sol Energy 221:559–582. https://doi.org/10.1016/j.solener.2021.04.056
4. Jha A, Tripathy PP (2021) Optimization of process parameters and numerical modeling of heat and mass transfer during simulated solar drying of paddy. Comput Electron Agric 187:106215. https://doi.org/10.1016/j.compag.2021.106215
5. UNEP (2021) Food Waste Index Report 2021. Nairobi
6. Kumar M, Sansaniwal SK, Khatak P (2016) Progress in solar dryers for drying various commodities. Renew Sustain Energy Rev 55:346–360
7. Ugwuoke IC, Ikechukwu IB, Ifianyi OE (2019) Design and development of a mixed-mode domestic solar dryer. Int J Eng Manuf 9:55–65. https://doi.org/10.5815/ijem.2019.03.05
8. Manchanda H, Kumar M, Gupta M, Shimpy (2019) No load testing of single slope solar distillation- cum-drying unit: an experimental study. Trends and Advances in Mechanical Engineering. Faridabad, pp 17–22
9. Visavale GL (2012) Principles, classification, and selection of dryers. Handbook of Industrial Drying. CRC, pp 29–58
10. Shimpy, Manchanda H, Kumar M, Gupta M (2019) Recent developments and comprehensive review of greenhouse dryers. Trends and Advances in Mechanical Engineering. Faridabad, pp 23–31

11. Shimpy KM, Kumar A (2022) Designs, performance and economic feasibility of domestic solar dryers. Food Eng Rev. https://doi.org/10.1007/s12393-022-09323-1
12. Fernandes L, Fernandes JR, Tavares PB (2022) Design of a Friendly Solar Food Dryer for Domestic Over-Production. Solar 2:495–508. https://doi.org/10.3390/solar2040029
13. Ampratwum DB, Dorvlo ASS (1998) Evaluation of a solar cabinet dryer as an air-heating system. Appl Energy 59:63–71. https://doi.org/10.1016/S0306-2619(97)00043-3
14. Mustayen AGMB, Mekhilef S, Saidur R (2014) Performance study of different solar dryers: a review. Renew Sustain Energy Rev 34:463–470. https://doi.org/10.1016/j.rser.2014.03.020
15. Malik A, Shimpy, Kumar M (2023) Advancements in ginger drying technologies. J Stored Prod Res 100:102058. https://doi.org/10.1016/j.jspr.2022.102058
16. Kumar M (2013) Experimental study on natural convection greenhouse drying of papad. J Energy South Africa 24:37–43. https://doi.org/10.17159/2413-3051/2013/v24i4a3144
17. Singh P, Gaur MK (2020) Review on development, recent advancement and applications of various types of solar dryers. Energy Sources A Recover Util Environ Eff. https://doi.org/10.1080/15567036.2020.1806951
18. 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
19. Lingayat AB, Chandramohan VP, Raju VRK, Meda V (2020) A review on indirect type solar dryers for agricultural crops – dryer setup, its performance, energy storage and important highlights. Appl Energy 258:114005. https://doi.org/10.1016/j.apenergy.2019.114005
20. Leon MA, Kumar S, Bhattacharya SC (2002) A comprehensive procedure for performance evaluation of solar food dryers. Renew Sustain Energy Rev 6:367–393
21. Dejchanchaiwong R, Arkasuwan A, Kumar A, Tekasakul P (2016) Mathematical modeling and performance investigation of mixed-mode and indirect solar dryers for natural rubber sheet drying. Energy Sustain Dev 34:44–53. https://doi.org/10.1016/j.esd.2016.07.003
22. Ekka JP, Muthukumar P, Bala K et al (2021) Performance studies on mixed-mode forced convection solar cabinet dryer under different air mass flow rates for drying of cluster fig. Sol Energy 229:39–51. https://doi.org/10.1016/j.solener.2021.06.086
23. Patel J, Andharia J, Georgiev A et al (2020) A review of phase change material based thermal energy accumulators in small-scale solar thermal dryers. Bulg Chem Commun 52:53–64
24. Kumar M, Sahdev RK, Tiwari S et al (2021) Enviro-economical feasibility of groundnut drying under greenhouse and indoor forced convection hot air dryers. J Stored Prod Res 93:1–10. https://doi.org/10.1016/j.jspr.2021.101848
25. Kumar M, Sahdev RK, Tiwari S et al (2019) Experimental free convection thin layer groundnut greenhouse drying. Agric Eng Int CIGR J 21:203–211
26. Shimpy, Kumar M, Kumar A et al (2022) 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
27. Sahdev RK, Kumar M, Dhingra AK (2018) Forced convection greenhouse groundnut drying: an experimental study. Heat Transf Res 49:309–325. https://doi.org/10.1615/HeatTransRes.2018018321
28. 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
29. Mugi VR, Chandramohan VP (2021) Energy and exergy analysis of forced and natural convection indirect solar dryers: estimation of exergy inflow, outflow, losses, exergy efficiencies and sustainability indicators from drying experiments. J Clean Prod 282:124421. https://doi.org/10.1016/j.jclepro.2020.124421
30. Kesavan S, Arjunan TV, Vijayan S (2019) Thermodynamic analysis of a triple-pass solar dryer for drying potato slices. J Therm Anal Calorim 136:159–171. https://doi.org/10.1007/s10973-018-7747-0
31. Selimefendigil F, Şirin C, Öztop HF (2022) Improving the performance of an active greenhouse dryer by integrating a solar absorber north wall coated with graphene nanoplatelet-embedded black paint. Sol Energy 231:140–148. https://doi.org/10.1016/J.SOLENER.2021.10.082
32. Gupta A, Das B, Biswas A, Mondol JD (2022) Sustainability and 4E analysis of novel solar photovoltaic-thermal solar dryer under forced and natural convection drying. Renew Energy 188:1008–1021. https://doi.org/10.1016/j.renene.2022.02.090
33. Chauhan PS, Kumar A, Singh P, Kumar A (2016) Performance analysis of greenhouse dryer by using insulated north-wall under natural convection mode. Energy Rep 2:107–116. https://doi.org/10.1016/j.egyr.2016.05.004
34. Vijayan S, Arjunan TV, Kumar A (2017) Fundamental concepts of drying. Green Energy and Technology. CRC, pp 3–38
35. 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
36. Chauhan PS, Kumar A, Nuntadusit C (2018) Heat transfer analysis of PV integrated modified greenhouse dryer. Renew Energy 121:53–65. https://doi.org/10.1016/j.renene.2018.01.017
37. Cengel Y, Ghajar A (2015) Heat and mass transfer fundamentals & applications. Mc Graw Hill Education, India
38. Tiwari GN, Goyal RK (1998) Greenhouse technology: fundamentals, design, modelling and applications. Narosa Publishing House
39. Goyal RK, Tiwari GN (1998) Heat and mass transfer relations for crop drying. Dry Technol 16:1741–1754. https://doi.org/10.1080/07373939808917490
40. Malik MAS, Tiwari GN, Kumar A, Sodha MS (1982) Solar distillation: a practical study of a wide range of stills and their optimum design, construction, and performance. Pergamon Press Ltd
41. 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
42. Sharma VK, Sharma S, Ray RA, Garg HP (1986) Design and performance studies of a solar dryer suitable for rural applications. Energy Convers Manag 26:111–119. https://doi.org/10.1016/0196-8904(86)90040-3
43. 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
44. Saleh A, Badran I (2009) Modeling and experimental studies on a domestic solar dryer. Renew Energy 34:2239–2245. https://doi.org/10.1016/j.renene.2009.03.001
45. 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
46. Nabnean S, Nimnuan P (2020) Experimental performance of direct forced convection household solar dryer for drying banana. Case Stud Therm Eng 22:100787. https://doi.org/10.1016/j.csite.2020.100787
47. Tiwari S (2020) ANN and mathematical modelling for moisture evaporation with thermal modelling of bitter gourd flakes drying in SPVT solar dryer. Heat Mass Transf 56:2831–2845. https://doi.org/10.1007/s00231-020-02886-x
48. Moghimi P, Rahimzadeh H, Ahmadpour A (2021) Experimental and numerical optimal design of a household solar fruit and

vegetable dryer. Sol Energy 214:575–587. https://doi.org/10.1016/j.solener.2020.12.023
49. Sharma M, Atheaya D, Kumar A (2022) Exergy, drying kinetics, and performance assessment of Solanum lycopersicum (tomatoes) drying in an indirect type domestic hybrid solar dryer (ITDHSD) system. J Food Process Preserv. https://doi.org/10.1111/jfpp.16988
50. Hallak H, Hilal J, Hilal F, Rahhal R (1996) The staircase solar dryer: design and characteristics. Renew Energy 7:177–183. https://doi.org/10.1016/0960-1481(95)00127-1
51. Thanvi KP, Pande PC (1987) Development of a low-cost solar agricultural dryer for arid regions of India. Energy Agric 6:35–40. https://doi.org/10.1016/0167-5826(87)90020-9
52. Bolaji B (2005) Development and performance evaluation of a box-type absorber solar air collector for crop drying. J Food Technol 3:595–600
53. Ezekoye BA, Enebe OM (2006) Development and performance evaluation of modified integrated passive solar grain dryer. Pacific J Sci Technol 7:185–190
54. Eke BA (2013) Development of small scale direct mode natural convection solar dryer for tomato, okra and carrot. Int J Eng Adv Technol 3:199–204
55. Eke BA (2014) Investigation of low cost solar collector for drying vegetables in rural areas. Agric Eng Int CIGR J 16:118–125
56. Navale SR, Harpale VM, Mohite KC (2015) Comparative study of open sun and cabinet solar drying for fenugreek leaves. Int J Renew Energy Technol Res 4:1–9
57. Borah A, Hazarika K, Khayer SM (2015) Drying kinetics of whole and sliced turmeric rhizomes (Curcuma longa L.) in a solar conduction dryer. Inf Process Agric 2:85–92. https://doi.org/10.1016/j.inpa.2015.06.002
58. Subedi TR, Bhattarai RN (2017) Experimental performance analysis of solar conduction dryer (SCD) for ginger drying. IOE Graduate Conference, pp 597–601
59. 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
60. Modi VM, Desai NN, Gora A (2017) Design and development of low cost solar dryer. AGRES An Int e J 6:329–336
61. Chaudhari RH, Gora A, Modi VM, Chaudhari H (2018) Economic analysis of hybrid solar dryer for ginger drying. Int J Curr Microbiol Appl Sci 7:2725–2731. https://doi.org/10.20546/ijcmas.2018.711.312
62. 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
63. Poonia S, Singh AK, Jain D (2018) Mathematical modelling and techno-economic evaluation of hybrid photovoltaic-thermal forced convection solar drying of Indian jujube (Zizyphus mauritiana). J Agric Eng 55:74–88
64. Nimnuan P, Nabnean S (2020) Solar drying of galangal slices (alpinia galangal (linn.) swartz.) using household solar dryer. Suranaree J Sci Technol 27:1–8
65. Safri NAM, Zainuddin Z, Azmi MSM et al (2020) Temperature performance of a portable solar greenhouse dryer with various collector design. Sains Malaysiana 49:2539–2545. https://doi.org/10.17576/jsm-2020-4910-19
66. Obajemihi OI, Olaoye JO, Cheng JH et al (2021) Optimization of process conditions for moisture ratio and effective moisture diffusivity of tomato during convective hot-air drying using response surface methodology. J Food Process Preserv 45:1–14. https://doi.org/10.1111/jfpp.15287
67. 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
68. Singh M, Sethi VP (2018) On the design, modelling and analysis of multi-shelf inclined solar cooker-cum-dryer. Sol Energy 162:620–636. https://doi.org/10.1016/j.solener.2018.01.045
69. Yağcioğlu A, Demir V, Günhan T (2007) Effective moisture diffusivity estimation from drying data. Tarım Makinaları Bilim Derg 3:249–256
70. Falade KO, Abbo ES (2007) Air-drying and rehydration characteristics of date palm (Phoenix dactylifera L.) fruits. J Food Eng 79:724–730. https://doi.org/10.1016/j.jfoodeng.2006.01.081
71. Djebli A, Hanini S, Badaoui O, Boumahdi M (2019) A new approach to the thermodynamics study of drying tomatoes in mixed solar dryer. Sol Energy 193:164–174. https://doi.org/10.1016/j.solener.2019.09.057
72. Gilago MC, Chandramohan VP (2022) Performance parameters evaluation and comparison of passive and active indirect type solar dryers supported by phase change material during drying ivy gourd. Energy 252:123998. https://doi.org/10.1016/j.energy.2022.123998
73. Fiorentini C, Demarchi SM, Quintero Ruiz NA et al (2015) Arrhenius activation energy for water diffusion during drying of tomato leathers: the concept of characteristic product temperature. Biosyst Eng 132:39–46. https://doi.org/10.1016/j.biosystemseng.2015.02.004
74. Getahun E, Delele MA, Gabbiye N et al (2021) Importance of integrated CFD and product quality modeling of solar dryers for fruits and vegetables: a review. Sol Energy 220:88–110. https://doi.org/10.1016/j.solener.2021.03.049
75. Alonge AF, Adeboye OA (2012) Drying rates of some fruits and vegetables with passive solar dryers. Int J Agric Biol Eng 5:83–90. https://doi.org/10.3965/j.ijabe.20120504.00
76. 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
77. 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
78. Rawat BS, Pant PC, Joshi GC (2009) Energetics study of a natural convection solar crop dryer. Int J Ambient Energy 30:193–198. https://doi.org/10.1080/01430750.2009.9675096
79. Alonge AF, Omoniwa AO (2012) Development and modification of a direct passive solar dryer. NABEC-CSBE/SCGAB 2012 Joint Meeting and Technical Conference Northeast Agricultural & Biological Engineering Conference Canadian Society for Bioengineering, Orillia, Ontario, pp 1–10
80. Seveda MS, Jhajharia D (2012) Design and performance evaluation of solar dryer for drying of large cardamom (Amomum subulatum). J Renew Sustain Energy 4:063129. https://doi.org/10.1063/1.4769199
81. Daud LEI, Simate IN (2017) Drying kinetics of sliced pineapples in a solar conduction dryer. Energy Environ Res 7:14. https://doi.org/10.5539/eer.v7n2p14
82. Chavan A, Vitankar V, Thorat B (2020) CFD modeling and experimental study of solar conduction dryer. Dry Technol 39:1087–1100. https://doi.org/10.1080/07373937.2020.1846051
83. Dubey S, Bhayani K, Bhatt C et al (2022) Design and development of domestic solar dryer with comparative analysis of nutritional aspect of dried raisins. J Food Eng Technol 11:13–21. https://doi.org/10.32732/jfet.2022.11.1.13
84. Ayyappan S (2018) Performance and CO 2 mitigation analysis of a solar greenhouse dryer for coconut drying. Energy Environ 29:1482–1494. https://doi.org/10.1177/0958305X18781891
85. 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
86. Shimpy KM, Sahdev RK et al (2022) 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

87. Rawat BS, Rawat PN, Pant PC, Joshi GC (2014) Evaluation of energetics and CO2 emission mitigation potential of natural convection solar dryer for amla. AU J Technol 18:75–81
88. Kumar M, Shimpy Sahdev RK et al (2023) Natural convective greenhouse vermicelli drying: thermo-environ-econo-kinetic analyses. Sustain Energy Technol Assess 55:103002. https://doi.org/10.1016/j.seta.2022.103002
89. Mustapha MK, Salako AF, Ademola SK, Adefila IA (2014) Qualitative performance and economic analysis of low cost solar fish driers in Sub-Saharan Africa. J Fish 2:64–69. https://doi.org/10.17017/jfish.v2i1.2014.23
90. Sandali M, Boubekri A, Mennouche D (2020) Thermal and economical study of a direct solar dryer with integration of different techniques of heat supply. Springer Proceedings in Energy, pp 585–595
91. Tefera A, Endalew W, Fikiru B (2013) Evaluation and demonstration of direct solar potato dryer. Livest Res Rural Dev 25:1–7
92. Poonia S, Singh AK, Santra P, Mishra D (2018) Design, development and performance evolution of a low-cost solar dryer. In: Chandra L, Dixit A (eds) Concentrated Solar Thermal Energy Technologies, Springer Proceedings in Energy. Springer Nature, Singapore, pp 219–223
93. Pervin S, Islam M, Islam M (1970) Study on rehydration characteristics of dried lablab bean (Lablab purpureus) seeds. J Agric Rural Dev 6:157–163. https://doi.org/10.3329/jard.v6i1.1673
94. Seerangurayar T, Al-Ismaili AM, Janitha Jeewantha LH, Al-Nabhani A (2019) Experimental investigation of shrinkage and microstructural properties of date fruits at three solar drying methods. Sol Energy 180:445–455. https://doi.org/10.1016/j.solener.2019.01.047
95. Kaveh M, Abbaspour-Gilandeh Y, Fatemi H, Chen G (2021) Impact of different drying methods on the drying time, energy, and quality of green peas. J Food Process Preserv 45:e15503. https://doi.org/10.1111/jfpp.15503
96. Vijayan S, Arjunan TV, Kumar A (2016) Mathematical modeling and performance analysis of thin layer drying of bitter gourd in sensible storage based indirect solar dryer. Innov Food Sci Emerg Technol 36:59–67. https://doi.org/10.1016/j.ifset.2016.05.014
97. Gyawali M, Acharya A, Adhikari T et al (2021) A mixed-mode ginger and turmeric solar dryer: design, simulation, biochemical and performance analysis. BIBECHANA 19:40–60. https://doi.org/10.3126/bibechana.v19i1-2.46386
98. Kondareddy R, Sivakumaran N, Nayak PK (2019) Drying kinetics of black turmeric (Curcuma caesia) with optimal controller assisted low cost solar dryer. Food Res 3:373–379. https://doi.org/10.26656/fr.2017.3(4)0.139
99. Watson AG, Aleckovic S, Nallamothu R (2021) A novel and improved solar drying system appropriate for smallholder farmers. Dry Technol. https://doi.org/10.1080/07373937.2021.1931295
100. Kumar M (2019) Fluid mechanics and hydraulic machines, 1st edn. Pearson
101. Raj AK, Srinivas M, Jayaraj S (2020) Transient CFD analysis of macro-encapsulated latent heat thermal energy storage containers incorporated within solar air heater. Int J Heat Mass Transf 156:119896. https://doi.org/10.1016/j.ijheatmasstransfer.2020.119896
102. Prakash O, Laguri V, Pandey A et al (2016) Review on various modelling techniques for the solar dryers. Renew Sustain Energy Rev 62:396–417. https://doi.org/10.1016/j.rser.2016.04.028
103. Dejchanchaiwong R, Tirawanichakul Y, Tirawanichakul S et al (2017) Conjugate heat and mass transfer modeling of a new rubber smoking room and experimental validation. Appl Therm Eng 112:761–770. https://doi.org/10.1016/j.applthermaleng.2016.10.108
104. Kam S, Kombasseré AME, Kaboré B et al (2017) Numerical simulation of greenhouse solar dryer in natural convection. Int J Dev Res 07:2–6
105. Andharia JK, Markam B, Dzhonova D, Maiti S (2022) A comparative performance analysis of sensible and latent heat based storage in a small-scale solar thermal dryer. J Energy Storage 45:103764. https://doi.org/10.1016/j.est.2021.103764
106. 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
107. Mahapatra A, Tripathy PP (2019) Thermal performance analysis of natural convection solar dryers under no load condition: experimental investigation and numerical simulation. Int J Green Energy 16:1448–1464. https://doi.org/10.1080/15435075.2019.1671417
108. Sethi VP, Dhiman M (2020) Design, space optimization and modelling of solar-cum-biomass hybrid greenhouse crop dryer using flue gas heat transfer pipe network. Sol Energy 206:120–135. https://doi.org/10.1016/j.solener.2020.06.006
109. Spall S, Sethi VP (2020) Design, modeling and analysis of efficient multi-rack tray solar cabinet dryer coupled with north wall reflector. Sol Energy 211:908–919. https://doi.org/10.1016/j.solener.2020.10.012
110. Rezrazi A, Hanini S, Laidi M (2016) An optimisation methodology of artificial neural network models for predicting solar radiation: a case study. Theor Appl Climatol 123:769–783. https://doi.org/10.1007/s00704-015-1398-x
111. Sadadou A, Hanini S, Laidi M, Rezrazi A (2021) ANN-based approach to model MC/DR of some fruits under solar drying. Kem u Ind 70:233–242. https://doi.org/10.15255/kui.2020.050
112. Dhalsamant K (2021) Development, validation, and comparison of FE modeling and ANN model for mixed-mode solar drying of potato cylinders. J Food Sci 86:3384–3402. https://doi.org/10.1111/1750-3841.15847
113. Milczarek RR, Alleyne FS (2017) Mathematical and computational modeling simulation of solar drying systems. Green Energy and Technology. Springer, Singapore, pp 357–379. https://doi.org/10.1007/978-981-10-3833-4_11
114. Terres H, Chavez S, Lopez R et al (2015) Study of the lemon drying process using a solar dryer. ASME 2015 9th International Conference on Energy Sustainability, San Diego, California, pp 1–6
115. Chavan A, Sikarwar A, Tidke V, Thorat B (2018) Augmenting — natural convection and conduction based solar dryer. 21st International Drying Symposium, València, Spain, pp 1357–1364
116. Sandali M, Boubekri A, Mennouche D (2018) Thermal behavior modeling of a cabinet direct solar dryer as influenced by sensible heat storage in a fractured porous medium. AIP Conf Proc 1968:020014. https://doi.org/10.1063/1.5039173
117. Alonge OI, Obayopo SO (2019) Computational fluid dynamics and experimental analysis of direct solar dryer for fish. Agric Eng Int CIGR J 21:108–117
118. Singh P, Gaur MK (2021) Environmental and economic analysis of novel hybrid active greenhouse solar dryer with evacuated tube solar collector. Sustain Energy Technol Assess 47:101428. https://doi.org/10.1016/j.seta.2021.101428
119. Sansaniwal SK, Kumar M, Sahdev RK et al (2022) Toward natural convection solar drying of date palm fruits (Phoenix dactylifera L.): An experimental study. Environ Prog Sustain Energy. https://doi.org/10.1002/ep.13862
120. Sreekumar A (2010) Techno-economic analysis of a roof-integrated solar air heating system for drying fruit and vegetables. Energy Convers Manag 51:2230–2238. https://doi.org/10.1016/j.enconman.2010.03.017

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.

biosensors

MDPI

Review

Recent Prospects of Carbonaceous Nanomaterials-Based Laccase Biosensor for Electrochemical Detection of Phenolic Compounds

Sakshi Verma [1], Deeksha Thakur [1], Chandra Mouli Pandey [2,*] and Devendra Kumar [1,*]

[1] Department of Applied Chemistry, Delhi Technological University, Delhi 110042, India
[2] Department of Chemistry, Faculty of Science, SGT University, Gurugram 122505, India
* Correspondence: cmp.npl@gmail.com (C.M.P.); dkumar@dce.ac.in (D.K.)

Abstract: Phenolic compounds (PhCs) are ubiquitously distributed phytochemicals found in many plants, body fluids, food items, medicines, pesticides, dyes, etc. Many PhCs are priority pollutants that are highly toxic, teratogenic, and carcinogenic. Some of these are present in body fluids and affect metabolism, while others possess numerous bioactive properties such as retaining antioxidant and antimicrobial activity in plants and food products. Therefore, there is an urgency for developing an effective, rapid, sensitive, and reliable tool for the analysis of these PhCs to address their environmental and health concern. In this context, carbonaceous nanomaterials have emerged as a promising material for the fabrication of electrochemical biosensors as they provide remarkable characteristics such as lightweight, high surface: volume, excellent conductivity, extraordinary tensile strength, and biocompatibility. This review outlines the current status of the applications of carbonaceous nanomaterials (CNTs, graphene, etc.) based enzymatic electrochemical biosensors for the detection of PhCs. Efforts have also been made to discuss the mechanism of action of the laccase enzyme for the detection of PhCs. The limitations, advanced emerging carbon-based material, current state of artificial intelligence in PhCs detection, and future scopes have also been summarized.

Keywords: laccase; biosensor; electrochemical; carbonaceous material; phenolic compounds

check for updates

Citation: Verma, S.; Thakur, D.; Pandey, C.M.; Kumar, D. Recent Prospects of Carbonaceous Nanomaterials-Based Laccase Biosensor for Electrochemical Detection of Phenolic Compounds. *Biosensors* **2023**, *13*, 305. https://doi.org/10.3390/bios13030305

Received: 14 January 2023
Revised: 10 February 2023
Accepted: 16 February 2023
Published: 22 February 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Phenolic compounds (PhCs) are extensively used in the production of dyes, chemical analysis, disinfectants, synthesis of artificial resins, petrochemicals, pharmaceuticals, textiles, paints, and medical industries [1,2]. These PhCs are present in body fluids, and certain PhCs have potential antioxidant properties [3]. A diverse variety of PhCs display varied impacts on biotic and abiotic components.

PhCs used for industrial purposes and a few released from the body (e.g., estrogen) are ubiquitous pollutants that are toxic even at low concentrations. The level of PhCs toxicity is in the range of 9–25 mg L^{-1}, while the water purity standards set by the US EPA allow a threshold value of approximately 1 ppb of phenol in the case of surface water [1]. The European Commission has prescribed a 0.001 mg L^{-1} concentration limit of PhCs in natural water, whereas the Central Pollution Control Board of India has restricted 1.0 mg L^{-1} of a phenolic compound as the benchmark for discharge of water to inland surfaces under the Environment (Protection) Rules, 1986. When discharged into the environment without prior treatment, these compounds lead to severe complications and long-term health issues for humans, animals, and marine systems [4]. Drinking water containing PhCs may cause diseases such as indigestion, skin burns, muscle tremor, liver problems and kidney damage in humans, and death of fish in aquatic regions [2]. In contrast, some other PhCs present in food and body fluid are responsible for preventing disease by balancing blood pressure and stress, reducing cardiovascular diseases, and providing healthy antioxidants [5,6]. The effects and sources of widely used PhCs have been summarized in Table 1.

Various chromatographic and spectroscopic techniques have been used to detect and quantify PhCs. These laboratory techniques require tedious process of sample pretreatment

from water bodies, which are time-consuming and costly [7]. For rapid and efficient detection, it is necessary to develop portable and more sensitive on-site tools comprising resourceful materials to reduce the devastating effects caused by PhCs. In this context, biosensors can be considered the most reliable techniques for on-site and specific detection of PhCs [8] (Figure 1a).

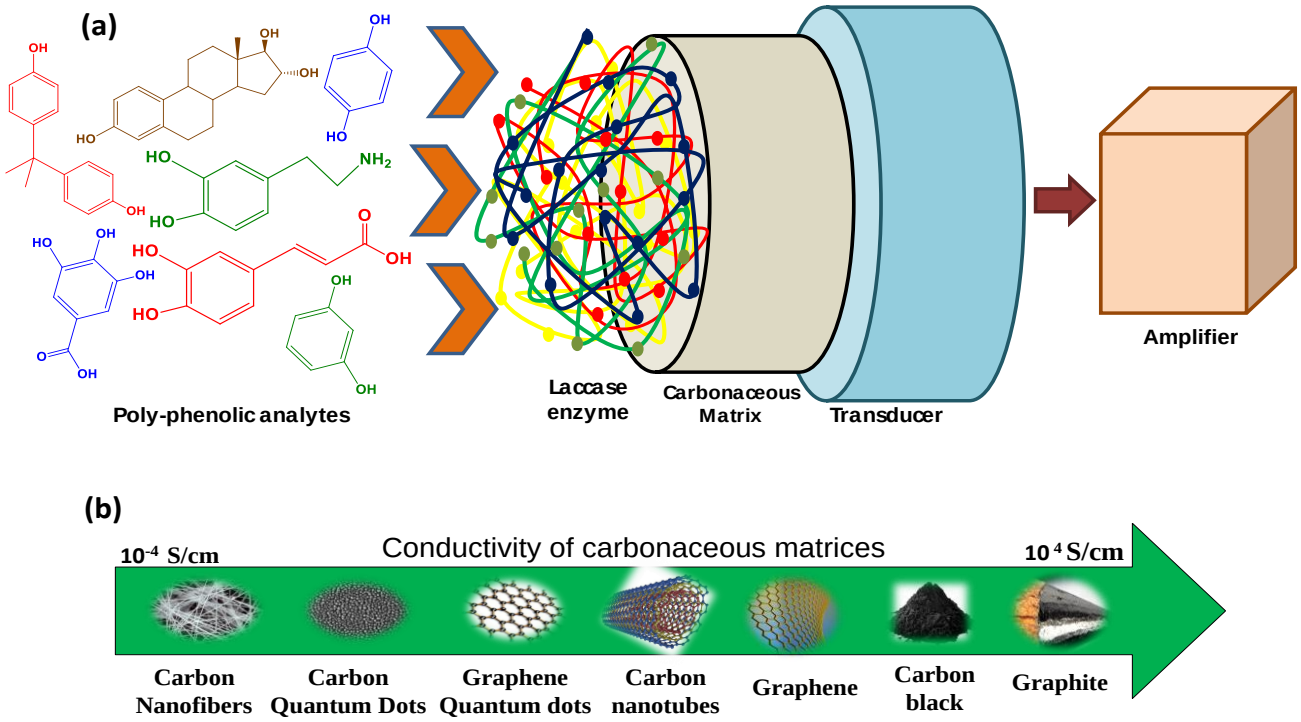


Figure 1. Scheme representing (**a**) a biosensor for PhCs detection using a carbon-based matrix to support the laccase enzyme and (**b**) the comparative conductivity of carbon−based materials.

Table 1. Harmful effects and sources of widely used PhCs on humans and animals.

S.No.	PhCs	Effects	Sources	Ref.
1.	Catechol	• Protein destruction in the body • Damage to DNA	• Coal conversion process • Crude wood tar • Coal tar production	[9]
2.	Bisphenol A	• Endocrine troublesome effects • Interruption of the onset of puberty	• Food and beverage packaging • Flame retardants • Building materials and electronic components • Paper coatings	[10]
3.	Caffeic and dihydrocaffeic acids	• Damages DNA in the presence of copper • Antioxidant property • Prevents cardiovascular disease • Reduces stress	• Wine • Coffee • Spices and herbs	[11]
4.	Chlorophenol	• Mouth scorching • Necrotic gashes in the respiratory canal	• Bleaching • Iron and steel industries • Paper pulp and paper board mills • Dye and pharmaceutical industries	[12]
5.	Hydroquinone	• Damages chromosome • Toxic to soil microbial activities • Skin irritation	• Food and Rubber industries as an antioxidant • Paint and fuel industries as a stabilizer • Cosmetic industries	[13]

Table 1. *Cont.*

S.No.	PhCs	Effects	Sources	Ref.
6.	Para-cresol	• Central nervous system disease • Cardiovascular system • Lungs and kidneys diseases	• Fumigants • Disinfectant • Explosives	[14]
7.	Estrogen	• Endocrine disability • Breast cancer	• Urine	[15]
8.	Dopamine	• Enhance brain functions and memory • Regulates blood pressure	• Brain	[16]
9.	Galliac acid	• Antioxidant • Reduce the risk of cardiovascular diseases • Prevent certain cancer	• Wine • Chestnut • Berries	[17]

The advantages of using biosensors include high sensitivity, accuracy, short response time, reliability, longer shelf life, and user-friendliness [7,18–20]. Biosensors can detect a very low amount of contaminant even from a composite medium, (wastewater) and are considered reliable for quantifying the total phenolic content in food by checking antioxidant properties of the foodstuff [21,22]. Several bio-recognition components have been used to detect polyphenols, such as microorganisms [23], DNA [24], whole cells [18], anti-bodies, and enzymes [25]. For monitoring PhCs, enzymatic biosensors based on laccase (Lac) [26], horseradish peroxidase [27], and tyrosinase [28,29] have proven to be the most effective due to their ability to directly catalyze electron transfer reactions without introducing cofactors into the reaction medium. Different supporting matrices such as conducting polymers, metal and metal oxide nanoparticles, carbon nano tubes (CNTs), graphite, silica gel, activated charcoal, reduced graphene oxide (RGO), glass surfaces, etc., have been used for the immobilization of these enzymes [30]. The comparative conductivities of such materials have been shown in Figure 1b.

In the current review article, we discuss:

- The recent progress in carbon-based nanomaterials (CNMs)-based Lac biosensors;
- The structure, mechanism of action, and immobilization methods of the Lac enzyme on CNMs;
- The application of CNMs-based Lac biosensors for the detection of PhCs present in food and body fluids;
- The limitations of highly utilized graphitic materials and challenges.

Recently, few articles have focused on discussing the importance of Lac enzymes for many applications. However, to the best of our knowledge, none of them focus on the application of CNMs, which can resolve the major challenge in the development of Lac-based biosensors for the detection of PhCs.

2. Transduction Principle for Monitoring PhCs

The biosensor has emerged as an interesting analytical approach for the detection of PhCs. It mainly consists of two components, including biological entities (analyte and receptor) and a transducer [31]. Enzyme-based sensors are prominently explored due to their ease of selectivity, sensitivity, reliability, and usage [32,33]. In this context, Lac has been considered a greener bio-recognition component supported on modified conducting carbon-based matrices for the detection of PhCs. An enzymatic biosensor produces signals targeted to the analyte's concentration by combining an enzyme with the transducer, and the transducer converts the signal received into a quantifiable response [34]. Amperometric biosensors aim to electrochemically convert non-active analytes into products by catalysis depending on the enzyme system that performs oxidation and reduction at the surface of the working electrode. Das et al. [35] fabricated Lac-based amperometric biosen-

sor based on an osmium tetroxide/poly 4-vinyl pyridine/multi-walled carbon nanotube (MWCNT)/Nafion/carbon black/GCE electrode for pyrocatechol detection in environmental samples. Kavetskyy et.al. [36] investigated an electroconductive immobilized matrix based on microporous carbon fibers and the Lac enzyme using the amperometric transduction principle. A successive assessment of biosensors was performed by detecting industrial pollutant catechol from actual communal wastewater samples. There is a limited number of optical biosensors reported in the literature that use CNMs for the detection of PhCs, one of which comprises carbon quantum dots (C-QDs) and the Lac enzyme [37]. An optical biosensor based on a sol-gel immobilized Lac was also developed for the detection of three isomeric PhCs (catechol, resorcinol, and hydroquinone) in real and tap water samples [38].

2.1. Electron Transfer and Reaction Mechanism in Lac

Lac is the largest subgroup of multi-copper oxidases, comprising four copper atoms in different oxidation states, as shown in Figure 2a. Lac has the capability to catalytically oxidize a variety of PhCs in the presence of molecular oxygen by performing a four-electron reduction of oxygen to water. The substrate undergoes reduction by the Lac enzyme conjoining its four-electron oxidation leading to the reductive cleavage of molecular oxygen (O=O) bonds by four electrons along with four Cu atoms. The Cu atoms forming the core of Lac are sub-divided into three types, based on the electron paramagnetic resonance (EPR) spectroscopic technique in the literature [39,40]:

- Type 1 (Cu C1) has trigonal coordination with one sulfur atom of cysteine (cys) and two nitrogen atoms of histidine imidazole units. The fourth coordination is with S of methionine attached axially and far more than the other three. Therefore, the structure seems distorted in the form of tetrahedral geometry with triagonal elongation [41]. It possesses intense blue color resulting from the strong electronic absorbance at 600 nm (charge transfer: S of cysteine to Cu, ε = 5000 M^{-1} cm^{-1}) in UV/visible spectroscopy, and its paramagnetic nature has been confirmed using EPR spectroscopy.
- Type 2 (Cu C2) has been found to be coordinated with two nitrogen atoms of histidine units and one water ligand. Although it is colorless, EPR studies have revealed its paramagnetic nature via ultrafine splitting.
- Type 3 (2Cu C3) comprises two anti-ferromagnetically coupled Cu atoms each tetragonally coordinated with three nitrogen atoms of histidine units and one bridged hydroxide group. The bridged hydroxide leads to electron-paired Cu sites. The oxidized form shows weak absorbance under UV-visible spectrum having a shoulder at around 330 nm (charge transfer: OH to Cu), and EPR studies show no signal signifying their diamagnetic nature.

The histidine units are, therefore, coordinated in a 2:2:6 ratio for Cu C1, Cu C2, and 2Cu C3, respectively. The bonds connecting Cu and N of histidine are slightly different in bond lengths. The approximate Cu–N bond length for all categories of Cu is 2 Å A, some Cu–N bonds are of greater bond length, and some are found to be slightly smaller. The distance between Cu C1 and binuclear Cu atoms (2Cu C3) is found to be around 12 Å, whereas, in the trinuclear Cu cluster, the distance between Cu C2 and binuclear Cu atoms (2Cu C3) is observed at about 4 Å.

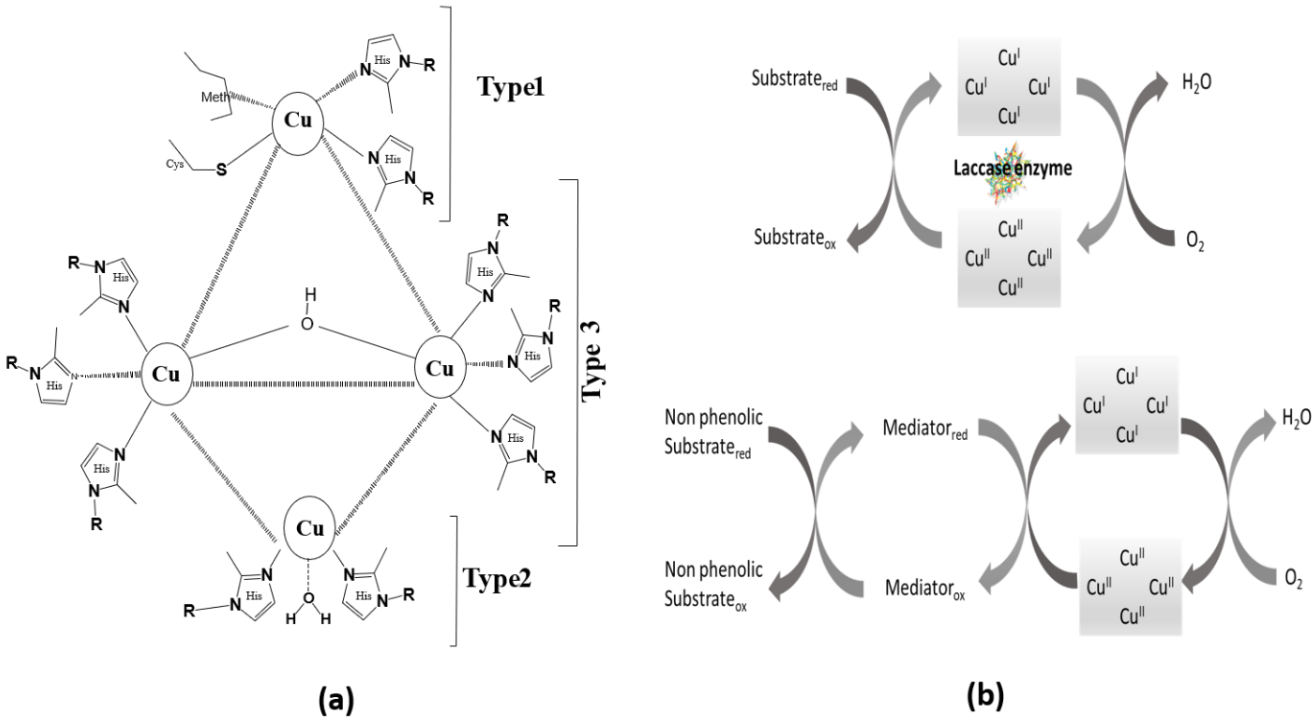


Figure 2. Scheme showing the (**a**) assembly of the Lac enzyme, adapted from ref. [39] and (**b**) the activity of the Lac enzyme for a phenolic and non-phenolic substrate [42].

The Electron Paramagnetic Resonance (EPR) of Lac have been observed to be perturbed in the presence of strong anion binding, leading to the disturbance of the EPR signal intensity of Cu C2 and in the anti-ferromagnetic coupling of 2Cu C3. At the same time, the electron reduction of molecular oxygen and discharge of water takes place at a trinuclear cluster of Cu C2 and 2Cu C3 atoms. Based on the physical and chemical properties of oxidoreductase Lac, the positioning of copper atoms has different well-defined redox potentials (E_0). Depending upon the chemical nature, the Lac enzymes obtained from basidiomycetes (mostly white rot fungi) are considered to have high redox potential, whereas low redox potential is found to be possessed by Lac obtained from bacterial and plant sources [40]. Other biotic sources of Lac enzymes have been listed in Table S1. However, the physical structure shows that Cu C1 and 2Cu C3 have variable low ($+E_0$ = 0.4–0.5 V) and high (E_0 = 0.7–0.8 V) redox potentials with respect to standard hydrogen electrodes. The low and high redox potential of Cu C2 appears at approximately 0.4 V [43].

Biosensors based on tyrosinase, horseradish peroxidase, Lac, and polyphenol oxidase enzymes have been broadly used for the detection of different PhCs. However, Lac possesses higher stability and better catalytic ability for electron transfer reactions [44]. Lac-based biosensors are modest in assembling over other reported enzymatic biosensor because they do not necessarily involve hydrogen peroxide as one of the co-substrates or some other co-factors for the catalytic process.

2.2. Activity of the Lac Enzyme

The catalytic activity of Lac depends on the copper (Cu) atoms. All the Cu sites show different functions. Cu C1 is the primary electron acceptor from the reduced substrate. The high redox potential (790 mV) of Cu C1 is the reason for substrate oxidation at its site [39]. Further, the electron received by Cu C1 is delivered via Cys-His pathways to the trinuclear Cu cluster that shows an inter-electron transfer mechanism, which takes place between Cu C1 and 2Cu C3 positions. Due to this transfer of an electron from one Cu site to another, the step where the oxygen molecule is reduced to water by four electrons replicates. It has been reported that H_2O_2 was not detected during the steady state [45]. At the same site, an

oxidized form of the enzyme has been generated. During this aerobic oxidation process, 2Cu C3 accepts two electrons in the presence of Cu C2. The presence of Cu C2 is found to be necessary for the acceptance of electrons by 2Cu C3. In short, the Lac enzyme is oxidized by reducing oxygen to water followed by the oxidation of the substrate leading towards the terminus of the catalytic mechanism by reducing the oxidized Lac enzyme to Lac again.

The Lac enzyme does not possess the ability to oxidize non-phenolic substrates directly. To increase the range of compounds that can be oxidized using the Lac enzyme, several Lac-based mediator compounds have been reported. These Lac-based mediator compounds consider the substrate as an intermediate, exhibiting higher redox potential and indirectly oxidizing non-phenolic substrates [43] (Figure 2b). The mediators enhance the catalytic performance of the Lac enzyme, acting as a co-catalyst. The catalytic mechanism for non-phenolic compounds initiates with the conversion of Lac into oxidized Lac on the same side, leading to the reduction of oxygen to release water. The mediator works by oxidizing the substrate and reducing itself to the original mediator, indicating completion of the co-catalytic mechanism [42]. A standard mediator must be low in molecular weight to avoid hindrance due to size and should possess a high redox potential to enhance the catalytic oxidation process. An efficient mediator must perform several consecutive cycles deprived of deterioration. For example, (2,2′-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid)) is one of the most frequently used mediator for oxidation of non-phenolic compounds [43].

Enzymes undergo specific interactions with certain chemical carriers and substrates to undergo genetic and chemical modifications to increase their catalytic activity. In this context, Zhang et.al. reported that flower-like $Cu_3(PO_4)_2.3H_2O$ nanocrystals integrated with Lac shows 6 times activation in the activity of the Lac enzyme [46]. Whereas, Wu et al. found a 1.5 to 4 times increase in the activity of fungal Lac after its pre-incubation with different organic solvents such as methanol, acetone, dimethylformamide, and dimethyl sulfoxide [47]. On the other hand, the influence on enzymatic activity has also been studied based on mutagenesis focusing on the residue at certain position and pH. Another method to enhance activity is through the use of multienzyme, which involves channeling different substrates, kinetics matching, and 3-D distribution of the involved enzymes. These methods can be utilized to engineer the best enzyme–substrate relationship to provide superior catalytic performance [46].

2.3. Immobilization Matrix

In addition to the several merits of the Lac enzyme, the use of the Lac enzyme for biosensor application has not been explored much because of the denaturation of the enzyme by the external environment. In order to increase the stability, lifetime, efficiency, and reusability, and to equilibrate the cost of developed biosensors for detection of PhCs, the efficient immobilization of the Lac enzyme onto suitable matrices is very important [48]. Immobilization provides a way to increase the ability of enzymes by reducing the cost of production [49], leading to better stability of the biosensor [50], strategies for reproducibility, ease of recovery [49], etc. Moreover, the involvement of new bonding, entrapment, and crosslinking also enhances the shelf life of enzymes [51,52]. Immobilization maintains the structural stability and functional possessions of the immobilized enzyme by proper attachment to the surface of the appropriate matrix so that the required activity of the enzyme can be retained throughout the repetitive use of sensor [53,54]. Immobilization methods vary from one enzyme to another and depend on the type of application for which the enzymes are being used. An ideal matrix must be physically rigid, chemically inert, and insoluble in enzymes to preserve its catalytic properties. The immobilization matrix should be thermally and photo-chemically stable, showing effective charge transfer capability. Furthermore, there must be permissible diffusion of a bio-catalytic reaction between the substrate and matrices on which the enzyme has to be immobilized [55].

Effective immobilization also helps to restrict the gross movement of biomolecules, thus leading to the fabrication of a stable and accurate biosensor. The different immobilization techniques include (i) covalent bonding, (ii) adsorption, (iii) cross-linking, (iv) en-

trapment, and (v) encapsulation (Figure 3a–c) [56]. One more method, i.e., the electrospray deposition technique, has also been reported for Lac immobilization on carbon black-nanomodified screen-printed electrodes, as shown in Figure 3d [57]. Several physical and chemical factors, such as pH, temperature, solubility, concentration, etc., influence the immobilization strategies. Large numbers of organic and inorganic matrices have been used for immobilizing the Lac enzyme. In particular, conducting polymers [58], metal nanoparticles [59], metal oxides [60], silica [61,62], clay [63], CNMs, metal–organic frameworks [64], etc. Conducting polymers (CPs) have been widely explored for the fabrication of Lac-based biosensors because of their good compatibility with enzymes and electron conduction ability. In the case of CPs, conductivity can also be varied by hosting doping materials. Calitri et al. [65] evaluated the total phenolic content using CPs poly(aniline-co-2-aminobenzylamine) supported on MWCNT by immobilizing Lac. Yaropolov et al. [66] fabricated a Lac biosensor using three CPs electrodes, viz., Nafion, poly(ethyleneimine), and Eastman AQ 29D supported with GCE for amperometric determination of PhCs. The synthesis of polydopamine nanoparticles based on a bacterial cellulose composite was also accomplished as a compatible matrix for Lac immobilization. Meso-porous silica sieve matrices possess compatibility and uniformity with respect to well-ordered pore structure, size, density, and surface characteristics. Amperometric detection of catechol has been reported where magnesium-doped mesoporous silica sieve–polyvinyl alcohol composite has been used as a matrix [67]. Li et al. constructed a disposable biosensor using silica sphere on the surface of MWCNT with a screen printed electrode (SPE) for immobilizing Lac and detecting dopamine [68]. The major problem with this sensor is the film formation ability and lower conductivity.

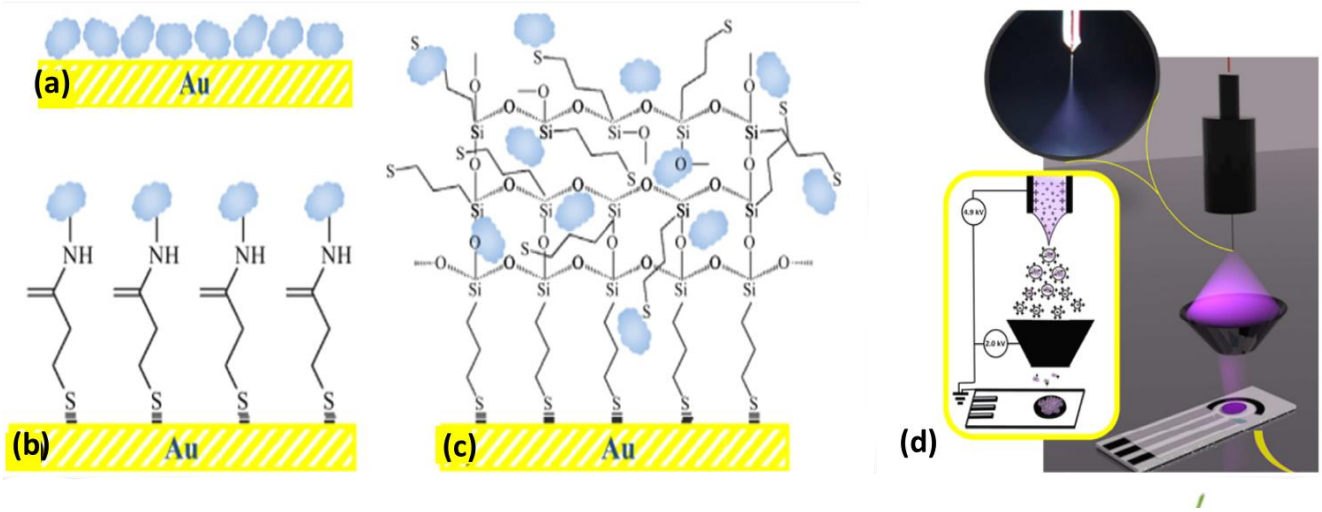


Figure 3. Enzyme immobilization methodologies employed for biosensor fabrication using Lac onto a Au matrix: (**a**) direct adsorption, (**b**) covalent binding, (**c**) sol-gel encapsulation, Adapted with permission from Ref. [56]. Copyright © 2013 Elsevier and (**d**) the electrospray deposition technique for Lac immobilization on carbon black-nanomodified SPE. Adapted with permission from Ref. [57]. Copyright © 2020 Elsevier.

The recent era specifically focuses more on immobilization matrices based on nanomaterials. Although many other matrices have been used to support enzymes, nanomaterials-based matrices are widely explored because of their numerous benefits. Introducing nanomaterial matrices into the enzyme-based biosensors led to improved sensitivity, lower detection limit and higher response time, better stability and rigidity, a longer shelf life, and speeding up the fabrication process [69]. There is a broad spectrum of nanomaterials, including metal nanoparticles, metal oxide nanoparticles, nanostructured polymer composites, nanofibers, nanorods, etc., utilized in the field of biosensing for polyphenol

detection. Sofia et al. demonstrated numerous methods for covalent immobilization of the Lac onto several matrices (Figure 4), and concluded that the approach based on the linkage of carbon-based materials through EDC-NHS to −COOH functionalized matrices was less effective than the Lac bonded with the -NH_2 group using glutaraldehyde to support amminated carbon based matrices [70].

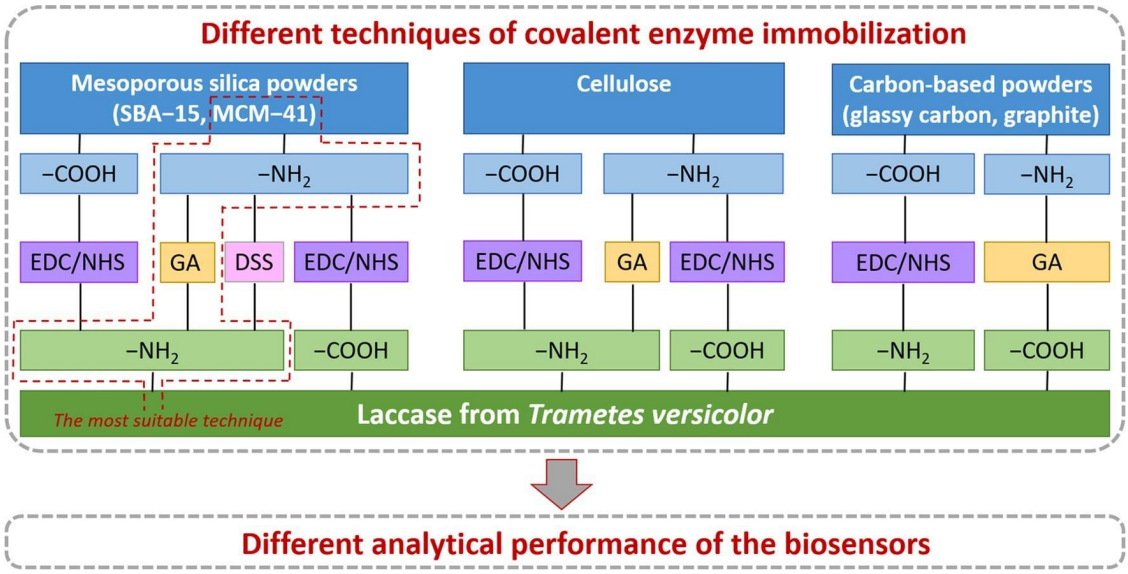


Figure 4. Illustration of different methods of covalent immobilization of laccase onto a variety of matrices Adapted with Permission from Ref. [70]. Copyright © 2022 Elsevier.

3. Carbonaceous Nanomaterial-Based Lac Biosensor

Carbon-based nanomaterials (CNMs) are one of the most explored nanomaterials for biosensor applications. Carbon nanomaterials have comparable dimensions to redox proteins, and thus can be used as an effective electrical connector with redox enzymes. The most common CNMs used to support the Lac enzyme in biosensing applications include graphene (Gr), reduced GO, carbon nanotubes, carbon quantum dots, and graphene quantum dots (Figure 5) as summarized in Table S2. These CNMs possess a tremendous ability to detect phenolic pollutants as they have a large surface-to-volume ratio, which in addition to sensing, helps in the adsorption of pollutants. Table 2 shows the application of different CNMs-based Lac biosensors for the detection of phenols.

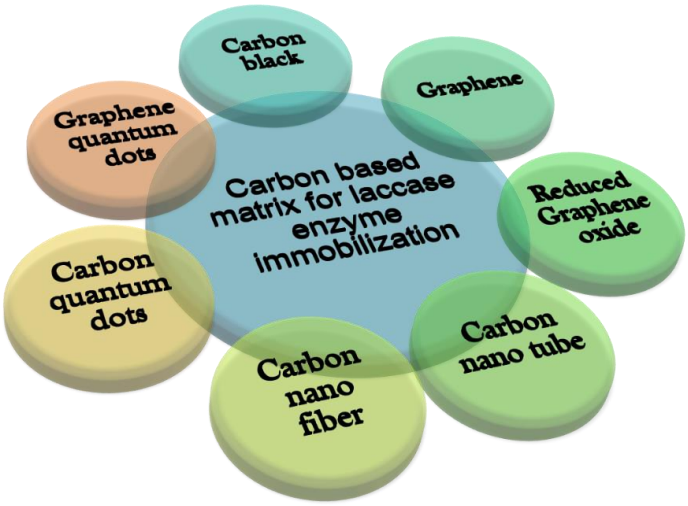


Figure 5. Types of carbon-based immobilization matrices.

Table 2. Biosensors used for the detection of phenols using the Lac enzyme.

S.No.	Matrix	Method of Analysis	Analyte	Matrices	Linear Range	LOD	Ref.
1.	PANI/MG composite	ChronoAmp	Hydroquinone	Gr	0.4–337.2 µM	2.94 µM	[71]
2.	Microporous carbon fibers	ChronoAmp	Catechol	C F	0.01–0.05 mM	<5 µM	[36]
3.	Au NP/graphene NP-SPE	ChronoAmp	Hydroquinone	Gr	4–130 µM	1.5 µM	[72]
4.	Polyvinylpyrrolidone/CS/RGO	Amp	17 α-Ethinylestradiol	RGO	0.25–20 pmol L^{-1}	0.15 pM	[73]
5.	Rh/GO	DPV	17β-estradiol	GO	0.9–11 pM	0.54 pM	[74]
6.	PEDOT, GO nano-sheets	DPV	Catechol	GO	0.036–0.35 µM and 0.35–2.5 µM,	0.032 µM	[75]
7.	RGO-MWCNT	ChronoAmp	Epicatechin equivalents	RGO & MWCNT	1–300 µM	0.3 µM	[76]
8.	Poly L-lysine/Citric acid-Gr/GCE	DPV	17β-estradiol	Gr	4×10^{-13}–5.7×10^{-11} M	1.3×10^{-13} M	[77]
9.	Poly(dithienotetraphenylsilane)	Fluorescence	Dopamine	G-QDs	1–200 µM	80 nM	[78]
10.	SPCE/anthraquinone-COOH-MWCNT	Amp	Catechol	MWCNT	0.002–0.061 µM		[79]
11.	G/PANABA/MWCNT	ChronoAmp	Phenol	G & MWCNT	0.0005–0.4 mM	0.5 µM	[65]
12.	Bacterial Cellulose/cMWCNTs/ZIF-8		Bisphenol A	MWCNT	0.01–0.4 mM	1.95 mM	[80]
13.	Pt NP/GO	SWV	Chlorogenic acid	GO	0.56–7.3 µmol L^{-1}	0.18 and 0.59 µM	[81]
14.	Au–ZnO/NP/ITO	Amp	Catechol	NC	75 nM–1100 µM	25 nM	[82]
15.	D glucan/carbon black paste/Au NP	SWV	Hydroquinone	C	2.00–56.5 µM	0.474 µM	[83]
16.	F, N-doped carbon dots	CV	Catechol	F,N-CD	0.1–0.45 mM	0.014 µM	[84]
17.	Thin polydopamine film/carbon surfaces	ChronoAmp	Caffeic acid, rosmarinic acid, and gallic acid	C	1–150 µM	0.29 µM	[85]
18.	PPy/GCE	Amp	Catechol	GCE	1–60 µM		[86]
19.	PANI/ CuCNFs	Amp	Hydroquinone	C/Cu NF	500 nM–110 µM	0.24 µM	[87]
20.	BMIMBF4-CS and MWCNT	SWV	Bisphenol A	MWCNT		8.4 ± 0.3 nM	[88]
21.	Plastic packaging waste derived CNTs/SPCE		Para cresol	CNT	0.2–25 ppm	0.05 ppm	[89]
22.	Au-MXene (Ti_3C_2)	Amp	Catechol	Ti_3C_2	0.05–0.15 µM	0.05 µM	[90]
23.	Enzyme POXA1b and POXA1b-Vmh2/MWCNT	ChronoAmp	Catechol and Dopamine	MWCNT	2–30 pM, 0.1–800 µM and 0.015–90 µM	2 pM and 15 nM	[91]
24.	BC/c-MWCNTs/LAC@ZIF-90 membrane	DPV	Catechol	MWCNT	20–400 µM	1.86 µM	[92]
25.	CMB/CBPE	SWV	Quercetin	Carbon black	4.98–50.0 × 10^{-8} M	2.6 × 10^{-8} M	[93]
26.	OMC-SPE	SWV	Serotinin	Mesoporous carbon	0.1–1.2 mu M	316 nM	[94]

Table 2. *Cont.*

S.No.	Matrix	Method of Analysis	Analyte	Matrices	Linear Range	LOD	Ref.
27.	TiO_2/nafion/graphitic	CV	Gallic acid	G	0.125–175 mu M	0.125 mu M	[95]
28.	Fe_3O_4-Pc-cMWCNTs	DPV	Rosenmerinic acid	cMWCNT	0.2–400 mu M	0.182 mu M	[96]
29.	MnO_2/GNP decorated SPCE	Amp	Caffeic acid	GNP	0.3 µM–0.4 mM	1.9 mu M	[97]
30.	Cellulase/c-MWCNTs	Amp	Catechol	cMWCNT	10–160 mu M	0.004 mu M	[98]
31.	Au-RGO/SPE	DPV	Catechol	RGO	1 mM–1 nM	3.3 µM	[99]
32.	SPCE modified Gr-AuNPs with CS	DPV	Bisphenol A	Gr	0.05–12 µM	0.023 mu M	[100]
33.	Ag-ZnO/ MWCNTs/SPE	DPV	Bisphenol A	MWCNT	0.5–2.99 mu M	6 nM	[101]
34.	CS-Fe_2O_3/RGO	ChronoAmp	Bisphenol A	RGO	6–228 ppb	18 nM	[102]
35.	GO/Fe Pc composite		Adrenaline	GO	1.8–92 µM		[103]
36.	TiO_2-GPE	DPV	Methyldopa	GPE	10–180 mu M	1 mu M	[104]
37.	G-QDs		Catecholamine	G-QDs	1–120 µM	83 nM	[105]
38.	Gr/Cu/Fe_3O_4 composites	DPV	Bisphenol A	Gr	7.2–18 µM	1.7 µM	[106]
39.	PEDOT/Au/cMWCNT		Catechol	cMWCNT	0.1–0.5 and 11.99–94.11 µM	0.11 and 12.26 µM	[107]
40.	Gr/PPy nanotubes/$SrCuO_2$	DPV	2,4-Di chlorophenol	Gr	1–50 mu M	0.18 mu M	[108]
41.	RGO-MoS_2	ChronoAmp	Hydroquinone	RGO	1–100 µM	0.1 µM	[109]

3.1. Carbon Nano-Tube (CNT)-Based Lac Biosensors

The benzene-type hexagonal arrangement of carbon atoms forming hefty cylindrical molecules constitutes a quasi-one-dimensional CNT [110,111]. This allotrope of carbon has an enormous surface area in terms of a large length-to-diameter ratio and is observed to be just one atom thick [111]. CNTs exhibit extraordinary properties, such as being lightweight and stable, with excellent electronic, thermal, and mechanical properties [112]. A CNT lacks solubility in aqueous media, and the introduction of chemical functionalization increases their solubility in different solvents and increases their potential for biosensor applications [113]. CNTs are conjugated with many metal and metal oxide nanoparticles such as Au, Fe, Pd, Pt, Cu, Fe_3O_4, ZrO_2, etc., for biosensing application [79,114]. However, certain factors such as the bundling effect, increment in the number of walls, and aggregation can cause problem while using CNTs.

3.2. Carbon Nanofibers (CNFs)-Based Lac Biosensors

A CNF is a one-dimensional nano range fiber with a hollow core with a diameter of 10 to 500 nm and a length in the range of 0.5 to 200 µm [115]. In CNFs, strands of layered stratified graphite sheets are stacked on a single molecule which have different variations such as platelets, ribbon, cones, herringbone, cups, etc., in the nano size range [116,117]. It has remarkable mechanical and chemical properties, due to which they are superior to other fibers. The presence of a graphitic structure enables their larger area and excellent thermal and electrical conductivities. In addition, they do not become oxidized easily. Other characteristics such as a low density, a high Young's modulus, thermally stability, low defects, a more significant aspect ratio, and a condensed structure make CNFs suitable matrix materials [118]. These advantages lead to their applications in ceramics, fuel cells, cement composition, etc. [119]. Yang et al., used electrospinning, carbonization, and the solvothermal technique to fabricate a novel Lac-based biosensor for monitoring hydroquinone using TiO_2 decorated copper and carbon composite nanofibers [120]. Recently, combined electrospinning and an in situ polymerization technique for synthesizing polyaniline-encapsulated carbon/copper composite nanofibers have been reported for the detection of hydroquinone in river water [87].

3.3. Carbon Quantum Dots (C-QDs)-Based Lac Biosensors

C-QDs are low-cost and water-soluble QDs that are easy to synthesize and have core carbon atoms organized in sp^2 hybridization bearing oxygen-containing functional groups (hydroxyl, carboxyl, carbonyl) [121,122]. The biocompatibility, electronic properties, photostability, and photo-luminescent radiation emission of C-QDs have been explored for different applications such as photodegradation, sensing, bioimaging, drug delivery, catalysis, energy conversion, optical electronics, etc. [123].

C-QDs do not interact with all analytes directly, and very few biosensors have been reported based on C-QDs as a biosensor matrix for the detection of phenols. Recently, a fluorescent sensor has been reported utilizing the unique optical properties of C-QDs for the detection of dopamine using photo-luminescent [37]. However, the quantum yield of C-QDs is generally low, and they are difficult to reproduce due to their geometric arrangement and challenging structural characterization [124].

3.4. Graphite-Based Lac Biosensors

Graphite has been utilized as electrode material in amperometric biosensors due to its exceptional electronic and catalytic properties. Cato et al. used bare and modified graphite-epoxy decorated with either tyrosinase or the Lac enzyme and copper nanoparticles for the voltammetry detection of antioxidant activity phenols that are present in rosé cava wines [125,126]. Ibarra Escutia et al. also developed an amperometric biosensor, entrapping the fungal Lac enzyme onto polyvinyl alcohol photopolymer and graphitic SPE for sensing PhCs in tea infusions [127]. However, graphite lacs in scope for the modification of an electrode because of its lower functionality and poor solubility in organic solvents, due

to which the derivatives of graphite offer more promising biosensing properties and are highly utilized.

3.5. Graphene (Gr) and Its Derivatives-Based Lac Biosensors

Gr is a single-layered sheet of carbon obtained after isolation from stacked layers of graphite having a thickness in atomic dimensions bearing a hexagonal layout of sp^2 hybridized carbon atoms. Gr provides a large surface area, excellent electrical conductivity due to its highly conjugated structure, comparable tensile strength, and remarkable biocompatibility [128]. Gr-based materials provide an enhanced signal response in electrochemical biosensing [129]. Due to these properties, Gr sheets have been widely utilized as an immobilization matrix for immobilizing various enzymes, including Lac, for detecting various phenolic compounds. Palanisamy et al. fabricated a Lac enzyme-based biosensor to detect catechol using Gr cellulose microfiber composite-improved screen-printed carbon electrodes [44]. A biosensor based on the Lac enzyme immobilized on a polymer-like polyaniline (PANI) and magnetic graphene electrode has been introduced recently by Lou et al. for detecting hydroquinone [71].

Graphene Oxide (GO) is a derivative of Gr having abundant epoxy functional groups on the surface of Gr sheets with a large surface-to-volume ratio. Although the conductivity of GO is lower than graphene (presence of oxygen functionality), it has been reported as an immobilization matrix due to its solubility in water and its potential for the immobilization of enzymes on its surface without the use of crosslinker [130,131]. Maleki, et al. fabricated a biosensor using poly(3,4-methylenedioxy-thiophene), GO nano-sheets, and the Lac enzyme to detect catechol [75]. An electrochemical sensor utilizing GO and a molecularly imprinted polymer was developed for selective and sensitive detection of 2,4-dichlorophenol in water. It was also optimized to estimate 2,4-dichlorophenol in real water samples [132]. The introduction of oxygen functionality to GO led to a reduction in the conductivity of GO, while the debut of RGO led to better conductivity compared to that of pristine graphene. Moreover, it provides promising characteristics such as easy assembling, cost-effectiveness, higher mobility, biocompatibility, etc. [133]. A large array of materials and composites have been used for doping RGO sheets to synergistically enhance the catalytic activity and electron transfer process of biosensors for the detection of different phenolic pollutants such as 17β-estradiol [74], catechol [134], estriol [135], etc. Using RGO as a matrix for the Lac enzyme, Mei et al. fabricated RGO palladium–copper electrodes for the sensitive determination of catechol [134]. Eremia et al. developed a disposable biosensor using platinum nanoparticles/RGO composite supported on carbon screen-printed electrodes for the assessment of caffeic acid. They reported the kinetics of the detection of polyphenol electrochemically and claimed that the fabricated biosensor can be used for evaluating the total phenolic content in tea infusions [136]. Despite all the benefits, there are certain limitations of using GO and RGO such as quick agglomeration, reduction in conductivity after undergoing functionalization, and the use of toxic chemicals during synthesis. These can perhaps be overcome by the use of advanced 2D materials.

3.6. Graphene Quantum Dots (G-QDs)-Based Lac Biosensors

G-QDs are considered zero-dimensional nanomaterials which are derived from the 2D sheet of graphene [137]. G-QDs exhibit better solubility than CNTs and displays biocompatibility, low toxicity, and stability. High quantum yield, edge effects, and confinement make it different from other carbon nanomaterials [138,139]. Top-down synthesis of G-QDs from graphene is a multistep and cost-effective process, while bottom-up synthesis of G-QDs from polycyclic aromatic compounds proceeds via complex steps and requires costly chemicals and equipment. G-QDs demonstrate remarkable photoluminescence and robust quantum properties [140]. Baluta, et al. used the fluorescence technique to detect dopamine using a Lac-based ceramic biosensor supported by graphene QDs. The fabricated biosensor was also modified using a conducting polymer [78]. Unlike graphene, G-QDs have zero band gap and can be modified using functional groups due to the edge effect [97].

Vasilescu et al. modified G-QDs electrodes using molybdenum disulfide, and the fabricated Lac-based biosensor has been used for the detection of caffeic acid and were proven to be useful for estimating the total phenolic content in red wine [141]. Inspite of several applications, the use of G-QDs is still challenging as the formation of a single layer of high-quality G-QDs is difficult. In addition, its synthesis suffers from toxic organic solvents, a long reaction time, and low yield, which limits its use for industrial purposes. Different applications of CNMs to detect PhCs have been summarized in Figure 6.

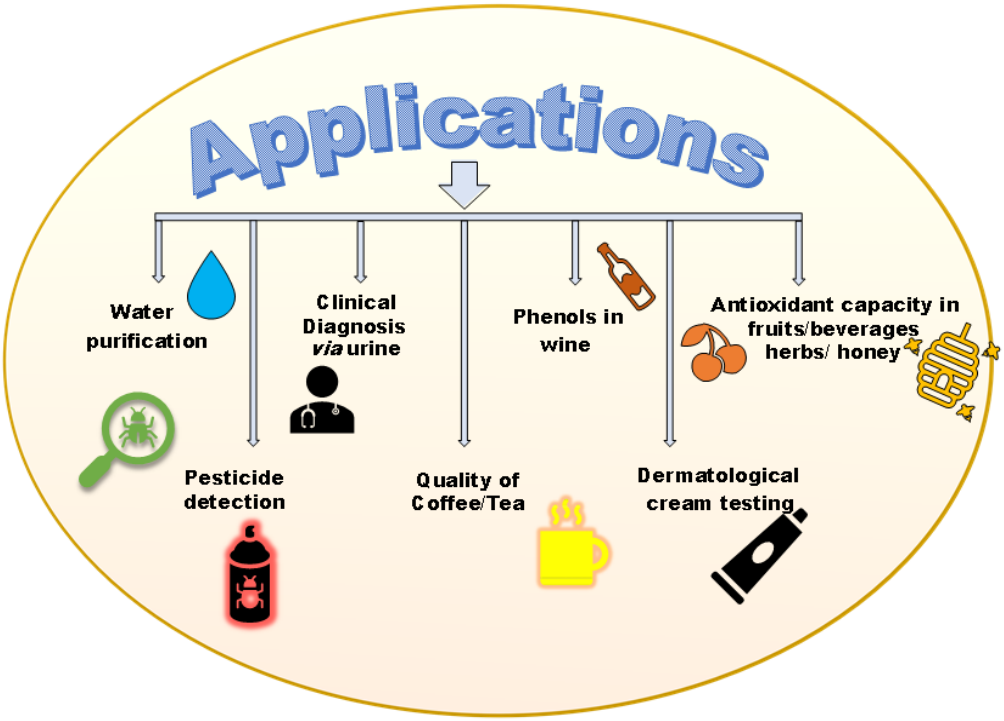


Figure 6. Applications of Lac enzyme-based biosensors.

4. Applications of Lac-Based Biosensors

4.1. Detection of Phenolic Pollutants in Wastewater

Water bodies contain waste from different sources in the form of dyes and harmful chemicals that affect various life forms (both flora and fauna). Toxic chemicals, dyes, and PhCs are major raw contributors from various textile, plastics, leather, and paper industries. PhCs also merge into water bodies from the oil, gas, and coal industries. Conventional methods such as photo-catalytic degradation, distillation, ozonation, extraction techniques, membrane processes, advanced oxidation process, and enzymatic biological methods have been used to remove these dyes and phenolic pollutants from wastewater [142,143]. Before their removal, it is important to implement a real-time monitoring device to determine which pollutant is to be treated and how much to reduce the cost and time for wastewater treatment methods is very crucial. The major water pollution-contributing phenolic compounds are catechol, hydroquinone, and Bisphenol-A.

Hydroquinone (Hq) is a widely used raw material in chemical industries as a dye intermediate and black and white photo developer and in cosmetics as a skin-lightening agent as well as a polymerization inhibitor. Studies reveal that high exposure to Hq may lead to skin-irritating effects, and it is also harmful to soil microorganisms. Recently, PANI/magnetic graphene immobilized Lac enzyme-modified electrodes were developed that show superior biosensing activity for Hq in real water samples [71]. Upan, et al. introduced a flow-injection amperometric sensor using a glutaraldehyde cross-linked Lac enzyme onto CNT-modified SPE for estimation of Hq in water sample [144]. rGO-MoS_2 nanocomposite has been used to detect Hq using chronoamperometry technique [109]. Synthetic catechol, on the other hand, has been utilized extensively in the production of

pesticides, pharmaceuticals, chemical perfumes, etc. [145,146]. Similar to Hq, it is also used in the field of black-and-white photography. Several biosensors have been developed for the evaluation of catechol. A biosensor based on an artificial neural network-integrated system has been formulated for the assessment of real water samples containing catechol under a wider linear range using poly(3,4-ethylenedioxythiophene) (PEDOT), GO nano-sheets, and Lac [75]. Kong, et al. used a graphite electrode-based intercalated montmorillonite for the electrochemical determination of catechol [147]. One of the fabricated biosensor-based on Lac/Polyvinylpyrrolidone/chitosan/rGO electrospun has been shown in Figure 7a [73]. Also, simultaneous detection of Hq and catechol has been carried out using Lac/aminopyrine/RGO/GCE electrode in water samples d [148] (Figure 7b). Bisphenol A is a component of hard plastic that are being used daily in different forms such as bottles, medical equipment, household electronic items, sports equipment, etc. [149]. Schematic representation of Lac/CNT/SPCE electrode for detecting p−cresol in real water samples and p−cresol produced during the Fenton process has been depicted in Figure 7c [89].

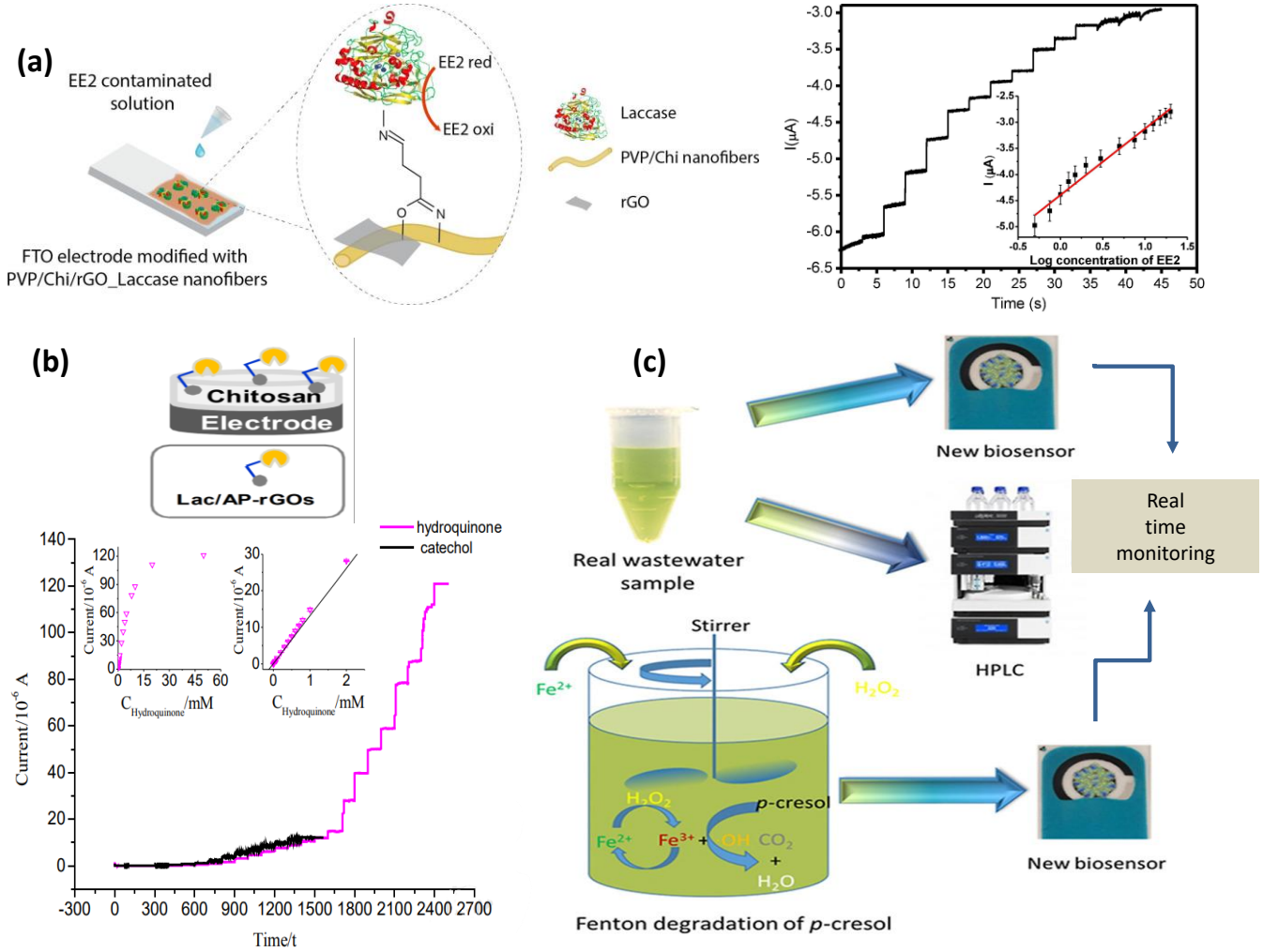


Figure 7. (**a**) Illustration for Lac/Polyvinylpyrrolidone/chitosan/rGO electrospun nanofibers for 17α− Ethinylestradiol electrochemical detection in synthetic and human urine. Adapted with permission from Ref. [73]. Copyright © 2018 Elsevier, (**b**) Schematic representation of Lac/aminopyrene/RGO/GCE electrode for the detection of phenols in water samples. Adapted with permission from Ref. [148]. Copyright © 2013 Elsevier (**c**) Lac/CNT/SPCE electrode for detecting p−cresol in real water samples and p−cresol produced during the Fenton process. Adapted with permission from Ref. [89]. Copyright © 2021 Elsevier.

Bisphenol A leaches into our bodies mainly through food containers as it is used to coat the inside of food containers and drink cans. Numerous studies reveal that it acts as an endocrine disrupter even at a very low dose [150]. Bisphenol A can cause DNA damage, so monitoring its amount in water is necessary. Jalalvand, et al. used biosensing electrodes as layer-by-layer modified GCE with methylene blue-DNA/MWCNTs-chitosan/palladium nanoparticles/fullerene C60 [151]. An rGO-(4-ferrocenylethyne) phenylamine/AuNPs/GCE electrode has been used for the detection of Bisphenol A infused in milk samples [152]. Lac–thionine–carbon black-modified SPE was formulated by Portaccio, M., et al. for the assessment of endocrine disruptors in tomato juice samples [153]. Paracresol is also a water-polluting phenol, as illustrated in Figure 6a, and quite recently, a Lac/CNT/SPCE electrode has been used for its detection in real water samples. More importantly, paracresol formed while performing the Fenton process has a good resemblance with that of the results obtained using the HPLC technique [89].

One of the main causes of the accumulation of phenolic compounds in water bodies is their heavy usage in the agricultural sector as insecticides, pesticides, and herbicides. Pesticides such as pentachlorophenol undergo degradation in water into different chlorophenols. Chlorophenols easily leach from soil into water bodies as they are water-soluble and have also been widely used for impregnating wood [154]. The washing away of byproducts from agricultural materials into water bodies plays a significant role in polluting water. Biosensors help in the detection of these pesticides with ease. Carbamates are among widely used pesticides in insecticides, nematicides, fungicides, acaricides, or herbicides for increasing the yield of crops as a prophylactic measure. However, for non-targeted organisms, carbamates are toxic due to their deliberate discharge into the soil and subsequently into water bodies, leading to ill effects on the environment and human health worldwide [155]. These are also among the list of endocrine disruptors as of Bisphenol A [156], and thus become an important analyte for intensive monitoring. A novel enzyme-based biosensor has been developed by Oliveira et al. for the estimation of carbamate pesticide applied on tomato and potato crops using direct immobilization of Prussian blue functionalized Lac on graphene-modified carbon paste electrodes [157]. They also detected pirimicarb pesticide by immobilizing Lac on composite carbon paste electrodes containing an MWCNTs paste electrode modified by the dispersion of Lac within the optimum composite matrix [158]. Pesticides cause a higher threat to vegetables and fruits, which easily carry food-borne microbes and pesticide residue into our bodies when eaten raw [159]. Therefore, it is necessary to investigate the amounts of pesticides to protect the endocrine system of human beings and animals. For governing food safety, Oliveira et al. constructed a dual enzyme (Lac and tyrosinase)-based biosensor immobilized on to Au nanoparticles, chitosan, and graphene-doped carbon paste electrodes in order to detect the number of different carbamate residues in fruits [160].

4.2. Detection of Phenolic Compounds in Food

PhCs from natural compounds are found in vegetables, fruits, cereals, beverages, and red wines. The nutritional value of plants, fruits, and vegetables can be determined by quantifying their PhCs content [161]. It has been reported that PhCs rich food items prevent oxidative damage leading to age-associated diseases by terminating free radicals generated after cell metabolism [162]. Biosensors have proven to be an important technique to check the monitoring of food quality by detecting PhCs in plants and food products. De Macêdo, et al. estimated the total phenolic content with the help of a Lac immobilized graphite electrode in dry extracts of red fruits [163]. Honey is commonly used as an alternative to normal sugar due to its good antioxidant properties [164]. The authenticity of honey can be evaluated by inspecting its total phenolic content and antioxidant capacity using a real-time biosensor. De Oliveira Neto reported an electroanalytical method for immobilizing the Lac enzyme onto a carbon paste electrode for detecting both the antioxidant nature and total phenolic content of honey. The results correlate with the data obtained using the spectroscopic method [165].

Among generally consumed beverages, tea contains various PhCs (phenolic acids, flavonoids, tannins, catechins, etc.), possessing a variety of biological activities that can lead to cancer prevention, anti-aging, cardiovascular diseases, and oxidative stress in the brain. Tea infusions have excellent antioxidant properties and contain a high percentage of caffeic acid. Caffeic acid was determined from a tea infusion by Eremia, who constructed disposable biosensors based on platinum nanoparticles-RGO-Lac biocomposite electrodes [136]. Rawal et al. constructed Lac/Fe_3O_4NPs/cMWCNT/PANI/Au, Lac/MnO_2NPs/cMWCNT/PANI/Au, and Lac/AgNPs/cMWCNT/PANI/Au electrodes for measuring the total content of phenols in tea leaf extract [59,166]. The total phenolic content in tea leaf extract and in black radish root was also quantified by Ibarra-Escutia, P., et al. with the help of an amperometric biosensor based on carbon SPE [167]. On the other hand, coffee, a worldwide beloved and traded beverage, holds potential for reducing high blood pressure and high cholesterol, risk to pancreatic cancer, and bone loss. All varieties of coffee contain lipids, vitamins, and carbohydrates accompanied by variable amounts of PhCs including chlorogenic acid and caffeic acid that are responsible for forecasting their virtuous qualities and categories [168]. Chlorogenic acid is the major phenol in coffee. Moreover, it is also a chief antioxidant among this beverage. GO paste/PtNPs, Lac, and BOT (botryosphaeran) electrodes were fabricated by Salamanca-Neto, et al. to discriminate the exceptional and traditional brewed coffee beverages by measuring chlorogenic acid [81].

PhCs were found to contribute to the color, astringency, mouth feel, and oxidative stabilization of wine [169,170]. PhCs in wine give an advantage to hypertensive patients based on their antioxidant nature by decreasing blood pressure [171]. Montereali et al. reported the detection of polyphenols in wine using an amperometric biosensor by immobilizing tyrosinase and the Lac enzyme onto ferrocene-modified graphite screen printed electrodes [172]. A bioelectric tongue was fabricated to detect gallic acid, a trihydroxy group containing compound in white wine, rose wine, and red wine [173]. The differential pulse voltametric technique has been used for assessing gallic acid using a $[Cu_2tpmc](ClO_4)_4$/PVC matrix coated on graphite or carbon rods [174]. However, Almeida et al. [85] used thin polydopamine film on carbon electrodes for the detection of gallic acid from the chestnut shell obtained from industrial waste extract [141].

Phenolic content can also be tested in medicinal plants. Lac-immobilized electrodeposited MWCNT and chitosan film were used to investigate phenols (caffeic acid, gallic acids, chlorogenic acid, and rosmarinic acid) present in the medicinal plants *Salvia officinalis* and *Menthapiperita* [175]. The oil of *Menthapiperita* (Pepperment) contains specific characteristic PhCs and is a well-known medicinal ingredient that is used in tea and herbs. It has strong antioxidant properties and possesses an excellent potential as a repellent of allergies, microbes, and tumors [176]. On the other hand, the herb *Salvia officinalis* (sage) is used as tonic in medicine to improve memory, learning power, and reduce stress [177]. PhCs present in sage (Salvia officinalis) show excellent antioxidant behavior. Rosmarinic acid is the chief phenol present on both herbs [178]. Eremia et al. determined rosmarinic acid from in vitro saliva culture with Lac–Nafion-based biosensors [179].

4.3. Detection of Analytes in Body Fluids

Dopamine is a significant neurotransmitter of the brain, which plays vital roles in the cardiovascular, central nervous, and endocrine systems [180]. Dopamine is found in lower concentrations in patients with dementia, Parkinson's disease, and schizophrenia [181]. The normal concentration of dopamine in the brain is necessary for maintaining accurate blood pressure, learning, enthusiasm, physical activities, and cognizance before releasing dopamine into the brain [37]. Therefore, a rapid clinical and sensitive diagnosis of dopamine is in considerable demand for monitoring Parkinson's disease and depression [182]. A fluorescence-based sensor has been developed using the Lac enzyme and polydopamine on the surface of G-QDs for dopamine detection [78]. Hua, et al. assessed the role of dopamine in human urine samples using a synthesized β-Cyclodextrin inclusion complex as the immobilization matrix onto RGO for the fabrication of a biosensor [183].

Estrogens are responsible for both environmental pollution and harmful effects on humans [184]. For example, 17 β-estradiol is an environmental pollutant that is of paramount concern amongst estrogens. It is also a natural steroid that regulates the female reproductive system by contributing to estrogen [185]. Studies reveal that the release of 17β-estradiol and estriol hormones is around 280–600 μg/day and 6000–10,000 μg/day for pregnant women [156]. The presence of pollutants in food such as meat and milk lead to severe diseases related to fertility and tumor issues in females [186,187]. Biosensors for detecting estrogens from blood and urine are considered superior over other conventional methods as they are simple, rapid, and sensitive [187,188]. Wang et al. developed an electrochemical Lac enzyme poly L-lysine/citric acid-graphene biosensor to detect 17β-estradiol in urine samples [77]. Urine contains conjugated forms of estriol accompanying water soluble sulfates that hydrolyze promptly, leading to regeneration of the free form of estriol, subsequently threatening human and marine lives and giving rise to water pollution [189]. A voltammetry sensor was constructed using cobalt-poly(methionine) modified GCE for determining estriol hormone in urine and pharmaceuticals [190]. Lac functionalized polyvinylpyrrolidone/chitosan/RGO hybrid nanofibers deposited onto fluorine doped tin oxide (FTO) were electrochemically able to diagnose 17α—Ethinylestradiol in synthetic and human urine samples [73] (Figure 6c).

5. Future Scope

It is believed that CNTs and graphitic materials offer faster response time and appreciable sensitivity compared to traditional electrodes at minimal working potentials. However, effort has been made to modify CNMs-based biosensors to have better control on chemical and physical properties. The limitations may be overcome using the integration of metal, conducting polymers, metal chalcogenides, etc., with CNMs, which can enhance the performance of biosensors to address future challenges.

A lot of interest is now been focused on advanced 2D materials such as MXenes (MXe), which was recently explored in the field of biosensors [191]. MXe has been found to be better than graphene-based nanomaterials in terms of improved electron transfer in the heterogeneous phase, low diffusion barrier, and stable dispersibility. This is due to the presence of metal at the center on MX, which does not reduce its conductivity or dispersibility even after functionalization and offers excellent biocompatibility. In this context, MXe based biosensors are progressing toward the development of lab-on-chip biosensing devices involving 5G communication [192].

Future generations demand quick and accurate detection of analytes [193]. New aspects such as the internet of things and machine learning involving artificial intelligence (AI) are being used in the advancement of point-of-care (POC) devices which are emerging more promising technologies as compared to conventional healthcare monitoring [194]. Emphasis on advanced biosensor fabrication must be given to use AI for detecting phenols and for their on-site and online monitoring in order to balance the content of phenol in diverse applications, as discussed in this review article.

6. Conclusions

The reliable detection of PhCs using Lac-based biosensors may result in maintaining permissible polyphenols in water, body fluids, and sustainable management of antioxidant quality. Developing a biosensor for PhCs detection will provide a way to overcome conventional techniques, which are time-consuming and expensive. The application of Lac-based biosensors is mainly focused on the electrochemical transduction method. Amperometric biosensors are among the most reported techniques for detecting PhCs using the Lac enzyme and carbonaceous matrices. The carbon-based matrix has shown enormous potential to support the Lac enzyme by providing enhanced surface area, superior conductivity, stability, and excellent mechanical strength. They have been used in diverse applications to detect polyphenols from wastewater, wine samples, urine, fruits, vegetables, juices, human serums, and beverages. The most significant limitation that remains is the

leaching of the enzyme from the substrate. This can perhaps be improved using suitable immobilization techniques and incorporation of nanomaterials to maintain the enzyme activity. The integration of CNMs with Lac-based biosensors can be crucial in developing POC devices for application in food, pharmaceuticals, and environmental monitoring. The oxidizing characteristics of Lac make it an effective biocatalyst, which can also be utilized for other applications related to the same mechanism. Enzyme modification, the multienzyme approach, AI, and advanced 2D materials should be explored for the development of POC devices.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/bios13030305/s1, Table S1: Sources of the laccase enzyme; Table S2: Doping materials used in carbon matrices [195–199].

Author Contributions: S.V.: Conceptualization, writing—original draft, and investigation. D.T.: validation and visualization. C.M.P.: supervision, validation, writing—reviewing and editing. D.K.: supervision, validation, writing—editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data will be made available on request.

Acknowledgments: Sakshi Verma thanks UGC-India for the Senior Research Fellowship award. Deeksha Thakur acknowledges the Department of Science and Technology (DST), New Delhi, India, for the DST-INSPIRE SRF award. C.M. Pandey acknowledges the Department of Science and Technology (DST), New Delhi, India, for the DST-INSPIRE Faculty award.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

PANI: Polyaniline, NP: nanoparticle, SPE: screen printed electrode, SPCE: screen printed carbon electrode, TPC: total phenolic content, GCE: glassy carbon electrode, CS: chitosan, ZIF: zeoliticimid–azolate framework, BC: bacterial cellulose, Pc: phthalocyanine, CMB: carboxymethyl-botryosphaeran, CBPE: carbon black paste electrode, GNPl: graphenenanoplatelets, BMIMBF4: 1-butyl-3-methylimidazolium tetrafluoroborate, PVP: polyvinylpyrrolidone, NF: nanofiber, GPE: graphene doped carbon paste electrode, Tyr: Tyrosinase, G: graphite, PANABA: poly-aniline-2-aminobenzylamine, OMC-SPE: organized mesoporous carbon-modified carbon screen-printed electrode, cMWCNT: carbonylated multi wall carbon nanotube, MG: magnetic graphene, PEDOT: poly(3,4-ethylenedioxythiophene), PPy: polypyrrole, G-QDs: graphene quantum dots, POXA1b: laccase Pleurotus ostreatus, DPV: Differential Pulse Voltametry, CV: Cyclic voltammetry, SWV: square wave voltammetry, Amp: Amperometry, ChronoAmp: Chronoamperometry

References

1. Villegas, L.G.C.; Mashhadi, N.; Chen, M.; Mukherjee, D.; Taylor, K.E.; Biswas, N. A Short Review of Techniques for Phenol Removal from Wastewater. *Curr. Pollut. Rep.* **2016**, *2*, 157–167. [CrossRef]
2. Anku, W.W.; Mamo, M.; Govender, P. Phenolic Compounds in Water: Sources, Reactivity, Toxicity and Treatment Methods. In *Phenolic Compounds—Natural Sources, Importance and Applications*; InTech Open: London, UK, 2017; pp. 420–443.
3. Kulys, J.; Bratkovskaja, I. Antioxidants determination with laccase. *Talanta* **2007**, *72*, 526–531. [CrossRef]
4. Michałowicz, J.; Duda, W. Phenols—Sources and Toxicity. *Pol. J. Environ. Stud.* **2007**, *16*, 347–362.
5. Rice-Evans, C.; Miller, N.; Paganga, G. Antioxidant properties of phenolic compounds. *Trends Plant Sci.* **1997**, *2*, 152–159. [CrossRef]
6. Meder, D.; Herz, D.M.; Rowe, J.B.; Lehéricy, S.; Siebner, H.R. The role of dopamine in the brain—Lessons learned from Parkinson's disease. *Neuroimage* **2019**, *190*, 79–93. [CrossRef]

7. Ejeian, F.; Etedali, P.; Mansouri-Tehrani, H.-A.; Soozanipour, A.; Low, Z.-X.; Asadnia, M.; Taheri-Kafrani, A.; Razmjou, A. Biosensors for wastewater monitoring: A review. *Biosens. Bioelectron.* **2018**, *118*, 66–79. [CrossRef]
8. Thévenot, D.R.; Toth, K.; Durst, R.A.; Wilson, G.S. Electrochemical biosensors: Recommended definitions and classification. *Biosens. Bioelectron.* **2001**, *16*, 121–131. [CrossRef]
9. Aghapour, A.A.; Moussavi, G.; Yaghmaeian, K. Biological degradation of catechol in wastewater using the sequencing continuous-inflow reactor (SCR). *J. Environ. Health Sci. Eng.* **2013**, *11*, 3. [CrossRef]
10. Groff, T. Bisphenol A: Invisible pollution. *Curr. Opin. Pediatr.* **2010**, *22*, 524–529. [CrossRef]
11. Li, Y.; Trush, M.A. Reactive oxygen-dependent DNA damage resulting from the oxidation of phenolic compounds by a copper-redox cycle mechanism. *Cancer Res.* **1994**, *54*, 1895s–1898s.
12. Kadmi, Y.; Favier, L.; Yehya, T.; Soutrel, I.; Simion, A.I.; Vial, C.; Wolbert, D. Controlling contamination for determination of ultra-trace levels of priority pollutants chlorophenols in environmental water matrices. *Arab. J. Chem.* **2019**, *12*, 2905–2913. [CrossRef]
13. Enguita, F.J.; Leitão, A.L. Hydroquinone: Environmental pollution, toxicity, and microbial answers. *BioMed Res. Int.* **2013**, *2013*, 542168. [CrossRef]
14. Singh, T.; Bhatiya, A.K.; Mishra, P.K.; Srivastava, N. Chapter 11—An effective approach for the degradation of phenolic waste: Phenols and cresols. In *Abatement of Environmental Pollutants*; Singh, P., Kumar, A., Borthakur, A., Eds.; Elsevier: Amsterdam, The Netherlands, 2020; pp. 203–243. [CrossRef]
15. Lecomte, S.; Habauzit, D.; Charlier, T.D.; Pakdel, F. Emerging Estrogenic Pollutants in the Aquatic Environment and Breast Cancer. *Genes* **2017**, *8*, 229. [CrossRef]
16. Lin, T.-W.; Kuo, Y.-M. Exercise benefits brain function: The monoamine connection. *Brain Sci.* **2013**, *3*, 39–53. [CrossRef]
17. Badhani, B.; Sharma, N.; Kakkar, R. Gallic acid: A versatile antioxidant with promising therapeutic and industrial applications. *RSC Adv.* **2015**, *5*, 27540–27557. [CrossRef]
18. Ye, Y.; Guo, H.; Sun, X. Recent progress on cell-based biosensors for analysis of food safety and quality control. *Biosens. Bioelectron.* **2019**, *126*, 389–404. [CrossRef]
19. Lu, X.; Sun, J.; Sun, X. Recent advances in biosensors for the detection of estrogens in the environment and food. *TrAC Trends Anal. Chem.* **2020**, *127*, 115882. [CrossRef]
20. María Isabel, G.-A.; Juan, B.C.; Gustavo, A.P. Nanomaterial-based fluorescent biosensors for monitoring environmental pollutants: A critical review. *Talanta Open* **2020**, *2*, 100006. [CrossRef]
21. Mello, L.D.; Kubota, L.T. Biosensors as a tool for the antioxidant status evaluation. *Talanta* **2007**, *72*, 335–348. [CrossRef]
22. Marco, M.-P.; Barceló, D. Chapter 22 Fundamentals and applications of biosensors for environmental analysis. In *Techniques and Instrumentation in Analytical Chemistry*; BarcelÓ, D., Ed.; Elsevier: Amsterdam, The Netherlands, 2000; Volume 21, pp. 1075–1105.
23. Makarewicz, M.; Drożdż, I.; Tarko, T.; Duda-Chodak, A. The Interactions between Polyphenols and Microorganisms, Especially Gut Microbiota. *Antioxidants* **2021**, *10*, 188. [CrossRef]
24. Zheng, Y.; Yang, C.; Pu, W.; Zhang, J. Carbon nanotube-based DNA biosensor for monitoring phenolic pollutants. *Microchim. Acta* **2009**, *166*, 21–26. [CrossRef]
25. Othman, A.; Karimi, A.; Andreescu, S. Functional nanostructures for enzyme based biosensors: Properties, fabrication and applications. *J. Mater. Chem. B* **2016**, *4*, 7178–7203. [CrossRef]
26. Sekretaryova, A.N.; Volkov, A.V.; Zozoulenko, I.V.; Turner, A.P.F.; Vagin, M.Y.; Eriksson, M. Total phenol analysis of weakly supported water using a laccase-based microband biosensor. *Anal. Chim. Acta* **2016**, *907*, 45–53. [CrossRef]
27. Cheng, H.; Hu, M.; Zhai, Q.; Li, S.; Jiang, Y. Polydopamine tethered CPO/HRP-TiO2 nano-composites with high bio-catalytic activity, stability and reusability: Enzyme-photo bifunctional synergistic catalysis in water treatment. *Chem. Eng. J.* **2018**, *347*, 703–710. [CrossRef]
28. de Matos Morawski, F.; Deon, M.; Nicolodi, S.; Menezes, E.; Costa, T.; Dias, S.; Benvenutti, E.; Arenas, L. Magnetic silica/titania xerogel applied as electrochemical biosensor for catechol and catecholamines. *Electrochim. Acta* **2018**, *264*, 319–328. [CrossRef]
29. Manan, F.A.A.; Hong, W.W.; Abdullah, J.; Yusof, N.A.; Ahmad, I. Nanocrystalline cellulose decorated quantum dots based tyrosinase biosensor for phenol determination. *Mater. Sci. Eng. C* **2019**, *99*, 37–46. [CrossRef]
30. Durán, N.; Rosa, M.A.; D'Annibale, A.; Gianfreda, L. Applications of laccases and tyrosinases (phenoloxidases) immobilized on different supports: A review. *Enzym. Microb. Technol.* **2002**, *31*, 907–931. [CrossRef]
31. Pandey, C. *Biosensors: Fundamentals and Applications*; Smithers Rapra: Shrewsbury, UK, 2017.
32. Scheller, F.W.; Hintsche, R.; Pfeiffer, D.; Schubert, F.; Riedel, K.; Kindervater, R. Biosensors: Fundamentals, applications and trends. *Sens. Actuators B Chem.* **1991**, *4*, 197–206. [CrossRef]
33. Sarkar, A.; Sarkar, K.D.; Amrutha, V.; Dutta, K. Chapter 15—An overview of enzyme-based biosensors for environmental monitoring. In *Tools, Techniques and Protocols for Monitoring Environmental Contaminants*; Kaur Brar, S., Hegde, K., Pachapur, V.L., Eds.; Elsevier: Amsterdam, The Netherlands, 2019; pp. 307–329. [CrossRef]
34. Mulchandani, A. Principles of Enzyme Biosensors. In *Enzyme and Microbial Biosensors: Techniques and Protocols*; Mulchandani, A., Rogers, K.R., Eds.; Humana Press: Totowa, NJ, USA, 1998; pp. 3–14.
35. Das, P.; Barbora, L.; Das, M.; Goswami, P. Highly sensitive and stable laccase based amperometric biosensor developed on nano-composite matrix for detecting pyrocatechol in environmental samples. *Sens. Actuators B Chem.* **2014**, *192*, 737–744. [CrossRef]

36. Kavetskyy, T.; Smutok, O.; Demkiv, O.; Mať ko, I.; Švajdlenková, H.; Šauša, O.; Novák, I.; Berek, D.; Čechová, K.; Pecz, M.; et al. Microporous carbon fibers as electroconductive immobilization matrixes: Effect of their structure on operational parameters of laccase-based amperometric biosensor. *Mater. Sci. Eng. C* **2020**, *109*, 110570. [CrossRef]
37. Tammina, S.K.; Yang, D.; Koppala, S.; Cheng, C.; Yang, Y. Highly photoluminescent N, P doped carbon quantum dots as a fluorescent sensor for the detection of dopamine and temperature. *J. Photochem. Photobiol. B Biol.* **2019**, *194*, 61–70. [CrossRef] [PubMed]
38. Lepore, M.; Portaccio, M. Optical detection of different phenolic compounds by means of a novel biosensor based on sol–gel immobilized laccase. *Biotechnol. Appl. Biochem.* **2017**, *64*, 782–792. [CrossRef]
39. Claus, H. Laccases: Structure, reactions, distribution. *Micron* **2004**, *35*, 93–96. [CrossRef]
40. Dwivedi, U.N.; Singh, P.; Pandey, V.P.; Kumar, A. Structure–function relationship among bacterial, fungal and plant laccases. *J. Mol. Catal. B Enzym.* **2011**, *68*, 117–128. [CrossRef]
41. Jones, S.M.; Solomon, E.I. Electron transfer and reaction mechanism of laccases. *Cell. Mol. Life Sci. CMLS* **2015**, *72*, 869–883. [CrossRef]
42. Chaurasia, P.; Yadav, R.; Yadava, S. A review on mechanism of laccase action. *Res. Rev. Biosci.* **2013**, *7*, 66–71.
43. Rodríguez-Delgado, M.M.; Alemán-Nava, G.S.; Rodríguez-Delgado, J.M.; Dieck-Assad, G.; Martínez-Chapa, S.O.; Barceló, D.; Parra, R. Laccase-based biosensors for detection of phenolic compounds. *TrAC Trends Anal. Chem.* **2015**, *74*, 21–45. [CrossRef]
44. Palanisamy, S.; Ramaraj, S.K.; Chen, S.-M.; Yang, T.C.K.; Yi-Fan, P.; Chen, T.-W.; Velusamy, V.; Selvam, S. A novel Laccase Biosensor based on Laccase immobilized Graphene-Cellulose Microfiber Composite modified Screen-Printed Carbon Electrode for Sensitive Determination of Catechol. *Sci. Rep.* **2017**, *7*, 41214. [CrossRef]
45. Castrovilli, M.C.; Bolognesi, P.; Chiarinelli, J.; Avaldi, L.; Calandra, P.; Antonacci, A.; Scognamiglio, V. The convergence of forefront technologies in the design of laccase-based biosensors—An update. *TrAC Trends Anal. Chem.* **2019**, *119*, 115615. [CrossRef]
46. Zhang, Y.; Ge, J.; Liu, Z. Enhanced Activity of Immobilized or Chemically Modified Enzymes. *ACS Catal.* **2015**, *5*, 4503–4513. [CrossRef]
47. Wu, M.-H.; Lin, M.-C.; Lee, C.-C.; Yu, S.-M.; Wang, A.H.J.; Ho, T.-H.D. Enhancement of laccase activity by pre-incubation with organic solvents. *Sci. Rep.* **2019**, *9*, 9754. [CrossRef]
48. Daronch, N.A.; Kelbert, M.; Pereira, C.S.; de Araújo, P.H.H.; de Oliveira, D. Elucidating the choice for a precise matrix for laccase immobilization: A review. *Chem. Eng. J.* **2020**, *397*, 125506. [CrossRef]
49. Prado Barragán, L.A.; Buenrostro-Figueroa, J.J.; Aguilar González, C.N.; Marañon, I. Chapter 10—Production, Stabilization, and Uses of Enzymes From Fruit and Vegetable Byproducts. In *Biotransformation of Agricultural Waste and By-Products*; Poltronieri, P., D'Urso, O.F., Eds.; Elsevier: Amsterdam, The Netherlands, 2016; pp. 271–286. [CrossRef]
50. Minteer, S.D. Chapter 16—Cell-Free Biotechnologies. In *Biotechnology for Biofuel Production and Optimization*; Eckert, C.A., Trinh, C.T., Eds.; Elsevier: Amsterdam, The Netherlands, 2016; pp. 433–448. [CrossRef]
51. Alarcon-Angeles, G.; Álvarez-Romero, G.A.; Merkoçi, A. Electrochemical Biosensors: Enzyme Kinetics and Role of Nanomaterials. In *Encyclopedia of Interfacial Chemistry*; Wandelt, K., Ed.; Elsevier: Amsterdam, The Netherlands, 2018; pp. 140–155. [CrossRef]
52. Salmon, S.; House, A. Chapter 2—Enzyme-catalyzed Solvents for CO_2 Separation. In *Novel Materials for Carbon Dioxide Mitigation Technology*; Shi, F., Morreale, B., Eds.; Elsevier: Amsterdam, The Netherlands, 2015; pp. 23–86. [CrossRef]
53. Sassolas, A.; Blum, L.J.; Leca-Bouvier, B.D. Immobilization strategies to develop enzymatic biosensors. *Biotechnol. Adv.* **2012**, *30*, 489–511. [CrossRef]
54. Cacicedo, M.L.; Manzo, R.M.; Municoy, S.; Bonazza, H.L.; Islan, G.A.; Desimone, M.; Bellino, M.; Mammarella, E.J.; Castro, G.R. Chapter 7—Immobilized Enzymes and Their Applications. In *Advances in Enzyme Technology*; Singh, R.S., Singhania, R.R., Pandey, A., Larroche, C., Eds.; Elsevier: Amsterdam, The Netherlands, 2019; pp. 169–200. [CrossRef]
55. Fernández-Fernández, M.; Sanromán, M.Á.; Moldes, D. Recent developments and applications of immobilized laccase. *Biotechnol. Adv.* **2013**, *31*, 1808–1825. [CrossRef]
56. Casero, E.; Petit-Domínguez, M.D.; Vázquez, L.; Ramírez-Asperilla, I.; Parra-Alfambra, A.M.; Pariente, F.; Lorenzo, E. Laccase biosensors based on different enzyme immobilization strategies for phenolic compounds determination. *Talanta* **2013**, *115*, 401–408. [CrossRef]
57. Castrovilli, M.C.; Bolognesi, P.; Chiarinelli, J.; Avaldi, L.; Cartoni, A.; Calandra, P.; Tempesta, E.; Giardi, M.T.; Antonacci, A.; Arduini, F.; et al. Electrospray deposition as a smart technique for laccase immobilisation on carbon black-nanomodified screen-printed electrodes. *Biosens. Bioelectron.* **2020**, *163*, 112299. [CrossRef]
58. Mazlan, S.Z.; Lee, Y.H.; Hanifah, S.A. A New Laccase Based Biosensor for Tartrazine. *Sensors* **2017**, *17*, 2859. [CrossRef]
59. Rawal, R.; Chawla, S.; Malik, P.; Pundir, C.S. An amperometric biosensor based on laccase immobilized onto MnO2NPs/cMWCNT/PANI modified Au electrode. *Int. J. Biol. Macromol.* **2012**, *51*, 175–181. [CrossRef]
60. Tortolini, C.; Bollella, P.; Zumpano, R.; Favero, G.; Mazzei, F.; Antiochia, R. Metal Oxide Nanoparticle Based Electrochemical Sensor for Total Antioxidant Capacity (TAC) Detection in Wine Samples. *Biosensors* **2018**, *8*, 108. [CrossRef]
61. Galliker, P.; Hommes, G.; Schlosser, D.; Corvini, P.; Shahgaldian, P. Laccase-modified silica nanoparticles efficiently catalyze the transformation of phenolic compounds. *J. Colloid Interface Sci.* **2010**, *349*, 98–105. [CrossRef]
62. Zdarta, J.; Feliczak-Guzik, A.; Siwińska-Ciesielczyk, K.; Nowak, I.; Jesionowski, T. Mesostructured cellular foam silica materials for laccase immobilization and tetracycline removal: A comprehensive study. *Microporous Mesoporous Mater.* **2020**, *291*, 109688. [CrossRef]

63. Cacciotti, I.; Lombardelli, C.; Benucci, I.; Esti, M. Clay/chitosan biocomposite systems as novel green carriers for covalent immobilization of food enzymes. *J. Mater. Res. Technol.* **2019**, *8*, 3644–3652. [CrossRef]
64. Liang, S.; Wu, X.-L.; Xiong, J.; Zong, M.-H.; Lou, W.-Y. Metal-organic frameworks as novel matrices for efficient enzyme immobilization: An update review. *Coord. Chem. Rev.* **2020**, *406*, 213149. [CrossRef]
65. Calitri, G.; Bollella, P.; Ciogli, L.; Tortolini, C.; Mazzei, F.; Antiochia, R.; Favero, G. Evaluation of different storage processes of passion fruit (*Passiflora edulis* Sims) using a new dual biosensor platform based on a conducting polymer. *Microchem. J.* **2020**, *154*, 104573. [CrossRef]
66. Yaropolov, A.I.; Shleev, S.V.; Morozova, O.V.; Zaitseva, E.A.; Marko-Varga, G.; Emneus, J.; Gorton, L. An Amperometric Biosensor Based on Laccase Immobilized in Polymer Matrices for Determining Phenolic Compounds. *J. Anal. Chem.* **2005**, *60*, 553–557. [CrossRef]
67. Dai, Z.; Guo, M.Q.; Wang, X.J.; Wang, H.F.; Chen, W.Y. Development of Amperometric Laccase Biosensor through Immobilizing Enzyme in Magnesium-Containing Mesoporous Silica Sieve (Mg-MCM-41)/Polyvinyl Alcohol Matrix. *J. Nanomater.* **2014**, *2014*, 458245. [CrossRef]
68. Li, Y.; Zhang, L.; Li, M.; Pan, Z.; Li, D. A disposable biosensor based on immobilization of laccase with silica spheres on the MWCNTs-doped screen-printed electrode. *Chem. Cent. J.* **2012**, *6*, 103. [CrossRef]
69. Kucherenko, I.S.; Soldatkin, O.O.; Kucherenko, D.Y.; Soldatkina, O.V.; Dzyadevych, S.V. Advances in nanomaterial application in enzyme-based electrochemical biosensors: A review. *Nanoscale Adv.* **2019**, *1*, 4560–4577. [CrossRef]
70. Tvorynska, S.; Barek, J.; Josypcuk, B. Influence of different covalent immobilization protocols on electroanalytical performance of laccase-based biosensors. *Bioelectrochemistry* **2022**, *148*, 108223. [CrossRef]
71. Lou, C.; Jing, T.; Zhou, J.; Tian, J.; Zheng, Y.; Wang, C.; Zhao, Z.; Lin, J.; Liu, H.; Zhao, C.; et al. Laccase immobilized polyaniline/magnetic graphene composite electrode for detecting hydroquinone. *Int. J. Biol. Macromol.* **2020**, *149*, 1130–1138. [CrossRef]
72. Zrinski, I.; Pungjunun, K.; Martinez, S.; Zavašnik, J.; Stanković, D.; Kalcher, K.; Mehmeti, E. Evaluation of phenolic antioxidant capacity in beverages based on laccase immobilized on screen-printed carbon electrode modified with graphene nanoplatelets and gold nanoparticles. *Microchem. J.* **2020**, *152*, 104282. [CrossRef]
73. Pavinatto, A.; Mercante, L.A.; Facure, M.H.M.; Pena, R.B.; Sanfelice, R.C.; Mattoso, L.H.C.; Correa, D.S. Ultrasensitive biosensor based on polyvinylpyrrolidone/chitosan/reduced graphene oxide electrospun nanofibers for 17α—Ethinylestradiol electrochemical detection. *Appl. Surf. Sci.* **2018**, *458*, 431–437. [CrossRef]
74. Povedano, E.; Cincotto, F.H.; Parrado, C.; Díez, P.; Sánchez, A.; Canevari, T.C.; Machado, S.A.S.; Pingarrón, J.M.; Villalonga, R. Decoration of reduced graphene oxide with rhodium nanoparticles for the design of a sensitive electrochemical enzyme biosensor for 17β-estradiol. *Biosens. Bioelectron.* **2017**, *89*, 343–351. [CrossRef]
75. Maleki, N.; Kashanian, S.; Maleki, E.; Nazari, M. A Novel Enzyme Based Biosensor for Catechol Detection in Water Samples Using Artificial Neural Network. *Biochem. Eng. J.* **2017**, *128*, 1–11. [CrossRef]
76. Vlamidis, Y.; Gualandi, I.; Tonelli, D. Amperometric biosensors based on reduced GO and MWCNTs composite for polyphenols detection in fruit juices. *J. Electroanal. Chem.* **2017**, *799*, 285–292. [CrossRef]
77. Wang, A.; Ding, Y.; Li, L.; Duan, D.; Mei, Q.; Zhuang, Q.; Cui, S.; He, X. A novel electrochemical enzyme biosensor for detection of 17β-estradiol by mediated electron-transfer system. *Talanta* **2019**, *192*, 478–485. [CrossRef]
78. Baluta, S.; Malecha, K.; Zając, D.; Sołoducho, J.; Cabaj, J. Dopamine sensing with fluorescence strategy based on low temperature co-fired ceramic technology modified with conducting polymers. *Sens. Actuators B Chem.* **2017**, *252*, 803–812. [CrossRef]
79. Othman, A.M.; Wollenberger, U. Amperometric biosensor based on coupling aminated laccase to functionalized carbon nanotubes for phenolics detection. *Int. J. Biol. Macromol.* **2020**, *153*, 855–864. [CrossRef]
80. Li, X.; Li, D.; Zhang, Y.; Lv, P.; Feng, Q.; Wei, Q. Encapsulation of enzyme by metal-organic framework for single-enzymatic biofuel cell-based self-powered biosensor. *Nano Energy* **2020**, *68*, 104308. [CrossRef]
81. Salamanca-Neto, C.A.R.; Marcheafave, G.G.; Scremin, J.; Barbosa, E.C.M.; Camargo, P.H.C.; Dekker, R.F.H.; Scarminio, I.S.; Barbosa-Dekker, A.M.; Sartori, E.R. Chemometric-assisted construction of a biosensing device to measure chlorogenic acid content in brewed coffee beverages to discriminate quality. *Food Chem.* **2020**, *315*, 126306. [CrossRef]
82. Liu, T.; Zhao, Q.; Xie, Y.; Jiang, D.; Chu, Z.; Jin, W. In situ fabrication of aloe-like Au–ZnO micro/nanoarrays for ultrasensitive biosensing of catechol. *Biosens. Bioelectron.* **2020**, *156*, 112145. [CrossRef]
83. Mattos, G.J.; Moraes, J.T.; Barbosa, E.C.M.; Camargo, P.H.C.; Dekker, R.F.H.; Barbosa-Dekker, A.M.; Sartori, E.R. Laccase stabilized on β-D-glucan films on the surface of carbon black/gold nanoparticles: A new platform for electrochemical biosensing. *Bioelectrochemistry* **2019**, *129*, 116–123. [CrossRef]
84. Liu, L.; Anwar, S.; Ding, H.; Xu, M.; Yin, Q.; Xiao, Y.; Yang, X.; Yan, M.; Bi, H. Electrochemical sensor based on F,N-doped carbon dots decorated laccase for detection of catechol. *J. Electroanal. Chem.* **2019**, *840*, 84–92. [CrossRef]
85. Almeida, L.C.; Correia, R.D.; Squillaci, G.; Morana, A.; La Cara, F.; Correia, J.P.; Viana, A.S. Electrochemical deposition of bio-inspired laccase-polydopamine films for phenolic sensors. *Electrochim. Acta* **2019**, *319*, 462–471. [CrossRef]
86. Apetrei, R.-M.; Cârâc, G.; Bahrim, G.; Camurlu, P. Utilization of enzyme extract self-encapsulated within polypyrrole in sensitive detection of catechol. *Enzym. Microb. Technol.* **2019**, *128*, 34–39. [CrossRef]
87. Fu, Y.; An, Q.; Ni, R.; Zhang, Y.; Li, Y.; Ke, H. Preparation of Polyaniline-Encapsulated Carbon/Copper Composite Nanofibers for Detection of Polyphenol Pollutant. *Colloids Surf. A Physicochem. Eng. Asp.* **2018**, *559*, 289–296. [CrossRef]

88. Bravo, I.; Prata, M.; Torrinha, Á.; Delerue-Matos, C.; Lorenzo, E.; Morais, S. Laccase bioconjugate and multi-walled carbon nanotubes-based biosensor for bisphenol A analysis. *Bioelectrochemistry* **2022**, *144*, 108033. [CrossRef]
89. Zhao, K.; Veksha, A.; Ge, L.; Lisak, G. Near real-time analysis of para-cresol in wastewater with a laccase-carbon nanotube-based biosensor. *Chemosphere* **2021**, *269*, 128699. [CrossRef]
90. Chandran, M.; Aswathy, E.; Shamna, I.; Vinoba, M.; Kottappara, R.; Bhagiyalakshmi, M. Laccase immobilized on Au confined MXene based electrode for electrochemical detection of catechol. *Mater. Today Proc.* **2021**, *46*, 3136–3143. [CrossRef]
91. Sorrentino, I.; Stanzione, I.; Piscitelli, A.; Giardina, P.; Le Goff, A. Carbon-nanotube-supported POXA1b laccase and its hydrophobin chimera for oxygen reduction and picomolar phenol biosensing. *Biosens. Bioelectron. X* **2021**, *8*, 100074. [CrossRef]
92. Li, D.; Cheng, Y.; Zuo, H.; Zhang, W.; Pan, G.; Fu, Y.; Wei, Q. Dual-functional biocatalytic membrane containing laccase-embedded metal-organic frameworks for detection and degradation of phenolic pollutant. *J. Colloid Interface Sci.* **2021**, *603*, 771–782. [CrossRef]
93. Gomes, A.; Mattos, G.J.; Coldibeli, B.; Dekker, R.F.H.; Barbosa Dekker, A.M.; Sartori, E.R. Covalent attachment of laccase to carboxymethyl-botryosphaeran in aqueous solution for the construction of a voltammetric biosensor to quantify quercetin. *Bioelectrochemistry* **2020**, *135*, 107543. [CrossRef]
94. Dăscălescu, D.; Apetrei, C. Development of a Novel Electrochemical Biosensor Based on Organized Mesoporous Carbon and Laccase for the Detection of Serotonin in Food Supplements. *Chemosensors* **2022**, *10*, 365. [CrossRef]
95. Ma Guadalupe, G.-R.; Mariana, R.-A.; Héctor Eduardo, M.-F. Characterization of nanostructures of TiO2 used as bioreceptor to immobilize laccase enzyme for detection of gallic acid. *Mater. Res. Express* **2022**, *9*, 095005. [CrossRef]
96. Wang, Z.; Wang, Y.; Yang, S.; Xue, L.; Feng, W.; Liu, X.; Li, B.; Yin, M.; Jiao, J.; Chen, Q. Electrochemical sensor based on magnetic nanohybrids of multiple phthalocyanine doped ferrites/CMWCNTs for detection of rosmarinic acid. *Talanta* **2021**, *226*, 122165. [CrossRef]
97. Đurđić, S.; Stanković, V.; Vlahović, F.; Ognjanović, M.; Kalcher, K.; Veličković, T.Ć.; Mutić, J.; Stanković, D.M. Laccase Polyphenolic Biosensor Supported on MnO2@GNP Decorated SPCE: Preparation, Characterization, and Analytical Application. *J. Electrochem. Soc.* **2021**, *168*, 037510. [CrossRef]
98. Wang, J. Study on Enzymatic and Electrochemical Properties of Cellulase Immobilized with Multi-Walled Carbon Nanotubes as Sensor for Catechol. *Int. J. Electrochem. Sci.* **2021**, *16*, 210447. [CrossRef]
99. Livia Alexandra Dinu, G.; Florina, P.; Sevinc, K.; Lucian-Barbu, T.; Andreea Bianca, S.; Irina, K.; Stela, P. Graphene-Gold Nanoparticles Nanozyme-Based Electrochemical Sensor with Enhanced Laccase-Like Activity for Determination of Phenolic Substrates. *J. Electrochem. Soc.* **2021**, *168*, 067523. [CrossRef]
100. Fartas, F.; Abdullah, J.; Yusof, N.; Sulaiman, Y.; Saiman, M.; Mat Zaid, M. Laccase Electrochemical Biosensor Based on Graphene-Gold/ Chitosan Nanocomposite Film for Bisphenol A Detection. *Curr. Anal. Chem.* **2019**, *15*, 570–579. [CrossRef]
101. Kunene, K.; Sabela, M.; Kanchi, S.; Bisetty, K. High Performance Electrochemical Biosensor for Bisphenol A Using Screen Printed Electrodes Modified with Multiwalled Carbon Nanotubes Functionalized with Silver-Doped Zinc Oxide. *Waste Biomass Valorization* **2020**, *11*, 1085–1096. [CrossRef]
102. Fernandes, P.M.V.; Campiña, J.M.; Silva, A.F. A layered nanocomposite of laccase, chitosan, and Fe3O4 nanoparticles-reduced graphene oxide for the nanomolar electrochemical detection of bisphenol A. *Microchim. Acta* **2020**, *187*, 262. [CrossRef]
103. Peng, C.; Tan, S.Y.; Li, M.T. Synthesis and Characterization of Iron Phthalocyanine Supported on Graphene Oxide and Catalysis of Adrenaline. *J. Nanosci. Nanotechnol.* **2020**, *20*, 2195–2204. [CrossRef]
104. Garcia, L.F.; da Cunha, C.E.P.; Moreno, E.K.G.; Vieira Thomaz, D.; Lobón, G.S.; Luque, R.; Somerset, V.; de Souza Gil, E. Nanostructured TiO_2 Carbon Paste Based Sensor for Determination of Methyldopa. *Pharmaceuticals* **2018**, *11*, 99. [CrossRef]
105. Baluta, S.; Lesiak, A.; Cabaj, J. Graphene Quantum Dots-based Electrochemical Biosensor for Catecholamine Neurotransmitters Detection. *Electroanalysis* **2018**, *30*, 1781–1790. [CrossRef]
106. Lou, C.; Jing, T.; Tian, J.; Zheng, Y.; Zhang, J.; Dong, M.; Wang, C.; Hou, C.; Fan, J.; Guo, Z. 3-Dimensional graphene/Cu/Fe_3O_4 composites: Immobilized laccase electrodes for detecting bisphenol A. *J. Mater. Res.* **2019**, *34*, 2964–2975. [CrossRef]
107. Albayati, S.A.R.; Kashanian, S.; Nazari, M.; Rezaei, S. Novel fabrication of a laccase biosensor to detect phenolic compounds using a carboxylated multiwalled carbon nanotube on the electropolymerized support. *Bull. Mater. Sci.* **2019**, *42*, 187. [CrossRef]
108. Yashas, S.R.; Sandeep, S.; Shivakumar, B.P.; Swamy, N.K. A matrix of perovskite micro-seeds and polypyrrole nanotubes tethered laccase/graphite biosensor for sensitive quantification of 2,4-dichlorophenol in wastewater. *Anal. Methods* **2019**, *11*, 4511–4519. [CrossRef]
109. Verma, S.; Pandey, C.; Kumar, D. A highly efficient rGO grafted MoS 2 nanocomposite for dye adsorption and electrochemical detection of hydroquinone in wastewater. *New J. Chem.* **2022**, *46*, 21190–21200. [CrossRef]
110. Holban, A.M.; Grumezescu, A.M.; Andronescu, E. Chapter 10—Inorganic nanoarchitectonics designed for drug delivery and anti-infective surfaces. In *Surface Chemistry of Nanobiomaterials*; Grumezescu, A.M., Ed.; William Andrew Publishing: Oxford, UK, 2016; pp. 301–327. [CrossRef]
111. Saifuddin, N.; Raziah, A.Z.; Junizah, A.R. Carbon Nanotubes: A Review on Structure and Their Interaction with Proteins. *Journal of Chemistry* **2013**, *2013*, 676815. [CrossRef]
112. Yu, M.-F.; Files, B.S.; Arepalli, S.; Ruoff, R.S. Tensile Loading of Ropes of Single Wall Carbon Nanotubes and their Mechanical Properties. *Phys. Rev. Lett.* **2000**, *84*, 5552–5555. [CrossRef]

113. Harik, V. Chapter 1—Nanotechnology of Carbon Nanotubes: Sensors, Transistors and Nanocomposites. In *Mechanics of Carbon Nanotubes*; Harik, V., Ed.; Academic Press: Cambridge, MA, USA, 2018; pp. 1–24. [CrossRef]
114. Korri-Youssoufi, H.; Zribi, B.; Miodek, A.; Haghiri-Gosnet, A.-M. Chapter 4—Carbon-Based Nanomaterials for Electrochemical DNA Sensing. In *Nanotechnology and Biosensors*; Nikolelis, D.P., Nikoleli, G.-P., Eds.; Elsevier: Amsterdam, The Netherlands, 2018; pp. 113–150. [CrossRef]
115. Han, B.; Yu, X.; Ou, J. Chapter 2—Compositions of Self-Sensing Concrete. In *Self-Sensing Concrete in Smart Structures*; Han, B., Yu, X., Ou, J., Eds.; Butterworth-Heinemann: Oxford, UK, 2014; pp. 13–43. [CrossRef]
116. Karuppanan, K.K.; Panthalingal, M.K.; Biji, P. Chapter 26—Nanoscale, Catalyst Support Materials for Proton-Exchange Membrane Fuel Cells. In *Handbook of Nanomaterials for Industrial Applications*; Mustansar Hussain, C., Ed.; Elsevier: Amsterdam, The Netherlands, 2018; pp. 468–495. [CrossRef]
117. Mohamed, A. Chapter Eight—Synthesis, Characterization, and Applications Carbon Nanofibers. In *Carbon-Based Nanofillers and Their Rubber Nanocomposites*; Yaragalla, S., Mishra, R., Thomas, S., Kalarikkal, N., Maria, H.J., Eds.; Elsevier: Amsterdam, The Netherland, 2019; pp. 243–257. [CrossRef]
118. Feng, L.; Xie, N.; Zhong, J. Carbon Nanofibers and Their Composites: A Review of Synthesizing, Properties and Applications. *Materials* **2014**, *7*, 3919–3945. [CrossRef]
119. Yazdani, N.; Brown, E. 3—Carbon nanofibers in cement composites: Mechanical reinforcement. In *Innovative Developments of Advanced Multifunctional Nanocomposites in Civil and Structural Engineering*; Loh, K.J., Nagarajaiah, S., Eds.; Woodhead Publishing: Oxford, UK, 2016; pp. 47–58. [CrossRef]
120. Yang, J.; Li, D.; Fu, J.; Huang, F.; Wei, Q. TiO2-CuCNFs based laccase biosensor for enhanced electrocatalysis in hydroquinone detection. *J. Electroanal. Chem.* **2016**, *766*, 16–23. [CrossRef]
121. Das, R.; Bandyopadhyay, R.; Pramanik, P. Carbon quantum dots from natural resource: A review. *Mater. Today Chem.* **2018**, *8*, 96–109. [CrossRef]
122. Feng, H.; Qian, Z. Functional Carbon Quantum Dots: A Versatile Platform for Chemosensing and Biosensing. *Chem. Rec.* **2018**, *18*, 491–505. [CrossRef]
123. Huang, S.; Li, W.; Han, P.; Zhou, X.; Cheng, J.; Wen, H.; Xue, W. Carbon quantum dots: Synthesis, properties, and sensing applications as a potential clinical analytical method. *Anal. Methods* **2019**, *11*, 2240–2258. [CrossRef]
124. Wang, Y.; Zhu, Y.; Yu, S.; Jiang, C. Fluorescent carbon dots: Rational synthesis, tunable optical properties and analytical applications. *RSC Adv.* **2017**, *7*, 40973–40989. [CrossRef]
125. Chung, D. Review Graphite. *J. Mater. Sci.* **2002**, *37*, 1475–1489. [CrossRef]
126. Cetó, X.; Capdevila, J.; Mínguez, S.; del Valle, M. Voltammetric BioElectronic Tongue for the analysis of phenolic compounds in rosé cava wines. *Food Res. Int.* **2014**, *55*, 455–461. [CrossRef]
127. Ibarra-Escutia, P.; Gómez, J.J.; Calas-Blanchard, C.; Marty, J.L.; Ramírez-Silva, M.T. Amperometric biosensor based on a high resolution photopolymer deposited onto a screen-printed electrode for phenolic compounds monitoring in tea infusions. *Talanta* **2010**, *81*, 1636–1642. [CrossRef]
128. Krishnan, S.K.; Singh, E.; Singh, P.; Meyyappan, M.; Nalwa, H.S. A review on graphene-based nanocomposites for electrochemical and fluorescent biosensors. *RSC Adv.* **2019**, *9*, 8778–8881. [CrossRef]
129. Peña-Bahamonde, J.; Nguyen, H.N.; Fanourakis, S.K.; Rodrigues, D.F. Recent advances in graphene-based biosensor technology with applications in life sciences. *J. Nanobiotechnol.* **2018**, *16*, 75. [CrossRef]
130. Zhang, J.; Zhang, F.; Yang, H.; Huang, X.; Liu, H.; Zhang, J.; Guo, S. Graphene oxide as a matrix for enzyme immobilization. *Langmuir ACS J. Surf. Colloids* **2010**, *26*, 6083–6085. [CrossRef]
131. Bolibok, P.; Wiśniewski, M.; Roszek, K.; Terzyk, A.P. Controlling enzymatic activity by immobilization on graphene oxide. *Sci. Nat.* **2017**, *104*, 36. [CrossRef]
132. Liang, Y.; Yu, L.; Yang, R.; Li, X.; Qu, L.; Li, J. High sensitive and selective graphene oxide/molecularly imprinted polymer electrochemical sensor for 2,4-dichlorophenol in water. *Sens. Actuators B Chem.* **2017**, *240*, 1330–1335. [CrossRef]
133. Lu, X.; Miodek, A.; Munief, W.-M.; Jolly, P.; Pachauri, V.; Chen, X.; Estrela, P.; Ingebrandt, S. Reduced graphene-oxide transducers for biosensing applications beyond the Debye-screening limit. *Biosens. Bioelectron.* **2019**, *130*, 352–359. [CrossRef]
134. Mei, L.-P.; Feng, J.-J.; Wu, L.; Zhou, J.-Y.; Chen, J.-R.; Wang, A.-J. Novel phenol biosensor based on laccase immobilized on reduced graphene oxide supported palladium–copper alloyed nanocages. *Biosens. Bioelectron.* **2015**, *74*, 347–352. [CrossRef]
135. Cincotto, F.H.; Canevari, T.C.; Machado, S.A.S.; Sánchez, A.; Barrio, M.A.R.; Villalonga, R.; Pingarrón, J.M. Reduced graphene oxide-Sb2O5 hybrid nanomaterial for the design of a laccase-based amperometric biosensor for estriol. *Electrochim. Acta* **2015**, *174*, 332–339. [CrossRef]
136. Eremia, S.A.V.; Vasilescu, I.; Radoi, A.; Litescu, S.-C.; Radu, G.-L. Disposable biosensor based on platinum nanoparticles-reduced graphene oxide-laccase biocomposite for the determination of total polyphenolic content. *Talanta* **2013**, *110*, 164–170. [CrossRef]
137. Ray, S.C. Chapter 1—Application and Uses of Graphene. In *Applications of Graphene and Graphene-Oxide Based Nanomaterials*; Ray, S.C., Ed.; William Andrew Publishing: Oxford, UK, 2015; pp. 1–38. [CrossRef]
138. Sun, H.; Wu, L.; Gao, N.; Ren, J.; Qu, X. Improvement of Photoluminescence of Graphene Quantum Dots with a Biocompatible Photochemical Reduction Pathway and Its Bioimaging Application. *ACS Appl. Mater. Interfaces* **2013**, *5*, 1174–1179. [CrossRef]
139. Tian, P.; Tang, L.; Teng, K.S.; Lau, S.P. Graphene quantum dots from chemistry to applications. *Mater. Today Chem.* **2018**, *10*, 221–258. [CrossRef]

140. Ramezani, M.; Alibolandi, M.; Nejabat, M.; Charbgoo, F.; Taghdisi, S.M.; Abnous, K. Chapter 6—Graphene-Based Hybrid Nanomaterials for Biomedical Applications. In *Biomedical Applications of Graphene and 2D Nanomaterials*; Nurunnabi, M., McCarthy, J.R., Eds.; Elsevier: Amsterdam, The Netherlands, 2019; pp. 119–141. [CrossRef]
141. Vasilescu, I.; Eremia, S.A.; Kusko, M.; Radoi, A.; Vasile, E.; Radu, G.L. Molybdenum disulphide and graphene quantum dots as electrode modifiers for laccase biosensor. *Biosens. Bioelectron.* **2016**, *75*, 232–237. [CrossRef]
142. Preis, S.; Kamenev, S.; Kallas, J.; Munter, R. Advanced Oxidation Processes Against Phenolic Compounds In Wastewater Treatment. *Ozone Sci. Eng.* **1995**, *17*, 399–418. [CrossRef]
143. Mohamed, A.; Yousef, S.; Nasser, W.S.; Osman, T.A.; Knebel, A.; Sánchez, E.P.V.; Hashem, T. Rapid photocatalytic degradation of phenol from water using composite nanofibers under UV. *Environ. Sci. Eur.* **2020**, *32*, 160. [CrossRef]
144. Upan, J.; Reanpang, P.; Chailapakul, O.; Jakmunee, J. Flow injection amperometric sensor with a carbon nanotube modified screen printed electrode for determination of hydroquinone. *Talanta* **2016**, *146*, 766–771. [CrossRef]
145. Lennard, L. 4.21—Methyltransferases. In *Comprehensive Toxicology*, 2nd ed.; McQueen, C.A., Ed.; Elsevier: Amsterdam, The Netherlands, 2010; pp. 435–457. [CrossRef]
146. Prakash, D.; Pandey, J.; Tiwary, B.N.; Jain, R.K. A process optimization for bio-catalytic production of substituted catechols (3-nitrocatechol and 3-methylcatechol. *BMC Biotechnol.* **2010**, *10*, 49. [CrossRef]
147. Kong, Y.; Xu, Y.; Mao, H.; Yao, C.; Ding, X. Expanded graphite modified with intercalated montmorillonite for the electrochemical determination of catechol. *J. Electroanal. Chem.* **2012**, *669*, 1–5. [CrossRef]
148. Zhou, X.-H.; Liu, L.-H.; Bai, X.; Shi, H.-C. A reduced graphene oxide based biosensor for high-sensitive detection of phenols in water samples. *Sens. Actuators B Chem.* **2013**, *181*, 661–667. [CrossRef]
149. Agrawal, V.; Ghaznavi, S.A.; Paschke, R. Environmental Goitrogens ☆. In *Encyclopedia of Endocrine Diseases*, 2nd ed.; Huhtaniemi, I., Martini, L., Eds.; Academic Press: Oxford, UK, 2018; pp. 506–511. [CrossRef]
150. Miyagawa, S.; Sato, T.; Iguchi, T. Subchapter 101C—Bisphenol A. In *Handbook of Hormones*; Takei, Y., Ando, H., Tsutsui, K., Eds.; Academic Press: San Diego, CA, USA, 2016; pp. 577–578. [CrossRef]
151. Jalalvand, A.R.; Haseli, A.; Farzadfar, F.; Goicoechea, H.C. Fabrication of a novel biosensor for biosensing of bisphenol A and detection of its damage to DNA. *Talanta* **2019**, *201*, 350–357. [CrossRef]
152. Huang, N.; Liu, M.; Li, H.; Zhang, Y.; Yao, S. Synergetic signal amplification based on electrochemical reduced graphene oxide-ferrocene derivative hybrid and gold nanoparticles as an ultra-sensitive detection platform for bisphenol A. *Anal. Chim. Acta* **2015**, *853*, 249–257. [CrossRef]
153. Portaccio, M.; Di Tuoro, D.; Arduini, F.; Moscone, D.; Cammarota, M.; Mita, D.G.; Lepore, M. Laccase biosensor based on screen-printed electrode modified with thionine–carbon black nanocomposite, for Bisphenol A detection. *Electrochim. Acta* **2013**, *109*, 340–347. [CrossRef]
154. Laine, M.; Jørgensen, K. Straw compost as an inoculum for the bioremediation of chlorophenol-contaminated soil. *Appl. Environ. Microbiol.* **1996**, *62*, 1507–1513. [CrossRef]
155. Morais, S.; Dias, E.; Pereira, M. Carbamates: Human Exposure and Health Effects. . Academy Press: Cambridge, MA, USA, 2012; pp. 21–38.
156. Nair, P.R.A.; Sujatha, C.H. Organic Pollutants as Endocrine Disruptors: Organometallics, PAHs, Organochlorine, Organophosphate and Carbamate Insecticides, Phthalates, Dioxins, Phytoestrogens, Alkyl Phenols and Bisphenol A. In *Environmental Chemistry for a Sustainable World: Volume 1: Nanotechnology and Health Risk*; Lichtfouse, E., Schwarzbauer, J., Robert, D., Eds.; Springer: Dordrecht, The Netherlands, 2012; pp. 259–309.
157. Oliveira, T.; Barroso, M.F.; Morais, S.; Araújo, M.; Freire, C.; Lima-Neto, P.; Correia, A.; Oliveira, M.; Delerue-Matos, C. Laccase-Prussian blue film-graphene doped carbon paste modified electrode for carbamate pesticides quantification. *Biosens. Bioelectron.* **2013**, *47C*, 292–299. [CrossRef]
158. Oliveira, T.M.B.F.; Fátima Barroso, M.; Morais, S.; de Lima-Neto, P.; Correia, A.N.; Oliveira, M.B.P.P.; Delerue-Matos, C. Biosensor based on multi-walled carbon nanotubes paste electrode modified with laccase for pirimicarb pesticide quantification. *Talanta* **2013**, *106*, 137–143. [CrossRef]
159. Kapeleka, J.A.; Sauli, E.; Sadik, O.; Ndakidemi, P.A. Co-exposure risks of pesticides residues and bacterial contamination in fresh fruits and vegetables under smallholder horticultural production systems in Tanzania. *PLoS ONE* **2020**, *15*, e0235345. [CrossRef]
160. Oliveira, T.M.B.F.; Barroso, M.F.; Morais, S.; Araújo, M.; Freire, C.; de Lima-Neto, P.; Correia, A.N.; Oliveira, M.B.P.P.; Delerue-Matos, C. Sensitive bi-enzymatic biosensor based on polyphenoloxidases–gold nanoparticles–chitosan hybrid film–graphene doped carbon paste electrode for carbamates detection. *Bioelectrochemistry* **2014**, *98*, 20–29. [CrossRef]
161. Ignat, I.; Volf, I.; Popa, V.I. A critical review of methods for characterisation of polyphenolic compounds in fruits and vegetables. *Food Chem.* **2011**, *126*, 1821–1835. [CrossRef]
162. Haminiuk, C.W.I.; Maciel, G.M.; Plata-Oviedo, M.S.V.; Peralta, R.M. Phenolic compounds in fruits—An overview. *Int. J. Food Sci. Technol.* **2012**, *47*, 2023–2044. [CrossRef]
163. de Macêdo, I.Y.L.; Garcia, L.F.; Oliveira Neto, J.R.; de Siqueira Leite, K.C.; Ferreira, V.S.; Ghedini, P.C.; de Souza Gil, E. Electroanalytical tools for antioxidant evaluation of red fruits dry extracts. *Food Chem.* **2017**, *217*, 326–331. [CrossRef]
164. Bogdanov, S.; Jurendic, T.; Sieber, R.; Gallmann, P. Honey for Nutrition and Health: A Review. *J. Am. Coll. Nutr.* **2008**, *27*, 677–689. [CrossRef]

165. de Oliveira Neto, J.R.; Rezende, S.G.; Lobón, G.S.; Garcia, T.A.; Macedo, I.Y.L.; Garcia, L.F.; Alves, V.F.; Torres, I.M.S.; Santiago, M.F.; Schmidt, F.; et al. Electroanalysis and laccase-based biosensor on the determination of phenolic content and antioxidant power of honey samples. *Food Chem.* **2017**, *237*, 1118–1123. [CrossRef]
166. Rawal, R.; Chawla, S.; Pundir, C.S. Polyphenol biosensor based on laccase immobilized onto silver nanoparticles/multiwalled carbon nanotube/polyaniline gold electrode. *Anal. Biochem.* **2011**, *419*, 196–204. [CrossRef]
167. Verrastro, M.; Cicco, N.; Crispo, F.; Morone, A.; Dinescu, M.; Dumitru, M.; Favati, F.; Centonze, D. Amperometric biosensor based on Laccase immobilized onto a screen-printed electrode by Matrix Assisted Pulsed Laser Evaporation. *Talanta* **2016**, *154*, 438–445. [CrossRef]
168. Bae, J.-H.; Park, J.-H.; Im, S.-S.; Song, D.-K. Coffee and health. *Integr. Med. Res.* **2014**, *3*, 189–191. [CrossRef]
169. Cetó, X.; Gutiérrez, J.M.; Gutiérrez, M.; Céspedes, F.; Capdevila, J.; Mínguez, S.; Jiménez-Jorquera, C.; del Valle, M. Determination of total polyphenol index in wines employing a voltammetric electronic tongue. *Anal. Chim. Acta* **2012**, *732*, 172–179. [CrossRef]
170. Kanner, J.; Frankel, E.; Granit, R.; German, B.; Kinsella, J.E. Natural antioxidants in grapes and wines. *J. Agric. Food Chem.* **1994**, *42*, 64–69. [CrossRef]
171. Snopek, L.; Mlcek, J.; Sochorova, L.; Baron, M.; Hlavacova, I.; Jurikova, T.; Kizek, R.; Sedlackova, E.; Sochor, J. Contribution of Red Wine Consumption to Human Health Protection. *Molecules* **2018**, *23*, 1684. [CrossRef]
172. Montereali, M.R.; Seta, L.D.; Vastarella, W.; Pilloton, R. A disposable Laccase–Tyrosinase based biosensor for amperometric detection of phenolic compounds in must and wine. *J. Mol. Catal. B Enzym.* **2010**, *64*, 189–194. [CrossRef]
173. Cetó, X.; Céspedes, F.; del Valle, M. BioElectronic Tongue for the quantification of total polyphenol content in wine. *Talanta* **2012**, *99*, 544–551. [CrossRef]
174. Petković, B.B.; Stanković, D.; Milčić, M.; Sovilj, S.P.; Manojlović, D. Dinuclear copper(II) octaazamacrocyclic complex in a PVC coated GCE and graphite as a voltammetric sensor for determination of gallic acid and antioxidant capacity of wine samples. *Talanta* **2015**, *132*, 513–519. [CrossRef]
175. Diaconu, M.; Litescu, S.C.; Radu, G.L. Laccase–MWCNT–chitosan biosensor—A new tool for total polyphenolic content evaluation from in vitro cultivated plants. *Sens. Actuators B Chem.* **2010**, *145*, 800–806. [CrossRef]
176. McKay, D.; Blumberg, J. A Review of the bioactivity and potential health benefits of peppermint tea *(Mentha piperita* L.). *Phytotherapy Res.* **2006**, *20*, 619–633. [CrossRef]
177. Lopresti, A.L. Salvia (Sage): A Review of its Potential Cognitive-Enhancing and Protective Effects. *Drugs R D* **2017**, *17*, 53–64. [CrossRef]
178. Wang, M.; Li, J.; Rangarajan, M.; Shao, Y.; LaVoie, E.J.; Huang, T.-C.; Ho, C.-T. Antioxidative Phenolic Compounds from Sage (*Salvia officinalis*). *J. Agric. Food Chem.* **1998**, *46*, 4869–4873. [CrossRef]
179. Eremia, S.A.V.; Radu, G.-L.; Litescu, S.-C. Monitoring of Rosmarinic Acid Accumulation in Sage Cell Cultures using Laccase Biosensor. *Phytochem. Anal.* **2013**, *24*, 53–58. [CrossRef]
180. Ko, J.H.; Strafella, A.P. Dopaminergic neurotransmission in the human brain: New lessons from perturbation and imaging. *Neurosci. A Rev. J. Bringing Neurobiol. Neurol. Psychiatry* **2012**, *18*, 149–168. [CrossRef]
181. Schultz, W. Chapter 2—Electrophysiological Correlates of Reward Processing in Dopamine Neurons. In *Decision Neuroscience*; Dreher, J.-C., Tremblay, L., Eds.; Academic Press: San Diego, CA, USA, 2017; pp. 21–31. [CrossRef]
182. Hou, S.; Kasner, M.L.; Su, S.; Patel, K.; Cuellari, R. Highly Sensitive and Selective Dopamine Biosensor Fabricated with Silanized Graphene. *J. Phys. Chem. C* **2010**, *114*, 14915–14921. [CrossRef]
183. Hua, Z.; Qin, Q.; Bai, X.; Wang, C.; Huang, X. β-Cyclodextrin inclusion complex as the immobilization matrix for laccase in the fabrication of a biosensor for dopamine determination. *Sens. Actuators B Chem.* **2015**, *220*, 1169–1177. [CrossRef]
184. Adeel, M.; Song, X.; Wang, Y.; Francis, D.; Yang, Y. Environmental impact of estrogens on human, animal and plant life: A critical review. *Environ. Int.* **2017**, *99*, 107–119. [CrossRef]
185. Vardanyan, R.S.; Hruby, V.J. 28—Female Sex Hormones. In *Synthesis of Essential Drugs*; Vardanyan, R.S., Hruby, V.J., Eds.; Elsevier: Amsterdam, The Netherlands, 2006; pp. 365–379. [CrossRef]
186. Das, A.; Sangaranarayanan, M.V. A sensitive electrochemical detection of progesterone using tin-nanorods modified glassy carbon electrodes: Voltammetric and computational studies. *Sens. Actuators B Chem.* **2018**, *256*, 775–789. [CrossRef]
187. Dai, Y.; Liu, C.C. Detection of 17 β-Estradiol in Environmental Samples and for Health Care Using a Single-Use, Cost-Effective Biosensor Based on Differential Pulse Voltammetry (DPV). *Biosensors* **2017**, *7*, 15. [CrossRef]
188. Yildirim-Tirgil, N.; Long, F.; Gao, C.; He, M.; Shi, H.; Gu, A. Aptamer-Based Optical Biosensor For Rapid and Sensitive Detection of 17β-Estradiol In Water Samples. *Environ. Sci. Technol.* **2012**, *46*, 3288–3294. [CrossRef]
189. Manjunatha, J.G. Electroanalysis of estriol hormone using electrochemical sensor. *Sens. Bio-Sens. Res.* **2017**, *16*, 79–84. [CrossRef]
190. Gomes, E.S.; Leite, F.R.F.; Ferraz, B.R.L.; Mourão, H.A.J.L.; Malagutti, A.R. Voltammetric sensor based on cobalt-poly(methionine)-modified glassy carbon electrode for determination of estriol hormone in pharmaceuticals and urine. *J. Pharm. Anal.* **2019**, *9*, 347–357. [CrossRef]
191. Khan, R.; Andreescu, S. MXenes-Based Bioanalytical Sensors: Design, Characterization, and Applications. *Sensors* **2020**, *20*, 5434. [CrossRef]
192. Chaudhary, V.; Kaushik, A.; Furukawa, H.; Khosla, A. Review—Towards 5th Generation AI and IoT Driven Sustainable Intelligent Sensors Based on 2D MXenes and Borophene. *ECS Sens. Plus* **2022**, *1*, 013601. [CrossRef]

193. Manickam, P.; Mariappan, S.A.; Murugesan, S.M.; Hansda, S.; Kaushik, A.; Shinde, R.; Thipperudraswamy, S.P. Artificial Intelligence (AI) and Internet of Medical Things (IoMT) Assisted Biomedical Systems for Intelligent Healthcare. *Biosensors* **2022**, *12*, 562. [CrossRef]
194. Chaudhary, V.; Khanna, V.; Ahmed Awan, H.T.; Singh, K.; Khalid, M.; Mishra, Y.K.; Bhansali, S.; Li, C.-Z.; Kaushik, A. Towards hospital-on-chip supported by 2D MXenes-based 5th generation intelligent biosensors. *Biosens. Bioelectron.* **2023**, *220*, 114847. [CrossRef]
195. Kumar, V.; Sonkar, P. Laccases: Sources and Their Environmental Application. *Int. J. Bioassays* **2013**, *2*, 909–911.
196. Viswanath, B.; Rajesh, B.; Janardhan, A.; Kumar, A.P.; Narasimha, G. Fungal Laccases and Their Applications in Bioremediation. *Enzym. Res.* **2014**, *2014*, 163242. [CrossRef]
197. Shraddha; Shekher, R.; Sehgal, S.; Kamthania, M.; Kumar, A. Laccase: Microbial Sources, Production, Purification, and Potential Biotechnological Applications. *Enzym. Res.* **2011**, *2011*, 217861. [CrossRef]
198. Ece, S.; Lambertz, C.; Fischer, R.; Commandeur, U. Heterologous expression of a Streptomyces cyaneus laccase for biomass modification applications. *AMB Express* **2017**, *7*, 86. [CrossRef]
199. Blánquez, A.; Ball, A.S.; González-Pérez, J.A.; Jiménez-Morillo, N.T.; González-Vila, F.; Arias, M.E.; Hernández, M. Laccase SilA from Streptomyces ipomoeae CECT 3341, a key enzyme for the degradation of lignin from agricultural residues? *PLoS ONE* **2017**, *12*, e0187649. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Environmental Science and Pollution Research
https://doi.org/10.1007/s11356-023-25572-x

RESEARCH ARTICLE

Role of nuclear energy in carbon mitigation to achieve United Nations net zero carbon emission: evidence from Fourier bootstrap Toda-Yamamoto

Akanksha Singh[1] · Shiv Lal[3] · Nand Kumar[1] · Rajan Yadav[2] · Shweta Kumari[1]

Received: 14 September 2022 / Accepted: 23 January 2023
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2023

Abstract

In this communication, the time series data of three major countries USA, France, and Japan from 1965 to 2020 for CO_2 emission, GDP, and nuclear energy (NE) are evaluated. It also analyzed and validated the EKC hypothesis while using nuclear energy for electricity generation. Fourier ARDL is used to investigate the hypothesis criteria, and the Fourier bootstrap Toda-Yamamoto (FBTY) causality test is used for causal linkage between the variables as well as the wavelet coherence; it is also presented the time and frequency dependency of the variables. The CO_2 mitigation by using the NE is also assessed for all three countries and assessed that the France, Japan, and USA mitigated the CO_2 per year is 0.0463 million metric ton (MMT), 0.0239 and 0.0728 MMT per year respectively. Similar to that the SO_2 is reduced by using the NE is 24.322, 43.527, and 132.592 MMT/year, and NO*x* is reduced by approximately 0.2847, 0.147, and 0.4478 MMT/year by France, Japan, and USA respectively by applying the NE for power generation. The evidence of the EKC, Fourier bootstrap and Toda-Yamamoto clarifies the important role of nuclear energy in terms of carbon mitigation to achieve UN net zero carbon emission by 2050. Hence, in order to meet the UN target of net zero carbon emission by 2050, the USA and Japan should increase the production of nuclear energy as France meets its 74.1% energy demand through NE by validating the EKC hypothesis; on the other hand, all the three countries should increase the production of tidal energy due to their geographical location as tides are much more predictable than wind and sun keeping in consideration to the expenses incurred and a full proof plan for disposing NE residuals in a safe place as NE residuals are highly radioactive and contains traces of thorium and uranium.

Keywords Nuclear energy · EKC hypothesis · Bootstrap ARDL Carbon-dioxide mitigation · Fourier bootstrap Toda-Yamamoto · Wavelet coherence · Carbon neutrality

Nomenclature

α	Deterministic term
β	Estimated parameter
γ	Slope coefficient
ϑ	Coefficient
ε	Error terms
ε_t	Stationary disturbance with variance
ω	Estimator
C	Cointegration
k	Single frequency selected for the approximation
n	Time index
R	Real part of complex variables
S	Smoothing operator
S_t	Partial sum of OLS residuals
s	Scale
t	Time
u_t	Stationary error term with variance
W	Wavelet
y	Casualty

Acronyms

GDP	Gross domestic product
ARDL	Autoregressive distributed lag
FBTY	Fourier bootstrap Toda-Yamamoto
EKC	Environmental Kuznets curve

Responsible Editor: Roula Inglesi-Lotz

✉ Shweta Kumari
singhkshweta@gmail.com

[1] Department of Humanities, Delhi Technological University, New Delhi, India 110042

[2] Delhi School of Management, Delhi Technological University, New Delhi, India 110042

[3] Department of Mechanical Engineering, Rajasthan Technical University, Kota, India 324010

Published online: 30 January 2023

LPG	Liquefied petroleum gas
BRICS	Brazil, Russia, India, China, and South Africa
OECD	Organization for Economic Co-operation and Development
GCC	Gulf Cooperation Council
NE	Nuclear energy
KPSS	Kwiatkowski–Phillips–Schmidt–Shin
VAR	Vector auto regression
OLS	Ordinary least square
SO_2	Sulfur dioxide
CO_2	Carbon dioxide
GLS	Generalized least squares
ECM	Error correction model
CIS	Commonwealth of independent states
FMOLS	Fully modified ordinary least squares
ECM	Error correction model
QLF	Quadratic functional form
VECM	Vector error correction model
DOLS	Dynamic ordinary least square
OLS	Ordinary least squares
EF	Ecological footprints
TYGC	Toda-Yamamoto Granger causality
AMG	Classical algebraic multigrid
CDS	Cross sectional dependence
PMG	Pooled mean group
CCEMG	Common correlated effect mean group
NO*x*	Nitrous oxide
ADF	Augmented Dickey–Fuller test
LM	Lagrange multiplier

Introduction

The industrial development is the main backbone of any country for their economic development and it requires electrical energy, so the energy is the wheel of any economy. Presently, conventional sources of energy like coal and petroleum fuels are the main sources of energy for a century and due to their negative impact on the environment, world's researchers are working on to find out the better alternatives. The whole community of researchers and scientist also wants to develop a technology for achieving net zero carbon emission. The nuclear energy, solar energy, and wind energy may have an important role to achieve the target of net zero carbon emission. It is observed that France meets its 70% of electricity generated through nuclear energy; in China, 19.7% is the total share of nuclear energy, and in Japan, 30% is the total share of nuclear energy in electricity generation. Obviously, the nuclear energy may be one source of the present as well as future energy (Worldbank.org). The annual CO_2 emission of four major power Counties like USA, France, China, and Japan is presented in Fig. 1.

It is observed that the CO_2 emission in China is not reducing even though using the nuclear energy for power generation; it means its power consumption is more than its generation, and demands have been fulfilled via commercial fuels like coal, gas, and petroleum. The addition of nuclear power generation in China is not so long that is why its long-term study is not viable to continue at present. The annual CO_2 emission of France, Japan, and USA has been reduced after using the nuclear power for this region; the authors used only these three countries for the study.

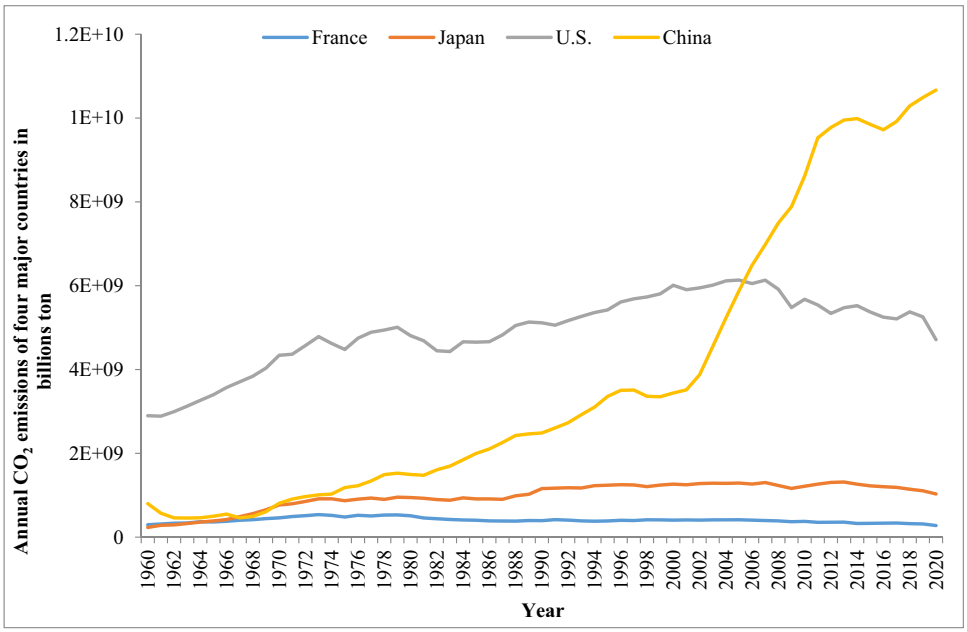


Fig. 1 Annual CO_2 emissions of France, Japan, USA, and China in billions of tons (worldbank.org)

Some of the researchers presented their views in the global meets to recognize pollution prevention measures like the Stockholm conference in 1972, and the Rio conference in 1992. The Kyoto Protocol encouraged and promoted the consumption of renewable energy for limiting GHGs (Bildirici and Gökmenoğlu 2017). The environmental Kuznets curve (EKC) hypothesis is a theory by which the relation between per capita pollutant like CO_2 and per capita gross development product (GDP) can be assessed, and it has an inverted U-shape. Grossman and Krueger (1991) developed the EKC hypothesis and checked the strength of the analysis. The EKC depicted an inverted U-shape curve relation between economic growth and transition in the emission levels (Heidari et al. 2015, Jebli et al. 2016, Leal and Marques 2020, Rafiq et al. 2016, Dong et al. 2017, Salim et al. 2019, Shahbaz et al. 2019, Koc and Bulus 2020, Murshed 2020a).

The EKC showed no observable effect between economic growth and renewable energy consumption for the top six hydropower countries, the five South East Asian countries, the Sub-Saharan African countries, the top 10 innovative countries, and the other 34 listed countries where the power generation by the conventional fuel does not influence the environment (Kisswani et al. 2018; Gorus and Aydin 2019; Pata and Aydin 2020; Jin and Kim 2020).

Furthermore, the causal relationship between energy, gross domestic products (GDP), and carbon dioxide (CO_2) emissions has an important role in econometric analysis. Some studies also presented the causal unidirectional link between energy use and GDP (Yang and Zhao 2014; Adebayo 2020; Adebayo et al. 2021). Few studies also presented the bidirectional link between GDP and energy (Mutascu 2016; Khobai and Roux 2017). Faisal et al. (2016) reflected no causal link between the variables used in the EKC hypothesis. Many studies were conducted by the researchers to check the causal relationship between GDP and CO_2 emissions (Jafari et al. 2015; Wang et al. 2019; Aydoğan and Vardar 2020; Kirikkaleli et al. 2020; Kirikkaleli and Adebayo 2021). On other side, some studies have been carried out and evaluated the unidirectional and bidirectional link between CO_2 emissions and GDP (Al-Mulali 2011; Wu et al. 2018, Adebayo et al. 2020; Gao and Zhang 2021). And some studies also presented their views towards the positive relation between GDP and CO_2 emissions (Pablo-Romero et al. 2016; Ahmed et al. 2019, 2020; Wasti and Zaidi 2020; Ahmed et al. 2020; Odugbesan and Adebayo 2020; Teng et al. 2021). Aye and Edoja (2017) studied 31 developing countries and show a negative relationship between GDP and CO_2 emissions, but those countries have very less clean energy production, whereas Iwata et al. (2010) and Ozcan et al.(2020) studied the validity of the EKC hypothesis while using nuclear energy.

Developing the better alternatives to conventional energy sources is at its peak level, and the validation of clean energy or renewable energy is being tested by the EKC hypothesis. The EKC hypothesis states that a co-integration relation exists between economic growth and environmental pollution. The validation of the EKC hypothesis has been carried out while using clean sources of energy; in this regard, liquefied petroleum gas (LPG) can also be approved as a transitional fuel to reducing the GHGs in the South Asian countries (Albela et al. 2012; Raslavicius et al. 2014; Ekbordin and Tami 2016; Murshed 2018; Murshed 2020b). Hydropower is also advantageous over solar energy and can be adopted as a good source of clean energy, and it provides a steady power supply as its another benefit (Solarin and Ozturk 2015). The review progress of various researchers is shown in Table 1 with the data taken for the period, studied countries, variables used in methodology, and the findings.

Adebayo (2022a) assessed the environmental sustainability and ecological footprint of Turkey linked with oil and hydro energy consumption. Adebayo (2022b) also studied the environmental consequences of fossil fuel in Spain by using the wavelet-based granger casualty approach and also find the effect of renewable energy consumption. Alam et al. (2022) presented the pathways for securing environmentally sustainable economic growth via efficient use of energy with the application of bootstrapped ARDL approach. Alam and Shabbir (2022) studied the economic growth and energy consumption for Bahrain with the application of the VECM and ARDL approach. Singh et al. (2022) analyzed India's logistic performance index using ARDL; it reveals that India is far behind in reaching the environmental sustainability of LPI. Alvarado et al. (2022) studied the impact of the informal economy on the environmental and ecological footprint via the globalization and urban concentration. Farooq et al. (2022) presented the effect of globalization on CO_2 emission via EKC. Hamid et al. (2022a, b) added their contribution to the field of decarburization pathways included the role of foreign direct investment, democracy, governance, renewable energy, and economic growth in Oman.

The role of energy efficiency, renewable energy production, and finance inclusion in the field of environmental sustainability is very important for the sustainable development of the earth. The similar study for 11 countries is presented by Khan et al. (2022). Muhammad et al. (2022) have been exploring the aggregate domestic consumption and CO_2 emission for environmental sustainability in Pakistan. Murshed et al. (2022a) studied the CO_2 restraining agenda for 11 countries with the implication of sustainable development and Argentina with the role of renewable energy transition and trade globalization (Murshed 2021; Murshed et al. 2022b).

The Kuznet curve hypothesis was applied by Murshed for studying the carbon and ecological trails. It also assessed the effect of LPG on the environment with the same hypothesis as the ARDL analysis in South Asia (Murshed 2022).

Table 1 Summary of various studies for EKC, ARDL, etc. methodologies used by researchers

S. no	Study	Data	Countries	Variables	Methodology	Findings
1	Apergis and Payne (2010)	1992–2004	11 CIS countries including Kazakhstan	QLF	FMOLS	IUS means valid EKC hypothesis
2	Shahbaz et al. (2013)	1965–2008	South Africa	CC, EN, FD, TO, CO_2	ARDL, ECM,	CO2 increasing with increasing CC and EN and decreasing with increasing FD
3	Li and Lin (2015)	1971–2010	Low, middle, and high-income countries	*Y*, URB, EN, CO_2	Dynamic panel threshold regression models	In low and middle income countries, when Y, EM, and URB increases means increases CO2 and CO2 increases with increasing Y and EN and decreasing with increasing URB in high income countries
4	Abbasi and Riaz (2016)	1971–2011	Pakistan	*Y*, FD, CO_2	VECM, Granger	CO2 increasing with increasing Y and not found a good relation with FD
5	Al-Mulali et al. (2016)	1980 (1990)–2010	107 countries including Kazakhstan	QLF	DOLS	IUS for the group with Kazakhstan
6	Shahbaz et al. (2016)	1970Q1–2011Q4	Malaysia	*Y*, EN, TO, URB, CO_2	Bayer–Hanck, ARDL bounds testing approach, VECM Granger causality	The CO2 increasing with increasing Y, EN and TO but decreasing with increasing URB
7	Lau et al. (2016)	1965–2010	Malaysia	HEC, *Y*, and CO_2	Granger causality	Y and HEC granger cause CO_2 in the long run, however, HEC granger causes CO_2 in the short run
8	Ozatac et al. (2017)	1960–2013	Turkey	*Y*, EN, TO, URB, FD,CO_2	ARDL, causality	EKC hold GDP → CO_2 (+) EN → CO_2 (+) URB → CO_2 (+)
9	Bekhet et al. (2017)	1980–2011	Gulf countries	*Y*, ENE, FD, CO_2	ARDL, causality	CO_2 → ENE in Saudi Arabia, UAE, and Qatar FD → CO_2 in UAE, Oman, and Kuwait
10	Saidi and Mbarek (2017)	1990–2013	19 emerging economies	*Y*, Y^2, TO, URB, FD, CO_2	Panel regression	EKC is held in these countries because CO2 is decreasing with increasing in FD and URB
11	Bildirici and Gökmenoğlu (2017)	1961–2013	G7 countries	HEC, *Y*, and CO_2	MSVAGC approach	HEC causes Y in overall, and two-way causation in some G7 economies
12	Salahuddin et al. (2018)	1980–2013	Kuwait	*Y*, EC, FDI, CO_2	ARDL, VECM	CO2 is increasing with increasing the Y,EC.FDI
13	Shittu et al. (2018)	1981–2014	Malaysia	URB, DEF, CO_2	Quantile regression and ARDL	CO2 increasing with increasing the DEF and decreasing with increasing URB

Table 1 (continued)

S. no	Study	Data	Countries	Variables	Methodology	Findings
14	Akbota and Baek (2018)	1991–2014	Kazakhstan		ARDL	IUS
15	Ummalla and Samal (2018)	1965–2016	China	CO_2, *Y*, and HEC	ARDL bound test,	EKC is invalid. CO_2 has a positive influence on HEC
16	Pata (2018)	1974–2014	Turkey	CO2, Y, FDI, REN, URB, and HYDEC	ARDL and FMOLS	HYDEC and REN do not affect CO_2; EKC is valid
17	Sinaga (2019)	1978–2016	Malaysia	HEC and *Y*	ARDL	HEC reduces environmental degradation
18	Charfeddine and Kahia (2019)	1980–2012	MENA region	*Y*, RENE, FD, CO_2	Panel vector	CO2 is increasing with increasing Y, REN, and FD
19	Dong et al. (2019)	1960–2012	Developed economies	*Y*, URB, EN, CO_2	Threshold regression model	CO_2 is increasing with increasing REN and URB
20	Jian et al. (2019)	1982–2017	China	*Y*, FD, EN, CO_2	VECM, impulse response	Increasing the EN means increasing CO_2 as well and FD but when increasing Y decreasing the CO_2
21	Koc and Bulus (2020)	1971–2017	South Korea	GDP, GDP^2, ENE, CO_2	ARDL,	*N*-shaped relationship CO2 is increasing with increasing the GDP, ENE and decreasing with increasing the TO and REN
22	Wang et al. (2020)	1990–2014	APEC nations	*Y*, IND, EN, CO_2	Dynamic-unrelated seemingly regression	CO2 is increasing with increasing Y, IND, and URB
23	Bekun et al. (2020)	1971–2015	Nigeria	*Y*, Y^2, FDI, EC, CO_2	ARDL	EKC hold CO2 is increasing with increasing Y, GDP, EC and decreasing with increasing the FDI
24	Nondo and Kahsai (2020)	1970–2016	South Africa	*Y*, URB, ENE, IND	ARDL, VECM	CO2 is increasing with increasing ENE, IND and URB
25	Rahman and Vu (2020)	1960–2015	Australia and Canada	*Y*, EN, URB, CO_2	ARDL, VECM causality	In Australia, CO2 is increasing with increasing *Y* In Canada, CO2 is increasing with increasing Y, and URB
26	Dogan and Inglesi-Lotz (2020)	1980–2014	European countries	*Y*, Y^2, EN, CO_2	Panel OLS, FMOLS	EKC does not hold CO2 is decreasing with increasing GDP
27	Pata and Aydin (2020)	1965–2016	Brazil, China, Canada, India, Norway, and the USA	*Y*, Y^2, EN, HENE, CO_2	Fourier bootstrap ARDL, Fourier TYGC	EKC hypothesis is not valid HENE → GDP in Brazil HENE ↔ GDP in China
28	Leal and Marques (2020)	1990–2014	OECD countries	*Y*, Y^2, TO, CO_2	FMOLS, DOLS, AMG, CSD	EKC hypothesis is not valid CO2 is increasing with increasing OPand REN

Table 1 (continued)

S. no	Study	Data	Countries	Variables	Methodology	Findings
29	Adebayo et al. (2020)	1980–2018	Mint nations	*Y*, TO, URB, EN, CO_2	Pedroni, Westerlund Coint., PMG, CSD, panel causality	GDP ≠ CO_2 TO → CO_2 (−) TO → ENE (+) URB → CO_2 (+)
30	Anser et al. (2020)	1995–2015	G-7 countries	CO_2, ITR, FDI, GEE, HEXP, INEQ, EG	Random effect, panel causality	CO2 is increasing with increasing ITR, GDP, FDI, and GEE
32	Khattak et al. (2020)	1980–2016	BRICS	CO_2, GDP, RE, YPC, YPC^2	CCEMG technique	EKC hold CO2 is increasing with increasing YPC and IN and decreasing the increasing RE
33	Rehman et al. (2020)	2000–2018	Pakistan	CO_2, EC, GDP, PG	GRA model Hurwicz method	CO2 is increasing with increasing GDP, EC, and PG
34	Zmami and Ben-Salha (2020)	1980–2017	GCC countries	CO_2, GDP, EC, URB, TRA, FDI	STIRPAT model, PMG	EKC valid CO2 is increasing with increasing EC and FDI
35	Pata and Kumar (2021)	1980–2016	China and India	FDI, HEC, coal, CO_2, and EFP	ARDL	HEC upsurges CO_2 and EFP for China but has no influence on India
36	Bilgili et al. (2021)	1980–2019	USA	HYDEC and CO_2	CWT approach Continuous Wavelet transformation	In the short run, HYDEC intensifies environmental degradation, however, in long run, it reduces
37	Tiwari et al. (2022)	1971Q1 to 2017Q4 1971Q1 to 2017Q4	Brazil and China	Hydro, EF, urban, and human capital	QARDL	Hydropower mitigates pollution
38	Jahanger et al. (2022)	1990–2016	73 developing countries	Technological innovation, natural resources, economic growth, development of human capital and financial development	PMG-ARDL	Globalization increases the ecological footprint in general

The similar studies have been carried out by the various researchers which presented the role of economic growth, globalization, energy consumption, and renewable energy applications, globalization, carbon emission, nuclear energy utilization, hydro energy, biomass energy, and agricultural productivity on the environmental sustainability in various countries (Nathaniel and Ahmad 2021; Ozturk et al. 2022a, b; Pan and Adebayo 2022; Ramzan and Adebayo 2022; Rehman et al. 2022a; Rehman et al. 2022b; Shi et al. 2022; Wang et al. 2023; Wu et al. 2022; Zhang 2021; Zhang et al. 2022; Zhou et al. 2022).

There is very few literatures available on the validity of the EKC hypothesis while using nuclear energy. The above studies do not take structural breaks into considerations. The smooth changes in causality and co-integrations are important and should be recognized. The time and frequency dependency of the variables is essential for strategic policy formulation that went unnoticed in the above literature.

In this study, the world's three biggest energy-producing countries have been studied for the use of nuclear energy and reducing the atmospheric pollution parameter like CO_2, SO_2, and NO_x. China is the third-highest power-producing country but it has recently started power generation through nuclear energy; therefore, long-term analysis is not possible. That is why the USA, France, and Japan have been taken for this study.

Methodology

The study has been conducted on a time series data from the period 1965 to 2020 for the top three energy-producing countries (USA, France, and Japan) by the nuclear capacity that was analyzed to check the validity of the EKC hypothesis.

EKC hypothesis

The environmental Kuznets curve (EKC) general model (supported the long-term logarithmic form) is adopted for this study which is suggested by Emrah and Salih (2021) and is given by Eq. 1 as follows:

$$lnCO_{2t} = \beta_0 + \beta_1 lnGDP_t + \beta_2 (lnGDP)_t^2 + \beta_3 (lnGDP * lnNE)_t + \beta_4 lnNE_t + \varepsilon_t \quad (1)$$

where β_0, β_1, β_2, β_3, and β_4 are estimated parameters and t is the time index, ε is error terms, and is NE an explanatory variable which can be nuclear energy production. The per capita carbon dioxide (per capita CO_2) is the environmental degradation factor, and per capita GDP is the economic activity factor.

The estimated parameters defined the five different forms of hypothesis based on the test or value limit between the environmental degradation factor (per capita CO_2) and economic activity factor (per capita GDP) which are given below (Yasin and Sehnem 2018).

(i) $\beta_1 = \beta_2 = \beta_3 = 0$ conclude a flat pattern of no relationship between the environmental degradation factor (per capita CO_2) and economic activity factor (per capita *GDP*).

(ii) $\beta_1 < 0$, $\beta_2 > 0$, and $\beta_3 = 0$ conclude in a U-shaped relationship between the environmental degradation factor (per capita CO_2) and economic activity factor (per capita *GDP*).

(iii) $\beta_1 > 0$, $\beta_2 < 0$ and $\beta_3 = 0$, and $\beta_4 > 0$ conclude an inverted U-shaped relationship i.e. EKC relationship between the environmental degradation factor (per capita CO_2) and economic activity factor (per capita *GDP*).

(iv) $\beta_1 > 0$, $\beta_2 < 0$, and $\beta_3 > 0$ conclude in an N-shaped relationship between the environmental degradation factor (per capita CO_2) and economic activity factor (per capita *GDP*).

(v) $\beta_1 < 0$, $\beta_2 > 0$, and $\beta_3 < 0$ conclude in an inverse N-shaped relationship (or EKC postulates) between the environmental degradation factor (per capita CO_2) and economic activity factor (per capita *GDP*).

The data of GDP per capita (current US $) were collected from World Bank 2020 (world bank 2020), and data for carbon dioxide (CO_2) emission in million tonnes and nuclear energy generation in terawatt-hours have been taken from BP stats 2021 (BP stats 2021).

Fourier ADF unit test

To capture the non-linear trends and gradual and sharp structural breaks, the Fourier ADF, Fourier GLS, Fourier LM, and Fourier KPSS are used to check the presence of unit root and stationarity. It may also solve the problem of over-fitting. The use of the Fourier form is suggested by Gallant and other researchers (Gallant 1981; Becker et al. 2006; Enders and Lee 2012a) in which trigonometric functions have been used to capture large variations from the mean of the dependent variable. It has such advantages over the traditional model as it takes into account a large number of soft transient structural breaks.

For this modelling using the Fourier function of Gallant (1981) and adding the structural changes to the KPSS test by Becker et al. (2006), so that time, number, and form were not predetermined after Ender and Lee (2012) added these time, number, and form nonlinear functions to the Dickey-Fuller (DF) type test (unit root test).

$$\Delta y_t = \alpha(t) + \delta t + \vartheta y_{t-1} + \sum_{i=1}^{p} \beta_i \Delta y_{t-i} + u_t \quad (2)$$

In above Eq. 2, $\alpha(t)$ is a deterministic term as a function of time, p is the optimal lag length determined by Akaike (AIC) and Schwarz (SIC), u_t is a stationary error term with variance, and β_i and ϑ are coefficients. To prevent the autocorrelation problem, delayed values of Δy are included in the model. For null hypothesis, $\vartheta = 0$ is the problematic ad per Enders and Lee (2012b), and this problem can be eliminated by using the Fourier function to the deterministic term and expressed as:

$$\alpha(t) = \alpha_0 + \sum_{k=1}^{n} \gamma_{1k} sin\left(\frac{2\pi kt}{T}\right) + \sum_{k=1}^{n} \gamma_{2k} cos\left(\frac{2\pi kt}{T}\right) \quad (3)$$

where n is the number of Fourier frequencies, the higher or excessive frequency components may cause the overfitting problem due to reducing the degree of freedom. That is why taking the Fourier component for a single frequency and found:

$$\alpha(t) = \alpha_0 + \gamma_1 sin\left(\frac{2\pi kt}{T}\right) + \gamma_2 cos\left(\frac{2\pi kt}{T}\right) \quad (4)$$

After a simple modification, it is to allow the deterministic term to the time-dependent function and the FADF unit root test with a single frequency equation is found by putting Eq. (5) into Eq. (2) and getting the new equation for optimal leg length as follows:

$$\Delta y_t = \alpha_1 + \delta t + \gamma_1 sin\left(\frac{2\pi kt}{T}\right) + \gamma_2 cos\left(\frac{2\pi kt}{T}\right) + \sum_{i=1}^{p} \beta_i \Delta y_{t-i} + \vartheta y_{t-i} + u_t \quad (5)$$

If the value of ϑ is zero, it states the series have a unit root, where k is Fourier frequency and T is the number of observations or data processed. By the bootstrap simulation, it is observed that the t statistic is greater than the table value which leads to the variables can be determined as a stationary process.

The Fourier Toda Yamamoto causality test

Toda and Yamamoto (1995), developed a method that depends on vector auto regression (VAR), applied for the level values of the series, and removes the long-term information loss. The TY casualty test with environmental degradation factor and GDP and NE can be written as follows:

$$y_t = \alpha_0 + \alpha_1 t + \alpha_2 sin\left(\frac{2\pi kt}{T}\right) + \alpha_3 cos\left(\frac{2\pi kt}{T}\right) + \varepsilon_t \quad (6)$$

$$y_t = \alpha_0 + \gamma_1 sin\left(\frac{2\pi kt}{T}\right) + \gamma_2 cos\left(\frac{2\pi kt}{T}\right) + v_t \quad (7)$$

where ε_t is a stationary disturbance with variance, and k represents the single frequency selected for the approximation,

The Fourier bootstrap Toda-Yamamoto causality test is employed as Fourier takes into consideration the abrupt and gradual structural shifts, while bootstrap increases the power of test statistics; therefore, both Fourier and bootstrap are used while analyzing the causal link between the variables (Nazlioglu et al. 2016) as shown in Eq. (8).

$$y_t = \alpha_0 + \gamma_1 sin\left(\frac{2\pi kt}{T}\right) + \gamma_2 cos\left(\frac{2\pi kt}{T}\right) + \beta_1 y_{t-1} + \cdots + \beta_{p+d} y_{t-(p+d)} + \varepsilon_t \quad (8)$$

Cointegration regression and Dynamic ordinary least square (DOLS) or F-test

Tsong et al. (2016) developed a Fourier approximation to test co-integration without estimating the specific break dates as the structural breaks of unknown form can be approximated with low-frequency Fourier components. While testing for co-integration between the variables the assumption of strict exogeneity is violated as it is rather restrictive in practice. Therefore, the asymptotic distribution of CI_f^m the test will depend on an unknown nuisance parameter, and the estimation of the co-integration vector is not efficient. There are several methods to deal with such problem like dynamic ordinary least square (DOLS). The DOLS regression, test statistics, and F test can be summarized as follows.

$$y_t = \sum_{i=0}^{m} \delta_i l^i + \alpha_k sin\left(\frac{2\pi kt}{T}\right) + \beta_k cos\left(\frac{2\pi kt}{T}\right) + x_t'\beta + \sum_{i=-1}^{l} \Delta x_{t-i}' \varnothing_i + \varepsilon_t^* \quad (9)$$

$$CI_f^m = T^{-2} \omega_\varepsilon^{*-2} \sum_{i-1}^{T} S_t^{*2} \quad (10)$$

where CI_f^m is the KPSS type cointegration statistic to test the null of COINTEGRATION with a structural break, S_t is the partial sum of OLS residuals, and ω_ε represents the estimator for long-run variance of ε_t. The test is verified with the null or absence of the Fourier component ($\alpha_k = \beta_k = 0$) against the alternative of nonlinear trend or structural changes. The F test for the hypothesis testing is evaluated by Eq. (11) given as follows:

$$F^m(k^*) = max_{k\epsilon\{1,2,3\}} F^m(k) \tag{11}$$

where F^m is FMOLS (fully modified ordinary least squares),$F^m(k)$ is the dynamic ordinary least squares (DOLS) estimated by:

$$F^m(k) = \frac{(SSE_0^m - SSE_1^m(k))/2}{SSE_1^m(k)/(T-q)} \tag{12}$$

The 0 and 1 subscripts in the SSE are denoted by the null and alternate hypotheses, q is the number of parameters, and k is unidentified under the null hypothesis.

Wavelet coherence

Drawing things to close wavelet coherence facilitates the study of time series data in the time and frequency domain. It helps to gain information that was formerly undetected. It analyzes the causal and correlation between the dependent and the independent variables from the short to long run (Torrence and Webster 1998).

$$R_n^2(s) = \frac{\left|S\left(s^{-1}W_n^{XY}(s)\right)\right|^2}{S\left(s^{-1}\left|W_n^X(s)\right|^2\right).S\left(s^{-1}\left|W_n^Y(s)\right|^2\right)} \tag{13}$$

From the above equation, it is observed that the time and location are the two main parameters of the wavelet. W^{xy} is the cross wavelet and it depends on individual wavelet W^x and W^y which are the wavelets for each time series. The S is the smoothing operator and s is the scale and n is the time index, s^{-1} is used to convert energy density, and the values of R^2 show the different features in the wavelet, and these are:

(i) $0 \le R^2(k,s) \le 1$ indicates the smoothing process overtime.
(ii) $R^2 \approx 1$ indicates a black line with red color picture.
(iii) $R^2 \approx 0$ indicates weakly correlated with a blue color picture.
(iv) $R^2 \approx$ *squaredvalue* provides only the strength of correlation but does not provide any information about the direction.

Torrence and Compo (1998) and Pal and Mitra (2017) developed the procedure to detect the wavelet coherence differences through the above-mentioned indications of deferrals in the wavering of the two time series.

Results and discussion

The results of the Fourier ADF, Fourier GLS, Fourier LM, and Fourier KPSS (Eqs. 2–4) test are presented in Tables 2 and 3. It is observed that the Fourier approximation has an extra edge over the conventional model as it solves the problems of unknown multiple breaks. The lags and Fourier have been automatically chosen by AIC and SCC criteria. Whereas, the maximum lags found 4 and the maximum Fourier found 3.

The unit root test by Fourier ADF, Fourier GLS, and Fourier LM rejected the null hypothesis of a unit root in a combination of I(0) and I(1) for all variables and all the three selected countries. Whereas, the null hypothesis of stationarity for the Fourier KPSS is rejected at I(0) but accepted at I(1) of the variables for all the three countries.

The Fourier ARDL model developed by Tsong et al. 2016 is adopted for the study. The model was enabled to approximate the breaks of unknown form. Therefore, estimating specific break dates has not been necessary as it is

Table 2 Unit root test by Fourier ADF and Fourier GLS

Countries and variables	FADF I(0)	FADF I(1)	Fourier component	FGLS I(0)	FGLS I(1)	Fourier component
USA						
CO_2	−3.943	−3.873**	1	−2.975	−2.639*	1
GDP	−4.189*	−2.859*	1	−4.055*	−4.144*	1
NE	−5.532***	−4.843**	1	−2.962	−4.945***	1
France						
CO_2	−2.956	−7.360***	1	−3.144*	−5.737***	2
GDP	−1.246	−4.855**	1	−1.592	−3.262*	1
NE	−2.805	−4.052*	2	−0.879	−4.224**	1
Japan						
CO_2	−6.707***	−6.983***	1	−3.161	−7.252***	1
GDP	−7.359***	−5.353***	2	−2.532	−5.532***	1
NE	−4.714***	−8.635***	2	−2.834	−3.546*	2

Optimal lag automatically by AIC ***, ** and * denote 1%, 5%, and 10% significance levels

Table 3 Unit root and stationarity test by Fourier LM and Fourier KPSS

Countries and variables	FLM I(0)	FLM I(1)	Fourier component	FKPSS I(0)	FKPSS I(1)	Fourier component
USA						
CO_2	−1.261	−2.696*	1	0.232***	0.034	1
GDP	−3.916*	−4.128**	1	0.099***	0.051	1
NE	−2.016	−5.545***	2	0.291***	0.068	3
France						
CO_2	−2.642	−7.494***	1	0.394***	0.280	2
GDP	−1.741	−4.414**	1	0.266***	0.029	1
NE	−2.330	−2.578*	3	0.170***	0.166	3
Japan						
CO_2	−1.298	−7.043***	1	0.213***	0.070	1
GDP	−1.833	−4.401**	1	0.120***	0.138	2
NE	−2.256	−2.659*	1	0.168***	0.081	3

Optimal lag automatically by AIC ***, **, and * denote 1%, 5%, and 10% significance levels

Table 4 Fourier ARDL (FARDL)

Countries and variables	*F*-Stats	CIF	Fourier component
USA			
FARDL	69.632	0.091	2
1%		0.198	
5%		0.124	
France			
FARDL	52.106	0.083	1
1%		0.163	
5%		0.099	
Japan			
FARDL	27.362	0.059	2
1%		0.163	
5%		0.099	

Optimal lag automatically by AIC *** and ** denote 1% and 5% significance levels

automatically included while adding Fourier components. The model is restricted up to four independent variables only and the maximum number of Fourier found 3. The optimal lag and lead are automatically chosen by AIC. The asymptotic critical values are being tabulated by simulation techniques.

Table 4 presents the results of Fourier ARDL. It expressed the Fourier components for approximating unknown structural breaks that provide strong evidence in favor of the validation of the EKC hypothesis while using nuclear energy in the long run for all three countries. Hence, the study proved that with economic growth and increasing the utilization of clean energy will also be decreasing the environmental pollution in the long run. It is also provides strong evidence of the existence of inverted U-shaped curve relation between income and environmental pollution as the null hypothesis of co-integration is accepted for all three countries.

The Toda-Yamamoto approach is used to evaluate the problem of the conventional model by estimating VAR (p + d) where *d* is the maximum integration order of variables. Whereas, Fourier Toda-Yamamoto takes into account for both the abrupt and gradual breaks. Furthermore, the Wald statistic may not follow an asymptotic chi-square distribution as it is dependent on the number of frequencies. Therefore to overcome the problem, bootstrapping is applied.

Table 5 presents the results of the Fourier bootstrap Toda-Yamamoto (Eqs. 5–9). These results are reflecting the bidirectional causal link from CO_2 emissions to nuclear energy and also a two-way link from GDP to nuclear energy for the USA. Furthermore, in the case of France, the consequence replicates the bidirectional linkage between GDP and nuclear energy, and in the case of Japan, the results expressed the unidirectional link between GDP and CO_2 emissions.

The Wavelet coherence shows the co-movement of the two time series data from short run to long run. The vertical axis and horizontal axis of wavelet coherence are represented by periods and years respectively. The short-run, medium-run, and long-run periods are referred to as 0–4, 4–8, and 8–16 respectively. The color band on the right side reflects the level of correlation between the variables where red hot refers to highly correlated and darkest blue represents the lowest correlation. The black contour represents the 5% significance level. The gray line is the cone of influence below; its interpretation is difficult due to discontinuity. The upward and downward arrows represent the causal linkage, whereas the rightward and leftward arrows refer to the phase in and phase out respectively.

Table 5 Fourier Bootstrap Toda-Yamamoto

Countries and variables	WaId	*p*-value	Bootstrap *p*-value	Fourier component
USA				
$CO_2 \rightarrow GDP$	0.029	0.866	0.851	1
$GDP \rightarrow CO_2$	0.275	0.600	0.598	1
$CO_2 \rightarrow NE$	4.269	0.038**	0.040**	1
$NE \rightarrow CO_2$	3.657	0.056**	0.060*	1
$GDP \rightarrow NE$	5.459	0.019**	0.029**	1
$NE \rightarrow GDP$	2.924	0.087*	0.090*	1
France				
$CO_2 \rightarrow GDP$	1.758	0.185	0.213	2
$GDP \rightarrow CO_2$	0.041	0.839	0.831	2
$CO_2 \rightarrow NE$	0.716	0.397	0.413	2
$NE \rightarrow CO_2$	1.106	0.293	0.314	2
$GDP \rightarrow NE$	3.483	0.062*	0.079*	2
$NE \rightarrow GDP$	5.081	0.024**	0.036**	2
Japan				
$CO_2 \rightarrow GDP$	0.355	0.552	0.544	1
$GDP \rightarrow CO_2$	7.264	0.007***	0.009***	1
$CO_2 \rightarrow NE$	0.000	1.000	0.999	1
$NE \rightarrow CO_2$	0.312	0.577	0.578	1
$GDP \rightarrow NE$	0.058	0.809	0.832	1
$NE \rightarrow GDP$	0.051	0.822	0.795	1

Optimal lag automatically by AIC ***, **, and * denote 1%, 5%, and 10% significance levels, bootstrap = 1000

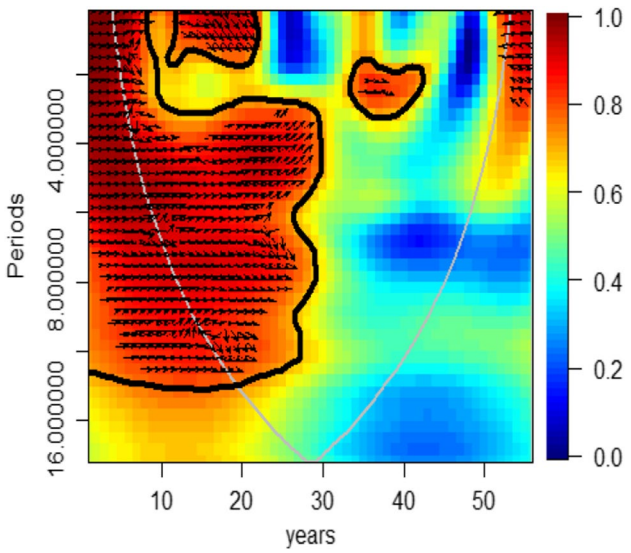


Fig. 2 Wavelet coherence USA CO_2 Vs GDP

In Fig. 2 (Eq. 10), wavelet coherence of the USA for CO_2 Vs GDP shows that for the first 25 years, the red hot color expressed a high correlation between CO_2 emission and GDP for the medium term. The rightward arrow in the middle

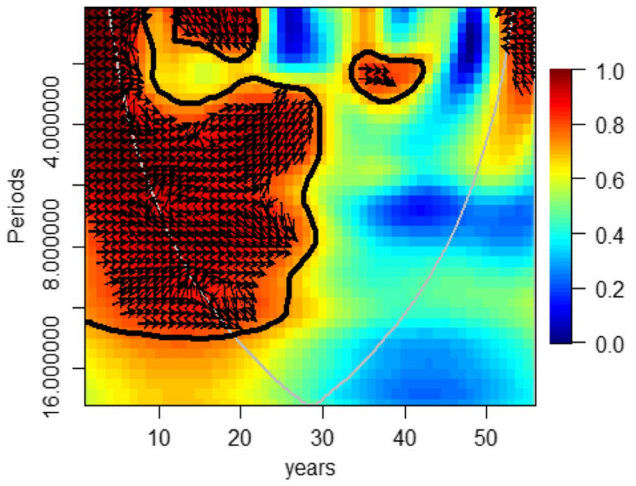


Fig. 3 Wavelet coherence USA CO_2 Vs NE

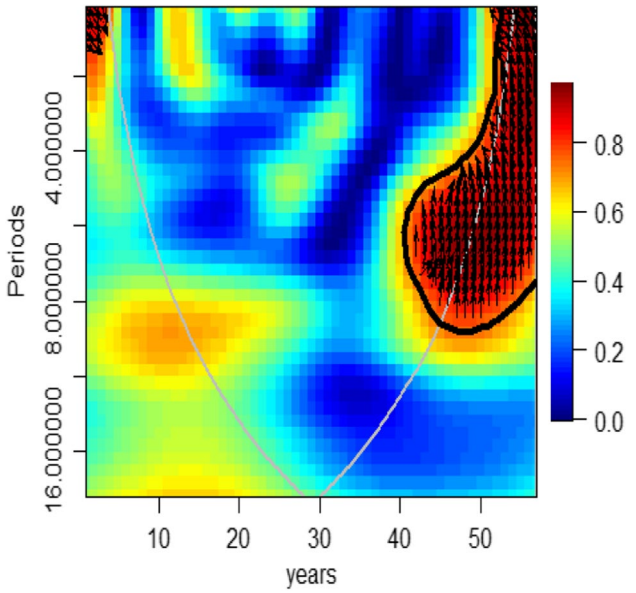


Fig. 4 Wavelet coherence France CO_2 Vs GDP

represents the phase-in, whereas the upward and downward arrows represent the causal link between CO_2 emission and GDP. Figure 3 shows the wavelet coherence between CO_2 emission and nuclear energy (NE), and it observed a healthy correlation between the two by bright red color from the short run to the medium run. The upward and downward arrows refer to the causal link from CO_2 emission to NE.

Figure 4 shows the wavelet coherence between CO_2 emission and GDP for France. In which the last 20 years show red hot color and upward arrows; it means the variables are highly correlated and also has a causal linkage relation. Figure 5 represents the wavelet coherence between CO_2 and NE for the period of the years from 2000 to 2020 of France,

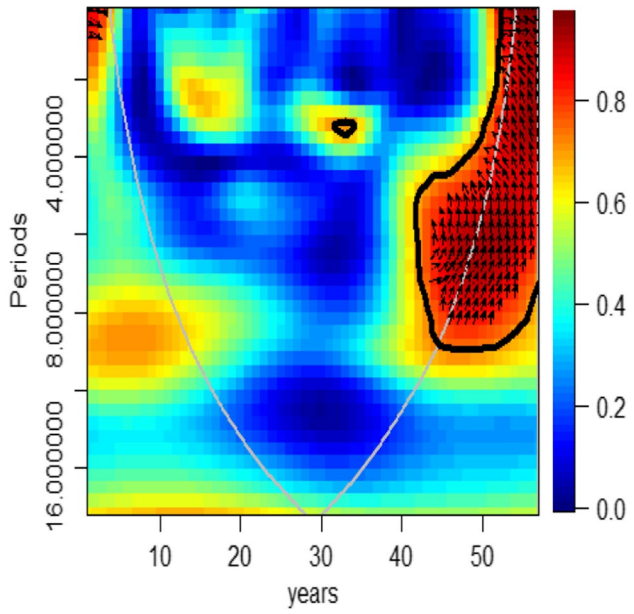


Fig. 5 Wavelet coherence France CO_2 Vs NE

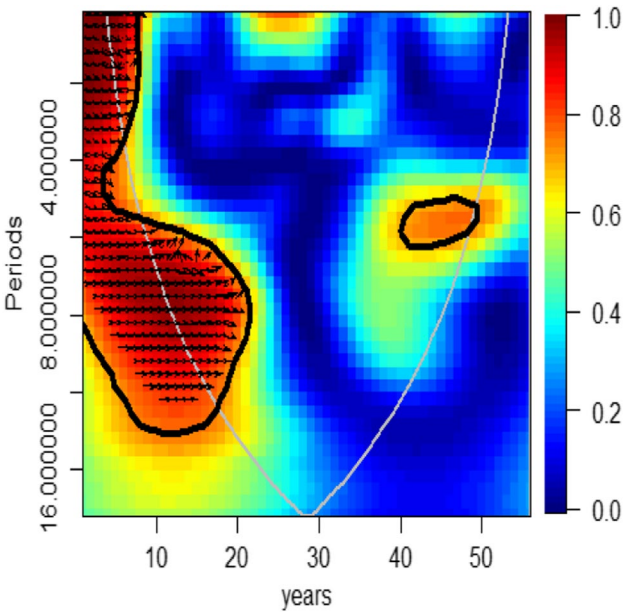


Fig. 7 Wavelet coherence Japan CO_2 Vs NE

and it is observed the bright red color which indicates that CO_2 emission and NE are well correlated. The upward arrow shows a causal link between CO_2 emission and NE.

Figure 6 shows the wavelet coherence of CO_2 emission and GDP for Japan. The red hot color and upward and downward arrows are spread from the short run to the long run from the year 1965 to 2020. It expressed the high correlation and causal linkage between CO_2 emission and the GDP of Japan. Figure 7 also represents the correlation between variables from the short run to the medium run for the first 20 years only. The rightward arrow refers to phase-in, and for the later 40 years, it reflects the dark blue color which means there is no correlation between the variables.

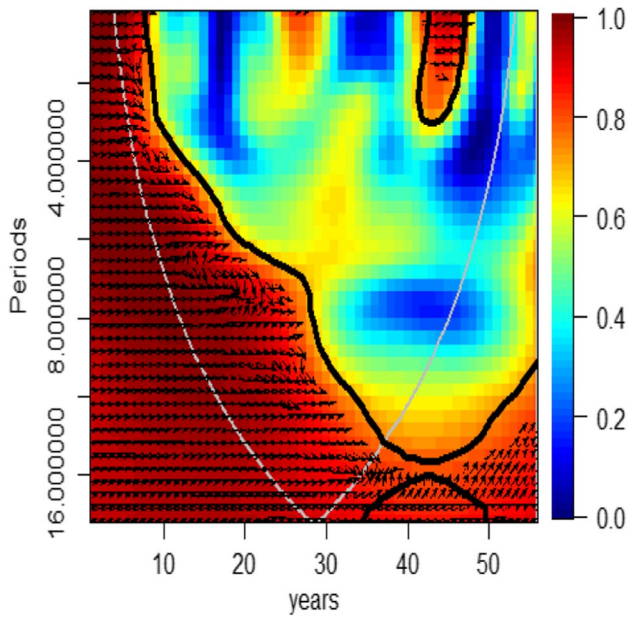


Fig. 6 Wavelet coherence Japan CO_2 Vs GDP

BAU estimation for energy demand meet out through NE and carbon mitigation

The 56-year analysis of France showed that power generation through nuclear energy took 29 years, i.e., 1993 to reach the breakeven point where CO_2 emission was 362.0301 million tonnes and electricity production via nuclear energy was 368.188 TWh as shown in Fig. 8. Furthermore, it took almost 8 years to cross the breakeven point after which the CO_2 emission started declining. According to thirty years (2021–2050) of business as usual (BAU) estimates if France continues to meet its increasing demand for electricity through nuclear energy, CO_2 emission will be mitigated by 15% declining to 212 million tonnes by 2050 generating 734.9321 TWh of electricity.

As we know, the USA produce's highest units of electricity through nuclear energy worldwide (U.S. Energy Information Administration 2022); yet, it is far away from reaching the breakeven point, i.e., where CO_2 emission is equal to electricity generation through nuclear energy. Figure 9 presents the BAU estimation of the USA for energy demand meet out via NE and CO_2 emission. In which, during the years 1974–1976, the NE production increased by 36.518% which had a positive impact and mitigated CO_2 emissions by 7.627%, after which the NE declined thus leading to an

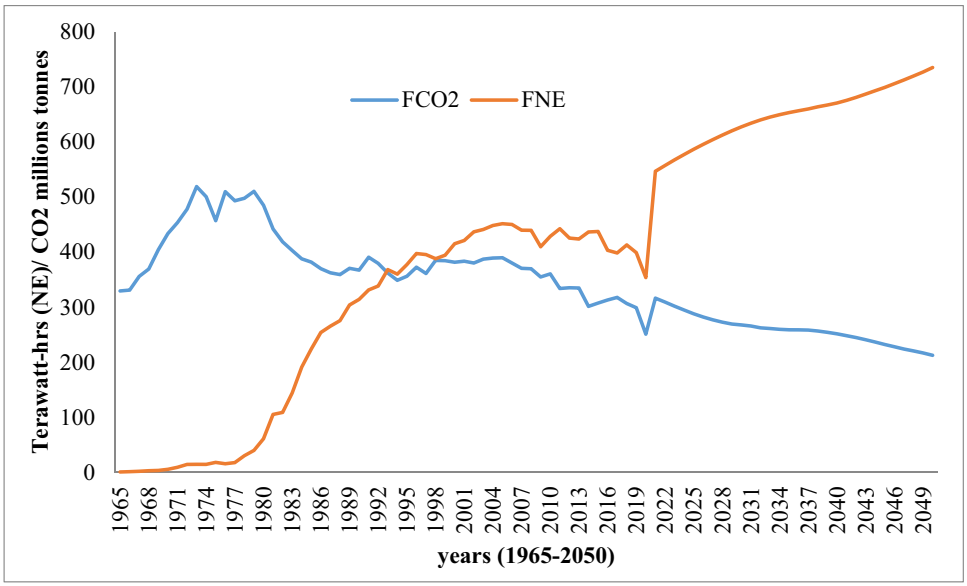


Fig. 8 BAU Estimation of France for energy demand meet out via NE and CO_2 emission

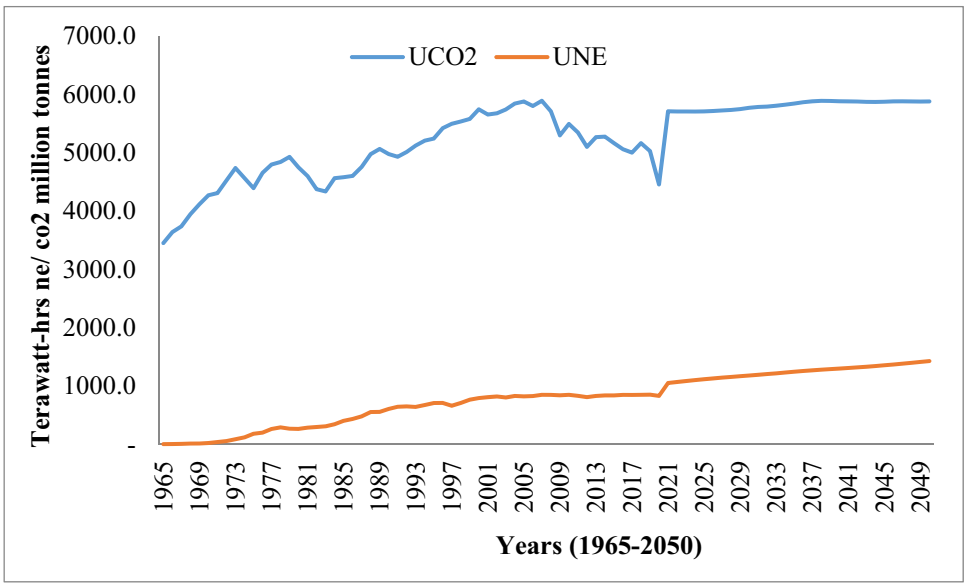


Fig. 9 BAU estimation of USA for energy demand meet out via NE and CO_2 emission

increase in CO_2 emission. The BAU estimate indicates that the USA needs to increase its electricity production through nuclear energy in order to meet its high electricity demand for offsetting the CO_2 emission by 2050. If the USA continues to produce electricity via nuclear energy at the same pace, then it will require more than 2500 years to reach the breakeven point.

Japan's first commercial nuclear power reactor began operating in mid-1966, and nuclear energy has been a national strategic priority since 1973 after which power generation via NE took an increasing trend, but the scenario totally changed after the 2011 Fukushima nuclear accident after which the use of NE declined from 162.9 to 18 TWh in 2012. It further declined to 0.5 TWh in 2014 touching the horizontal axis. It has been noted that almost after a decade Japan took an initiative to increase the production of electricity via nuclear energy. From Fig. 10, the BAU estimates that NE will be in a declining tread till the year 2050. Hence, to mitigate the CO_2 emission to reach the net zero carbon emission, Japan needs to increase the pace of electricity generation via NE.

CO_2 mitigation

The power generation from nuclear energy is shown in Table 5 for France Japan and USA for the year 2020. The data for nuclear energy power generation is taken from the source (IEA 2020).

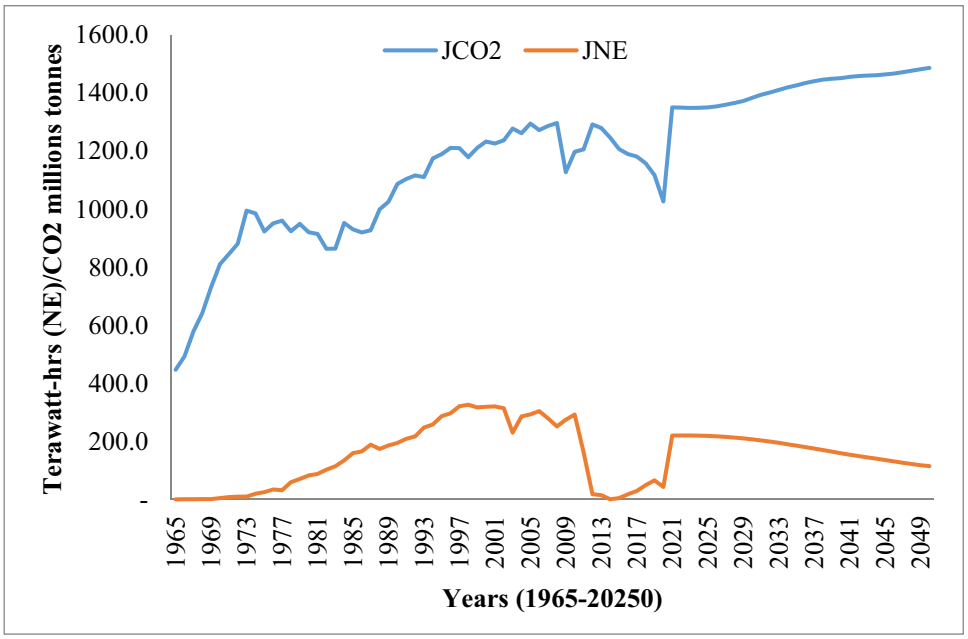


Fig. 10 BAU estimation of Japan for energy demand meet out via NE and CO_2 emission

The Mittal et al. (2012) estimated the specific coal consumption per kWh as 0.77 kg/kWh, and we have selected the same value for the calculation of specific coal consumption. Kaushik et al. (2014) and Lal et al. (2014) studied the CO_2 equivalent, whereas the average CO_2 equivalent for electricity has been taken as 0.98 kg/kWh CO_2. The total coal required to produce the electrical power (equivalent to nuclear power generation) in France, Japan, and USA for the year 2020 is 47,254.9, 24,392.8, and 74,305.46 MT respectively. The CO_2 emission is to be reduced by using the nuclear fuel (clean energy fuel) in France, Japan, and USA, and during the year 2020, it is 46,309.8, 23,904.97, and 72,819.35 MT/year. The CO2 reduction from the year 1980 to 2020 by using the nuclear energy to produce electrical energy is shown in Fig. 11. The other pollutants can also be saved like SO_2 and NO*x* by all three countries given in Table 6. It is noted that the huge sum of carbon dioxide production can be avoided by using the NE for power generation which is beneficial to reduce the global warming also.

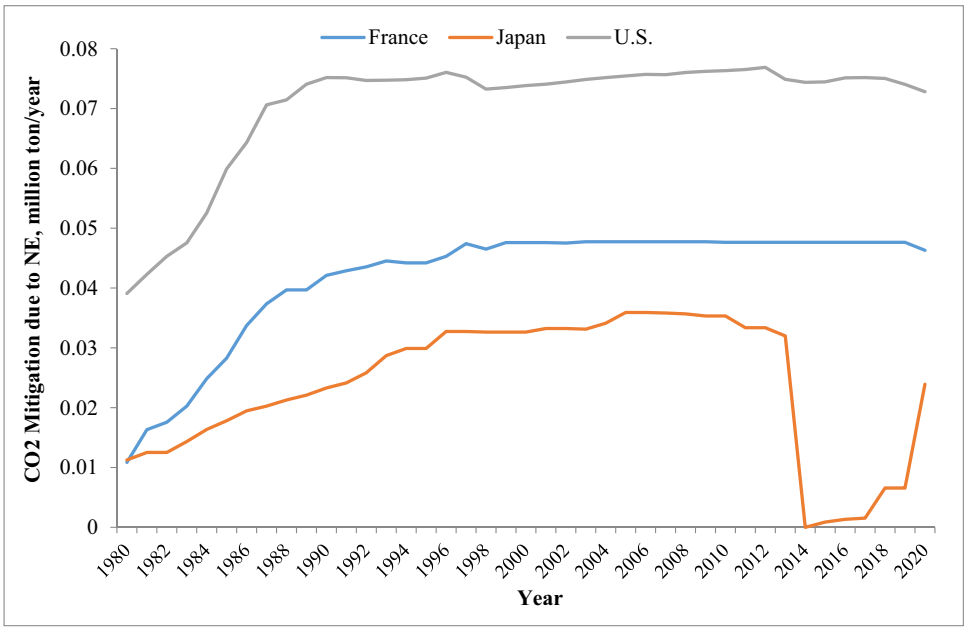


Fig. 11 Carbon dioxide mitigation year-wise during 1980 to 2020

Table 6 Electrical energy generation by nuclear energy during the year 2020

Country	Nuclear electricity generation capacity (million kilowatts)	Nuclear electricity generation (billion kilowatt hours)	CO2 Mitigation MMT/year	SO2 emission saved MMT/Year	NO*x* emission saved per year MMT/year
France	61.37	379.50	0.046309802	84.32238	0.284757
Japan	31.68	43.00	0.023904973	43.52695	0.146991
USA	96.50	789.88	0.072819353	132.5918	0.447763

**MMT*, million metric ton; *MT*, metric ton

Conclusion and policy suggestion

Among all the three countries, only France has been able to reach the breakeven point way back in 1993. France has also shown a positive impact of the use of NE for electricity generation as with the increase in NE CO_2 emission has declined over the period of time. Although the USA is the highest producer of electricity through NE in the world, still, its major demands are met by 36% petroleum, 32% natural gas, 11% coal, and only 8% from NE. Japan's national strategic priority was the production of electricity through nuclear energy in 1973, but the scenario totally changed after the Fukushima nuclear accident in 2011; the percentage of NE use dropped to 99.95%. Almost after a decade, Japan government took an initiative to increase the production of electricity via NE.

The USA and Japan meet their energy requirements 36% and 40% respectively via petroleum. Whereas the coal and petroleum are the major emitters of carbon dioxide (CO_2), sulphur dioxide (SO_2), and nitrous oxide (NO_x) which not only harm the environment and increase global warming but also have a negative impact on human health. Therefore the USA and Japan need to develop carbon capture utilization and storage technology (CCUS) which captures carbon before it enters the atmosphere, and it can be further utilized to produce by-products like sodium bicarbonate, ink, carbon fibers, methanol, and bio plastics; at present, CCUS technology captures 40mt CO_2 all around the world. According to the geographical location of all the three countries, the study suggests promoting tidal energy for electricity generation as tides are more predictable than wind and sun.

The study has been conducted to check the validity of the EKC hypothesis while using nuclear energy for USA, France, and Japan. The results showed strong evidence in favor of the EKC hypothesis while using nuclear energy, and it is observed that the utilization of clean energy (nuclear energy) for electrical power generation will reduce the environmental pollution in the long run. It will also help the nations in achieving the sustainable developmental goals of net zero carbon emission by 2050. The CO2 mitigation by using the NE is also assessed for all three countries and it is assessed that France, Japan, and the USA mitigated the CO2 per year by 0.0463 million metric ton (MMT), 0.0239, and 0.0728 MMT per year respectively. Similarly to that, the SO2 is reduced by using the NE is 24.322, 43.527, and 132.592 MMT/year, and NO*x* is reduced by approximately 0.2847, 0.147, and 0.4478 MMT/year by France, Japan, and the USA respectively by applying the NE for power generation.

Although nuclear energy is a clean energy, but its residuals are highly radioactive as it has the traces of thorium and uranium. Therefore, it is needed to develop a full-proof plan for disposing of NE residuals in a safe place. The cost of disposing of the radioactive waste should also be added to the total cost of electricity generation via nuclear energy. So that, there is no fund problem occurring in the future for disposing of the waste in a safe place. Future research can be done on minimizing the cost of disposing of the leftovers.

Author contribution Akanksha Singh: conceptualization and writing. Shiv Lal: writing, review, and editing. Nand Kumar and Rajan Yadav: supervision. Shweta Kumari: software, data, methodology, conceptualization, and writing.

Data availability World Bank and BP statistical review of the world.

Declarations

Ethical approval Throughout the drafting process, all ethical practices were followed.

Consent for publication Not applicable.

Competing interests The authors declare no competing interests.

References

Abbasi F, Riaz K (2016) CO2 emissions and financial development in an emerging economy: an augmented VAR approach. Energy Policy 90:102–114. https://doi.org/10.1016/j.enpol.2015.12.017

Adebayo TS (2020) Dynamic relationship between oil price and inflation in oil exporting economy: empirical evidence from wavelet coherence technique. Energy Econ Lett 7(1):12–22. https://doi.org/10.18488/journal.82.2020.71.12.22

Adebayo TS, Awosusi AA, Adeshola I (2020) Determinants of CO 2 emissions in emerging markets: an empirical evidence from

MINT economies. International Journal of Renewable Energy Development 9(3):411–422. https://doi.org/10.14710/ijred.2020.31321

Adebayo TS, Kirikkaleli D, Adeshola I, Oluwajana D, Akinsola GD, Osemeahon OS (2021) Coal consumption and environmental sustainability in South Africa: the role of financial development and globalization. Int J Renew Energy Dev 10(3):527–536. https://doi.org/10.14710/ijred.2021.34982

Adebayo TS (2022a) Assessing the environmental sustainability corridor : linking oil consumption, hydro energy consumption, and ecological footprint in Turkey. Environ Sci Pollut Res 1–11. Available from: https://doi.org/10.1007/s11356-022-23455-1

Adebayo TS (2022b) Environmental consequences of fossil fuel in Spain amidst renewable energy consumption : a new insights from the wavelet-based Granger causality approach. Int J Sustain Dev World Ecol 29(7):579–92.: https://doi.org/10.1080/13504509.2022b.2054877

Ahmed Z, Wang Z, Ali S (2019) Investigating the non-linear relationship between urbanization and CO 2 emissions: an empirical analysis. Air Qual Atmos Health 12(8):945–953. https://doi.org/10.1007/s11869-019-00711-x

Ahmed Z, Zafar MW, Ali S (2020) Linking urbanization, human capital, and the ecological footprint in G7 countries: an empirical analysis. Sustainable Cities and Society 55:102064. https://doi.org/10.1016/j.scs.2020.102064

Akbota A, Baek J (2018) The environmental consequences of growth: empirical evidence from the Republic of Kazakhstan. Economies 6(1):1–11. https://doi.org/10.3390/economies6010019

Alam, Shabbir Md (2022) Is trade, energy consumption and economic growth threat to environmental quality in Bahrain–evidence from VECM and ARDL bound test approach. International Journal of Emergency Services 11(3):396–408. https://doi.org/10.1108/IJES-12-2021-0084

Alam S, Noor M, Muntasir A, Mahmood H, Alam R (2022) Pathways to securing environmentally sustainable economic growth through efficient use of energy : a bootstrapped ARDL analysis. Environ Sci Pollut Res 29:50025–50039. https://doi.org/10.1007/s11356-022-19410-9

Albela HP, Lawankar SM, Deshmukh S (2012) Performance and emissions of LPG fueled internaL combustion engine: a review. Int J Sci Eng Res 3(3):1–7

Al-Mulali U (2011) Oil consumption, CO2 emission and economic growth in MENA countries. Energy 36(10):6165–6171. https://doi.org/10.1016/j.energy.2011.07.048

Al-Mulali U, Solarin SA, Ozturk I (2016) Investigating the presence of the environmental Kuznets curve (EKC) hypothesis in Kenya: an autoregressive distributed lag (ARDL) approach. Nat Hazards 80(3):1729–1747. https://doi.org/10.1007/s11069-015-2050-x

Alvarado R, Tillaguango B, Murshed M (2022) Impact of the informal economy on the ecological footprint : the role of urban concentration and globalization. Econ Anal Policy 75:750–67. https://doi.org/10.1016/j.eap.2022.07.001

Anser MK, Yousaf Z, Nassani AA, Abro MMQ, Zaman K (2020) International tourism, social distribution, and environmental Kuznets curve: evidence from a panel of G-7 countries. Environ Sci Pollut Res 27(3):2707–2720. https://doi.org/10.1007/s11356-019-07196-2

Apergis N, Payne JE (2010) The emissions, energy consumption, and growth nexus: evidence from the commonwealth of independent states. Energy Policy 38:650–655. https://doi.org/10.1016/j.enpol.2009.08.029

Aydoğan B, Vardar G (2020) Evaluating the role of renewable energy, economic growth and agriculture on CO2 emission in E7 countries. Int J Sustain Energ 39(4):335–348. https://doi.org/10.1080/14786451.2019.1686380

Aye GC, Edoja PE (2017) Effect of economic growth on CO2 emission in developing countries: evidence from a dynamic panel threshold model. Cogent Economics & Finance 5(1):1379239. https://doi.org/10.1080/23322039.2017.1379239

Becker R, Enders W, Lee J (2006) A stationarity test in the presence of an unknown number of smooth breaks. J Time Ser Anal 27(3):381–409. https://doi.org/10.1111/j.1467-9892.2006.00478.x

Bekhet HA, Matar A, Yasmin T (2017) CO_2 emissions, energy consumption, economic growth, and financial development in GCC countries: dynamic simultaneous equation models. Renew Sust Energ Rev 70:117–132. https://doi.org/10.1016/j.rser.2016.11.089

Bekun FV, Agboola MO, Joshua U (2020) Fresh insight into the EKC hypothesis in Nigeria: accounting for total natural resources rent. In econometrics of green energy handbook. Springer, Cham, Corpus ID 226718802:221–243. https://doi.org/10.1007/978-3-030-46847-7_11

Bildirici ME, Gökmenoğlu SM (2017) Environmental pollution, hydropower energy consumption and economic growth: evidence from G7 countries. Renew Sustain Energy Rev 75(C):68–85. https://doi.org/10.1016/j.rser.2016.10.052

Bilgili F, Lorente DB, Kuşkaya S, Ünlü F, Gençoğlu P, Rosha P (2021) The role of hydropower energy in the level of CO_2 emissions: an application of continuous wave- let transform. Renew Energy 178:283–294. https://doi.org/10.1016/j.renene.2021.06.015

BP stats (2021) https://www.bp.com/content/dam/bp/business-sites/en/global/corporate/pdfs/energy-economics/statistical-review/bp-stats-review-2021-full-report.pdf. Assessed 11 August 2022

Charfeddine L, Kahia M (2019) Impact of renewable energy consumption and financial development on CO2 emissions and economic growth in the MENA region: a panel vector autoregressive (PVAR) analysis. Renew Energy 139(C):198–213. https://doi.org/10.1016/j.renene.2019.01.010

Dogan E, Inglesi-Lotz R (2020) The impact of economic structure to the environmental Kuznets curve (EKC) hypothesis: evidence from European countries. Environ Sci Pol Res 27(11):12717–12724. https://doi.org/10.1007/s11356-020-07878-2

Dong K, Sun R, Hochman G (2017) Do natural gas and renewable energy consumption lead to less CO2 emission? Empirical evidence from a panel of BRICS countries. Energy 141(C):1466–1478. https://doi.org/10.1016/j.energy.2017.11.092

Dong F, Wang Y, Su B, Hua Y, Zhang Y (2019) The process of peak CO2 emissions in developed economies: a perspective of industrialization and urbanization. Resour Conserv Recycl 141:61–75. https://doi.org/10.1016/j.resconrec.2018.10.010

Ekbordin W, Tami C Bond (2016) Emissions from residential combustion considering end-uses and spatial constraints: part II, emission reduction scenarios. Atmospheric Environment 124(A):1–11. https://doi.org/10.1016/j.atmosenv.2015.10.011

Emrah B, Salih K (2021) Environmental Kuznets curve (EKC): empirical relationship between economic growth, energy consumptioN, and CO2 emissions: evidence from 3 developed countries. Panoeconomicus 64(4):483–506. https://doi.org/10.2298/PAN180503004B

Enders W, Lee J (2012a) The flexible Fourier form and Dickey-Fuller type unit root tests. Econ Lett 117(1):196–199. https://doi.org/10.1016/j.econlet.2012.04.081

Enders W, Lee J (2012b) A unit root test using Fourier series to approximate smooth breaks. Oxford Bull Econ Stat 74(4):574–599. https://doi.org/10.1111/j.1468-0084.2011.00662.x

Faisal, Tursoy T, Resatoglu NG (2016) Energy consumption, electricity, and GDP causality; the case of Russia, 1990–2011. Proc Econ Finance 39:653–659. https://doi.org/10.1016/j.procs.2017.11.294

Farooq S, Ozturk I, Tariq M, Akram R (2022) Globalization and CO2 emissions in the presence of EKC : A global panel data analysis.

Gondwana Res 106:367–78. https://doi.org/10.1016/j.gr.2022.02.002
Gallant AR (1981) On the bias in flexible functional forms and an essentially unbiased form: the flexible Fourier form. J Econom 15(2):211–245. https://doi.org/10.1016/0304-4076(81)90115-9
Gao J, Zhang L (2021) Does biomass energy consumption mitigate CO2 emissions? The role of economic growth and urbanization: evidence from developing Asia. Journal of the Asia Pacific Economy 26(1):96–115. https://doi.org/10.1080/13547860.2020.1717902
Gorus MS, Aydin M (2019) The relationship between energy consumption, economic growth, and CO2 emission in MENA countries: causality analysis in the frequency domain. Energy 168:815–822. https://doi.org/10.1016/j.energy.2018.11.139
Grossman GM, Krueger AB (1991) Environmental impacts of a North American free trade agreement. National Bureau of economic research 3914. https://doi.org/10.3386/w3914
Hamid I, Alam S, Kanwal A, Kumar P, Muntasir J (2022a) Decarbonization pathways : the roles of foreign direct investments, governance, democracy, economic growth, and renewable energy transition. Environ Sci Pollut Res 29:49816–49831. https://doi.org/10.1007/s11356-022-18935-3
Hamid I, Alam S, Murshed M, Kumar P, Nadia J (2022) The roles of foreign direct investments, economic growth, and capital investments in decarbonizing the economy of Oman. Environ Sci Pollut Res 29:22122–22138. https://doi.org/10.1007/s11356-021-17246-3
Heidari H, Katircioglu ST, Saeidpour L (2015) Economic growth, CO2 emissions, and energy consumption in the five ASEAN countries. Electr Power Energy Syst 64:785–791. https://doi.org/10.1016/j.ijepes.2014.07.081
https://data.worldbank.org/indicator/EN.ATM.CO2E.KD.GD?end=2019&locations=FR-IN-JP-US&start=1990&type=shaded&view=chart&year=2019. Assessed on 22 July 2022
International Energy Agency (2020) https://www.iea.org/topics/net-zero-emissions, https://www.iea.org/data-and-statistics
Iwata H, Okada K, Samreth S (2010) Empirical study on the environmental Kuznets curve for CO2 in France: the role of nuclear energy. Energy Policy 38:4057–4063. https://doi.org/10.1016/j.enpol.2010.03.031
Jafari Y, Ismail MA, Othman J, Mawar MY (2015) Energy consumption, emissions and economic growth in Bahrain. Chinese Journal of Population Resources and Environment 13(4):297–308. https://doi.org/10.1080/10042857.2015.1078493
Jahanger, A., Yang, B., Huang WC, Murshed M,Usman M, Raghulescu M, (2022), Dynamic linkages between globalization, human capital, and carbon dioxide emissions: empirical evidence from developing economies. Environ Dev Sustain (2022). https://doi.org/10.1007/s10668-022-02437-w
Jebli MB, Youssef SB, Ozturk I (2016) Testing environmental Kuznets curve hypothesis: The role of renewable and non-renewable energy consumption and trade in OECD countries. Ecological Indicators 60:824–831. https://doi.org/10.1016/j.ecolind.2015.08.031
Jian J, Fan X, He P, Xiong H, Shen H (2019) The effects of energy consumption, economic growth and financial development on CO2 emissions in China: a VECM Approach. Sustainability 11(18):4850. https://doi.org/10.3390/su11184850
Jin T, Kim J (2020) Investigating the environmental Kuznets curve for annex I countries using heterogeneous panel data analysis. Environ Sci Pollut Res 27:10039–10054. https://doi.org/10.1007/s11356-020-07668-w
Kaushik SC, Garg T, Lal S (2014) Thermal performance prediction and energy conservation potential of earth air tunnel heat exchanger for thermal comfort in building. J Renew Sustain Energy 6(1–13):013107. https://doi.org/10.1063/1.4872362
Khan S (2022) Murshed M, Ozturk I, Khudoykulov K. The roles of energy efficiency improvement, renewable electricity production, and financial inclusion in stimulating environmental sustainability in the Next Eleven countries. Renew Energy 193:1164–1176. https://doi.org/10.1016/j.renene.2022.05.065
Khattak SI, Ahmad M, Khan ZU, Khan A (2020) Exploring the impact of innovation, renewable energy consumption, and income on CO2 emissions: new evidence from the BRICS economies. Environ Sci Pollut Res 27(12):13866–13881. https://doi.org/10.1007/s11356-020-07876-4
Khobai H, Le Roux P (2017) The relationship between energy consumption, economic growth and carbon dioxide emission: the case of South Africa. International Journal of Energy Economics and Policy 7(3):102–109
Kirikkaleli D, Adebayo TS (2021) Do renewable energy consumption and financial development matter for environmental sustainability? New Global Evidence Sustainable Development 29(4):583–594. https://doi.org/10.1002/sd.2159
Kirikkaleli D, Adebayo TS, Khan Z, Ali S (2020) Does globalization matter for ecological footprint in Turkey? Evidence from dual adjustment approach. Environ Sci Pollut Res 28(11):14009–14017. https://doi.org/10.1007/s11356-020-11654-7
Kisswani KM, Harraf A, Kisswani AM (2018) Revisiting the environmental kuznets curve hypothesis: evidence from the ASEAN-5 countries with structural breaks. Appl Econ 51(2):1–14. https://doi.org/10.1080/00036846.2018.1529399
Koc S, Bulus GC (2020) Testing validity of the EKC hypothesis in South Korea: role of renewable energy and trade openness. Environ Sci Pollut Res 27(23):29043–29054. https://doi.org/10.1007/s11356-020-09172-7
Lal S, Balam NB, Jain HK (2014) Performance evaluation, energy conservation potential, and parametric study of borehole heat exchanger for space cooling in building. Journal of renewable and sustainable energy 6(1–13):023123. https://doi.org/10.1063/1.4872362
Lau E, Tan CC, Tang CF (2016) Dynamic linkages among hydroelectricity consumption, economic growth, and carbon dioxide emission in Malaysia. Energy Sources, Part B: Economics, Planning, and Policy 11(11):1042–1049. https://doi.org/10.1080/15567249.2014.922135
Leal PH, Marques AC (2020) Rediscovering the EKC hypothesis for the 20 highest CO2 emitters among OECD countries by level of globalization. International Economics 164:36–47
Li K, Lin B (2015) Impacts of urbanization and industrialization on energy consumption/CO2 emissions: does the level of development matter? Renew Sust Energ Rev 52:1107–1122
Mittal ML, Sharma C, Singh R (2012) Estimates of emissions from coal fired thermal power plants in India, presented in International Emission Inventory Conference, Tampa, Florida, August 14–16 (2012), p. 1–22
Muhammad ZC, Naushad A, Murshed M, Rehman A, Lorente DB (2022) Pathways towards environmental sustainability: exploring the influence of aggregate domestic consumption spending on carbon dioxide emissions in Pakistan. Environ Sci Pollut Res 29:45013–45030. https://doi.org/10.1007/s11356-022-18919-3
Murshed M (2018) Does improvement in trade openness facilitate renewable energy transition? Evidence from selected South Asian economies. South Asia Economic Journal 19(2):151–170. https://doi.org/10.1177/1391561418794691
Murshed M (2020b) Are trade liberalization policies aligned with renewable energy transition in low and middle income countries? An instrumental variable approach. Renewable Energy 151:1110–1123. https://doi.org/10.1016/j.renene.2019.11.106
Murshed M (2022) The Environmental Kuznets Curve hypothesis for carbon and ecological footprints in South Asia : the role of

renewable energy. Geo Journal 87(3):2345–2372. https://doi.org/10.1007/s10708-020-10370-6

Murshed M, Mahmood H, Ahmad P, Rehman A, Alam S (2022a) Pathways to Argentina's 2050 carbon - neutrality agenda : the roles of renewable energy transition and trade globalization. Environ Sci Pollut Res 2022:29949–29966. https://doi.org/10.1007/s11356-021-17903-7

Murshed M, Uzma Khan, Aarif Mohammad Khan IO (2022b) Can energy productivity gains harness the carbon dioxide-inhibiting agenda of the Next 11 countries? Implications for achieving sustainable development. Sustain Dev -Wiley.1–14. https://doi.org/10.1002/sd.2393

Murshed M (2020a) Revisiting the deforestation-induced EKC hypothesis: the role of democracy in Bangladesh. GeoJournal 1–22. https://doi.org/10.1007/s10708-020-10234-z

Murshed M (2021) LPG consumption and environmental Kuznets curve hypothesis in South Asia : a time-series ARDL analysis with multiple structural breaks 8337–72. https://doi.org/10.1007/s11356-020-10701-7

Mutascu M (2016) A bootstrap panel Granger causality analysis of energy consumption and economic growth in the G7 countries. Renew Sustain Energy Rev 63:166–171. https://doi.org/10.1016/j.rser.2016.05.055

Nathaniel SP, Ahmad P (2021) The roles of nuclear energy, renewable energy, and economic growth in the abatement of carbon dioxide emissions in the G7 countries. Environ Sci Pollut Res 28:47957–47972. https://doi.org/10.1007/s11356-021-13728-6

Nazlioglu S, Gormus NA, Soytas U (2016) Oil prices and real estate investment trusts (REITs): gradual-shift causality and volatility transmission analysis. Energy Economics 60:168–175. https://doi.org/10.1016/j.eneco.2016.09.009

Nondo C, Kahsai MS (2020) The impact of energy intensity, urbanization, industrialization, and income on CO2 emissions in South Africa: an ARDL bounds testing approach. Afr J Econ Sustain Dev 7(4):307–330

Odugbesan JA, Adebayo TS (2020) Modeling CO2 emissions in South Africa: empirical evidence from ARDL based bounds and wavelet coherence techniques. Environ Sci Pollut Res 28(8):9377–9389. https://doi.org/10.1007/s11356-020-11442-3

Ozatac N, Gokmenoglu KK, Taspinar N (2017) Testing the EKC hypothesis by considering trade openness, urbanization, and financial development: the case of Turkey. Environ Sci Pollut Res 24(20):16690–16701

Ozcan B, Tzeremes PG, Tzeremes NG (2020) Energy consumption, economic growth and environmental degradation in OECD countries. Econ Model 84:203–213. https://doi.org/10.1016/j.econmod.2019.04.010

Ozturk I, Aslan A, Altinoz B (2022) Investigating the nexus between CO 2 emissions, economic growth, energy consumption and pilgrimage tourism in Saudi Arabia. Econ Res Istraživanja. 35(1):3083–98. https://doi.org/10.1080/1331677X.2021.1985577

Ozturk S, Cetin M, Demir H (2022b) Income inequality and CO 2 emissions : nonlinear evidence. Environ Dev Sustain 24(10):11911–11928. https://doi.org/10.1007/s10668-021-01922-y

Pablo-Romero, del Maria P, De Jesús Josué (2016) Economic growth and energy consumption: the energy-environmental Kuznets curve for Latin America and the Caribbean. Renewable and Sustainable Energy Reviews 60:1343–50. https://doi.org/10.1016/j.rser.2016.03.029

Pal D, Mitra SK (2017) Time-frequency contained co-movement of crude oil and world food prices: a wavelet-based analysis. Energy Economics 6(2):230–239. https://doi.org/10.1016/j.eneco.2016.12.020

Pan B, Adebayo TS (2022) Does nuclear energy consumption mitigate carbon emissions in leading countries by nuclear power consumption ? Evidence from quantile causality approach. Energy & Environ.1–23. https://doi.org/10.1177/0958305X221112910

Pata UK (2018) Renewable energy consumption, urbanization, financial development, income and CO_2 emissions in Turkey: testing EKC hypothesis with structural breaks. J Clean Prod 187:770–779. https://doi.org/10.1016/j.jclepro.2018.03.236

Pata UK, Aydin M (2020) Testing the EKC hypothesis for the top six hydropower energy-consuming countries: evidence from Fourier bootstrap ARDL procedure. Journal of Cleaner Production 264:121699. https://doi.org/10.1016/j.jclepro.2020.121699

Pata UK, Kumar A (2021) The influence of hydropower and coal consumption on greenhouse gas emissions: a comparison between China and India. Water 13(10):1387. https://doi.org/10.3390/w13101387

Rafiq S, Salim R, Apergis N (2016) Agriculture, trade openness and emissions: an empirical analysis and policy options. Australian Journal of Agricultural and Resource Economics 60(3):348–365. https://doi.org/10.1111/1467-8489.12131

Rahman MM, Vu XB (2020) The nexus between renewable energy, economic growth, trade, urbanization and environmental quality: a comparative study for Australia and Canada. Renew Energy 155(617–627):v

Ramzan M, Adebayo TS (2022) Do green innovation and financial globalization contribute to the ecological sustainability and energy transition in the United Kingdom ? Policy insights from a bootstrap rolling window approach.1–22.: https://doi.org/10.1002/sd.2399

Raslavicius L (2014) Arturas Kersys, Saulius Mockus, Neringa Kersienė, Martynas Starevicius (2014), Liquefied petroleum gas (LPG) as a medium-term option in the transition to sustainable fuels and transport. Renew Sustain Energy Rev 32:513–525. https://doi.org/10.1016/j.rser.2014.01.052

Rehman E, Ikram M, Feng MT, Rehman S (2020) Sectoral-based CO2 emissions of Pakistan: a novel grey relatioN ANALYSIS (GRA) approach. Environ Sci Pollut Res 27:29118–29129

Rehman A, Ma H, Ozturk I, Ulucak R. (2022a) Sustainable development and pollution : the effects of CO2 emission on population growth, food production, economic development, and energy consumption in Pakistan. Environ Sci Pollut Res. 2022a:17319–30: https://doi.org/10.1007/s11356-021-16998-2

Rehman A, Mahtab M, Ilhan A, Alvarado R, Murshed M, Işık C(2022b b). Globalization and renewable energy use : how are they contributing to upsurge the CO 2 emissions ? A global perspective. Environ Sci Pollut Res. 2022b:(0123456789):1–14. https://doi.org/10.1007/s11356-022-22775-6

Saidi K, Mbarek MB (2017,), The impact of income, trade, urbanization, and financial development on CO 2 emissions in 19 emerging economies. Environ Sci Pollut Res 24(14):12748–12757. doi: https://doi.org/10.1007/s11356-016-6303-3

Salahuddin M, Alam K, Ozturk I, Sohag K (2018) The effects of electricity consumption, economic growth, financial development and foreign direct investment on CO2 emissions in Kuwait. Renew Sust Energ Rev 81:2002–2010. https://doi.org/10.1016/j.rser.2017.06.009

Salim R, Rafiq S, Shafiei S, Yao Y (2019) Does urbanization increase pollutant emission and energy intensity? Evidence from some Asian developing economies. Appl Econ 51(36):4008–4024. https://doi.org/10.1080/00036846.2019.1588947

Shahbaz M, Tiwari AK, Nasir M (2013) The effects of financial development, economic growth, coal consumption and trade openness on CO2 emissions in South Africa. Energy Policy 61:1452–1459. https://doi.org/10.1016/j.enpol.2013.07.006

Shahbaz M, Loganathan N, Muzaffar AT, Ahmed K, Jabran MA (2016) How urbanization affects CO2 emissions in Malaysia?

The application of STIRPAT model. Renew Sust Energ Rev 57:83–93. https://doi.org/10.1016/j.rser.2015.12.096

Shahbaz M, Balsalobre-Lorente D, Sinha A (2019) Foreign direct investment–CO2 emissions nexus in Middle East and North African countries: importance of biomass energy consumption. J Clean Prod 217:603–614. https://doi.org/10.1016/j.jclepro.2019.01.282

Shi Q, Alvarado R, Cuesta L, Tillaguango B (2022) Asymmetric impacts of foreign direct investment inflows, financial development, and social globalization on environmental pollution. Econ Anal Policy 76:236–51. https://doi.org/10.1016/j.eap.2022.08.008

Shittu WO, Musibau H, Hassan S (2018) Revisiting the environmental Kuznets curve in Malaysia: the interactive roles of deforestation and urbanisation. Int J Green Econ 12(3–4):272–293. https://doi.org/10.1504/IJGE.2018.097872

Sinaga O (2019) The impact of hydropower energy on the envi- ronmental Kuznets curve in Malaysia. Digitales Archiv. http://zbw.eu/econis-archiv/bitstream/11159/2735/1/1046870351.pdf

Singh A, Kumari S, Singh B, Kumar N (2022) Investigating the economic and environmental sustainability of logistic operations in India using ARDL procedure. In Advanced Production and Industrial Engineering (pp. 81–90). IOS Press. https://doi.org/10.3233/ATDE220725

Solarin SA, Ozturk I (2015) On the causal dynamics between hydroelectricity consumption and economic growth in Latin America countries. Renew Sustain Energy Rev 52:1857–1868. https://doi.org/10.1016/j.rser.2015.08.003

Teng JZ, Khan MK, Khan MI, Chishti MZ, Khan MO (2021) Effect of foreign direct investment on CO 2 emission with the role of globalization, institutional quality with pooled mean group panel ARDL. Environ Sci Pollut Res 28(5):5271–5282. https://doi.org/10.1007/s11356-020-10823-y

Tiwari AK, Kocoglu M, Banday UJ, Awan. (2022) Hydropower, human capital, urbanization and ecological footprints nexus in China and Brazil: evidence from quantile ARDL. Advance online publication, Environmental Science and Pollution Researc. https://doi.org/10.1007/s11356-022-20320-z

Toda HY, Yamamoto (1995) Statistical inference in Vector Autoregressions with possibly integrated processes. Journal of Econometrics 66:225–250. https://doi.org/10.1016/0304-4076(94)01616-8

Torrence and Compo (1998) A practical guide to wavelet analysis. Bulletin of the American Meteorological Society 79(1):61–78. https://doi.org/10.1175/1520-0477(1998)079<0061:APGTWA>2.0.CO;2

Torrence C, Webster PJ (1998) The annual cycle of persistence in the El Nino/Southern Oscillation. Q J R Meteorol Soc 124:1985–2004

Tsong CC, Lee C-F, Tsai L-J, · Te-Chung Hu, (2016) The Fourier approximation and testing for the null of cointegration. Empir Econ 51:1085–1113. https://doi.org/10.1007/s00181-015-1028-6

U.S. Energy Information Administration (2022) International Energy Statistics, as of July 11. https://www.eia.gov/international/data/world/electricity/electricity-capacity?pd=2&p=0000000000000000000060000000c&u=0&f=A&v=mapbubble&a=-&i=none&vo=value&t=C&g=00000000000000000000000000000000000000000000000001&l=249-ruvvvvvfvtvnvv1vrvvvvfvvvvvvfvvvou20evvvvvvvvvvvnvvvs0008&s=315532800000&e=1577836800000&,nuclearenergy. Accessed 12 Aug 2022

Ummalla M, Samal A (2018) The impact of hydropower energy consumption on economic growth and CO_2 emissions in China. Environ Sci Pollut Res 25(35):35725–35737. https://doi.org/10.1007/s11356-018-3525-6

Wang Z, Ahmed Z, Zhang B, Wang B (2019) The nexus between urbanization, road infrastructure, and transport energy demand: empirical evidence from Pakistan. Environ Sci Pollut Res 26(34):34884–34895. https://doi.org/10.1007/s11356-019-06542-8

Wang Z, Rasool Y, Zhang B, Ahmed Z, Wang B (2020) Dynamic linkage among industrialisation, urbanisation, and CO2 emissions in APEC realms: evidence based on DSUR estimation. Struct Chang Econ Dyn 52:382–389. https://doi.org/10.1016/j.strueco.2019.12.001

Wang C, Ali S, Sunday T, Yi S, Ibrahim M (2023) (2023), The roles of hydro, nuclear and biomass energy towards carbon neutrality target in China : a policy-based analysis. Energy 262(PA):125303. https://doi.org/10.1016/j.energy.2022.125303

Wasti SKA, Zaidi SW (2020) An empirical investigation between CO2 emission, energy consumption, trade liberalization and economic growth: a case of Kuwait. Journal of Building Engineering 28:101104. https://doi.org/10.1016/j.jobe.2019.101104

World Bank (2020) World development indicators. The World Bank. Available at: https://databank.worldbank.org/source/worlddevelopment-indicators#

Wu Y, Zhu Q, Zhu B (2018) Decoupling analysis of world economic growth and CO2 emissions: a study comparing developed and developing countries. J Clean Prod 190:94–103

Wu L, Adebayo TS, Yue X, Umut A, Wu L, Adebayo TS (2022) (2022), The role of renewable energy consumption and financial development in environmental sustainability: implications for the Nordic Countries. Int J Sustain Dev World Ecol 00(00):1–16. https://doi.org/10.1080/13504509.2022.2115577

Yang Z, Zhao Y (2014) Energy consumption, carbon emissions, and economic growth in India: evidence from directed acyclic graphs. Econ Model 38:533–540. https://doi.org/10.1016/j.econmod.2014.01.030

Yasin TG, Sehnem E (2018) Environmental Kuznets curve for CO2 emissions: an analysis for developing, Middle East, OECD and OPEC countries. Environmental and Socio-Economic Studies-Sciendo 6(4):48–58. https://doi.org/10.2478/environ-2018-0027

Zhang J (2021) Environmental Kuznets curve hypothesis on CO2 emissions: evidence for China. Journal of Risk and Financial Management 14(1–16):93. https://doi.org/10.3390/jrfm14030093

Zhang Q, Adebayo TS, Ibrahim RL (2022) Do the asymmetric effects of technological innovation amidst renewable and nonrenewable energy make or mar carbon neutrality targets ? Int J Sustain Dev World Ecol 00(00):1–13. https://doi.org/10.1080/13504509.2022.2120559

Zhou G, Li H, Ozturk I, Ullah S (2022) Shocks in agricultural productivity and CO2 emissions : new environmental challenges for China in the green economy. Econ Res Istraživanja 0(0):1–17. https://doi.org/10.1080/1331677X.2022.2037447

Zmami M, BenSalha O (2020) An empirical analysis of the determinants of CO2 emissions in GCC countries. Int J Sust Dev World Ecol 1–12:469–480. https://doi.org/10.1080/13504509.2020.1715508

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.

Running vacuum in Brans & Dicke theory: a possible cure for the σ_8 and H_0 tensions

Javier de Cruz Pérez[2,1] Joan Solà Peracaula[1] and C. P. Singh[3]
[1] *Departament de Física Quàntica i Astrofísica, and Institute of Cosmos Sciences, Universitat de Barcelona, Av. Diagonal 647*
[2] *Department of Physics, Kansas State University, 116 Cardwell Hall, Manhattan, KS 66506, USA*
[3] *Department of Applied Mathematics, Delhi Technological University, Bawana Road, New Delhi-110042, India**
(Dated: February 10, 2023)

Extensions of the gravitational framework of Brans & Dicke (BD) are studied by considering two different scenarios: i) 'BD-ΛCDM', in which a rigid cosmological constant, Λ, is included, thus constituting a BD version of the vanilla concordance ΛCDM model (the current standard model of cosmology with flat three-dimensional geometry), and ii) 'BD-RVM', a generalization of i) in which the vacuum energy density (VED), $\rho_{\rm vac}$, is a running quantity evolving with the square of the Hubble rate: $\delta\rho_{\rm vac}(H) \propto \nu\, m_{\rm Pl}^2(H^2 - H_0^2)$ (with $|\nu| \ll 1$). This dynamical scenario is motivated by recent studies of quantum field theory (QFT) in curved spacetime, which lead to the running vacuum model (RVM). In both cases, rigid or running $\rho_{\rm vac}$, the GR limit can be recovered smoothly. We solve the background as well as the perturbation equations for each cosmological model and test their performance against the modern wealth of cosmological data, namely a compilation of the latest SNIa+$H(z)$+BAO+LSS+CMB observations. We utilize the AIC and DIC statistical information criteria in order to determine if they can fit better the observations than the concordance model. The two BD extensions are tested by considering three different datasets. According to the AIC and DIC criteria, both BD extensions i) and ii) are competitive, but the second one (the BD-RVM scenario) is particularly favored when it is compared with the vanilla model. This fact may indicate that the current observations favor a mild dynamical evolution of the Newtonian coupling G_N as well as of the VED. This is in agreement with recent studies suggesting that the combination of these two features can be favorable for a possible resolution of the σ_8 and H_0 tensions. In this work, we show that the Brans & Dicke theory with running vacuum has the potential to alleviate the two tensions at the same time.

arXiv:2302.04807v1 [astro-ph.CO] 9 Feb 2023

I. INTRODUCTION

The so-called concordance model of cosmology, or standard ΛCDM model, has been a rather successful paradigm for the phenomenological description of the universe for more than three decades [1], but it became consolidated only in the late nineties [2]. Crucial ingredients of it, however, such as the hypothetical existence of dark matter (DM), still lack of direct observational evidence. No less worrisome is the theoretical status of the cosmological constant (CC) term, Λ, in Einstein's equations. Despite it is sustained by general covariance, nothing is firmly known about its possible physical origin. The roots of the problem most likely reside in its interpretation as a quantity which is connected with the vacuum energy density (VED), $\rho_{\rm vac}$, a fundamental concept in Quantum Field Theory (QFT). In fact, for more than half of a century, the notion of vacuum energy in cosmology has been a most subtle concept which has challenged theoretical physicists and cosmologists, especially with the advent of Quantum Theory and in general with the development of the more sophisticated conceptual machinery of QFT. The proposed connection between the two quantities is well known: $\rho_{\rm vac} = \Lambda/(8\pi G_N)$, where G_N is Newton's coupling, usually assumed to be constant. Consistent measurements of Λ (treating it as a fit parameter) made independently in the last twenty years using distant type Ia supernovae (SNIa) and the anisotropies of the cosmic microwave background (CMB), have put the foundations of the concordance ΛCDM model of cosmology [3–10].

The above mentioned cosmological observations have indeed provided strong evidence for a spatially flat and accelerating universe in the present time. The ultimate cause of the acceleration is currently unknown, but it is attributed to an energy component which is popularly called "dark energy" (DE). The DE constitutes $\sim 70\%$ of the total energy density of the universe and presumably possessing negative pressure, thereby capable of producing the observed accelerating universe. The very nature of the DE still remains a complete mystery. The simplest candidate is the aforementioned cosmological term, Λ, usually assumed constant and that is why is called the cosmological constant [11, 12]. A flat Friedmann-Lemaître-Robertson-Walker (FLRW) model with a cosmological term (ΛCDM) fits rather accurately the current observational data. However, this model traditionally suffers from the so-called cosmological constant problem [13], which manifests itself in a dual manner: the fine-tuning problem associated to the value of Λ ('the old CC

* decruz@phys.ksu.edu, sola@fqa.ub.edu, cpsingh@dce.ac.in

problem") and what is called the 'coincidence problem', see for instance [14, 15] for a discussion on the various aspects of the CC problem. In the literature, many other cosmological models have been proposed to describe the DE. In particular, models with time-dependent vacuum energy density seem to be promising, since $\rho_{\rm vac}(t)$ could have had a high enough value in the very early universe to drive inflation, and subsequently decay along the expansion history to its small value observed today. Many of these models, however, are of pure phenomenological nature since the time dependence of $\Lambda(t)$ is parameterized in a totally ad hoc manner and having no obvious connection with any fundamental theory, see e.g. [16–31] and the old review [32]. In many cases is to go testing since the parameterization is performed through a direct function of the cosmic time or the scale factor, sometimes also as a function of the Hubble parameter, or even some hybrid combination of these possibilities.

However, recent calculations involving the renormalization of quantum field theory in FLRW spacetime yield a time-varying Λ, and hence a time-varying vacuum as well, in which Λ acquires a dynamical component through quantum effects: $\Lambda \to \Lambda + \delta\Lambda$. The latter evolves through the Hubble rate $H(t)$ as follows: $\delta\Lambda \sim \nu\,(H^2 - H_0^2)$ ($|\nu| \ll 1$). This is the characteristic form of the "running vacuum model" (RVM). The connection of the latter with QFT can be motivated from semi-qualitative renormalization group arguments, see [14] and references therein. However, an explicit QFT calculation leading to that form has appeared only very recently [33–36].

Let us also mention a popular class of cosmological scalar field models which were proposed long ago to solve some aspects of the CC problem. They go under the name of quintessence, see [11] for a review and references. These models mimic a time-evolving Λ term through a cosmic scalar field and focused essentially on ameliorating the situation with the mentioned cosmic coincidence problem, viz. why $\rho_{\rm vac}^0$ is so close to today's value of the matter density, which falls rapidly with the expansion as $\rho_m \sim a^{-3}$. But again these models are essentially phenomenological in nature since we do not have a viable scalar field candidate to play the role of quintessence in a fundamental context.

All in all, such attempts in different directions have shown that the dynamical DE models may offer an interesting possibility not only to alleviate some aspects of the cosmological constant and coincidence problems, but also to modulate the processes of structure formation. A note of caution is in order here, though. Among the above proposals, the very large family of time-varying '$\Lambda(t)-$models" should not be confused with the restricted class of the RVMs mentioned above, in which the dynamical dependence of the VED is derived from the computation and proper renormalization of quantum effects in QFT in curved spacetime. As mentioned, these developments are actually quite recent [33–36]. Therefore, despite some existing confusion in the literature, the RVMs are to be understood in that (much more) restricted sense, hence closer to fundamental theory, see [14] for an introduction and [15] for a review of the recent developments. Let us also note that, apart from the mentioned QFT framework, a 'stringy' version of the RVMs is also available [37–40]. The potential dynamics of the cosmic vacuum is therefore well motivated from different fundamental perspectives, and this fact further enhances the interest for the current study.

In this work, indeed, we combine the modern viewpoint of the RVM to revisit the old Brans & Dicke (BD) framework [41] with the purpose of exploring the consequences of having an evolving gravitational coupling and at the same time a dynamical vacuum. The BD theory is the simplest extension/modification of Einstein's general theory of relativity (GR). In fact it is the prototype of scalar-tensor theory of gravitation based on Mach's principle and Dirac's large number hypothesis. Besides, the BD field ϕ is not of quantum origin. Rather it is classical in nature and hence can be expected to serve as a relevant candidate to play some role in the late-time evolution of the universe. In this theory, the gravitational constant G is associated with the BD field as $\phi(t) \sim 1/G(t)$ and is coupled to curvature. There is also a new coupling parameter $\omega_{\rm BD}$ in front of the BD kinetic term. It is generally assumed that GR is recovered in the limit $\omega_{\rm BD} \to \infty$, but we actually need also that $\phi(t)$ provides the (inverse) measured value of G_N at present ($t = t_0$). BD theory is in fact the first historical attempt to extend GR to accommodate a dynamical cosmic variable in the Newtonian coupling G, but at that time no cosmological term was included in the study. The interest on this kind of theory was renewed later on, in part owing to its association with superstrings theories, extra-dimensional theories and models with inflation or accelerated expansion [42–55]. In recent years, different studies [56–63] have incorporated the CC term in BD theory as a possibility for explaining the DE phenomena. It remains, of course, to see if realistic mechanisms can be devised that are capable of screening the modified gravity effects at local (astrophysical) scales where deviations from GR are are highly constrained see e.g. [64–67] and references therein. We concentrate here only on the cosmological consequences of BD-like theories of gravity.

Specifically, in this work we further dwell upon the BD theory with a cosmological term and investigate in detail two basic cosmological scenarios with flat FLRW spacetime: i) First we reconsider the BD theory with a rigid CC, Λ, along the lines of previous studies such as [58, 61] and [63]. This will serve as a fiducial BD scenario maximally close to the ΛCDM. It obviously represents a BD generalization of it, which is why we call it BD-ΛCDM; ii) The second scenario adds up more dynamics to the first one along the lines of the RVM, that is to say, we admit that the VED evolves with the cosmic expansion as $\delta\rho_{\rm vac} \propto \nu\, m_{\rm Pl}^2 (H^2 - H_0^2)$ (with $|\nu| \ll 1$), just as found in the QFT framework of [33–35]. In this way we can explore the BD theory (being the canonical scalar-tensor framework possessing a mild time-evolving gravitational

coupling) by further equipping it with a smooth dynamical vacuum component [68]. We perform the joint statistical analysis using the latest observational data including Type Ia Supernova data, Hubble data, baryon acoustic oscillation data, large scale structure (LSS) formation data and cosmic microwave background data.

At this point a very much practical motivation for studying the BD-RVM framework must be stood up. Let us recall that apart from the aforementioned theoretical aspects of the cosmological constant problem, it has long been known that there appear to exist potentially serious discrepancies or "tensions" between the CMB observations from the Planck Collaboration, which are based on the vanilla ΛCDM, and the local direct (distance ladder) measurements of the Hubble parameter today [69]. This brings about the so called H_0 tension, which leads to a severe discrepancy at $\sim 5\sigma$ c.l. Among many possibilities debated in the literature, it has been suggested e.g. that a possible intrinsic evolutionary behavior of $H_0(z)$ could be underlying this tension [70], in which case it would not be a discrepancy between local and CMB data, but an effect potentially observable at any redshift. In addition, there exist a smaller but appreciable and persistent ($\sim 2-3\,\sigma$) tension in the realm of the large scale structure (LSS) growth data, the σ_8 tension [71]. Specifically, it is concerned with the measurements of weak gravitational lensing at low redshifts ($z < 1$). Such a tension is gauged through the parameter S_8 or, alternatively, by means of σ_8; recall that $S_8 \equiv \sigma_8\sqrt{\Omega_m^0/0.3}$. It turns out that these measurements favor weaker matter clustering than that expected from the concordance cosmological model using parameters determined by CMB measurements. For detailed reviews on these tensions and other challenges afflicting the concordance ΛCDM model, see e.g. [72–74] and the long list of references therein.

It has been known for quite some time in the literature that the RVM-type of cosmological models can help in improving the overall fit to the cosmological observations and also in smoothing out these tensions as compared to the ΛCDM. Within the GR-RVM (viz. running vacuum in GR) it was investigated for instance in [75–82], see [83, 84] for a summary. When we move to BD theory with a cosmological constant, i.e. the BD-ΛCDM, the model turns out to mimic the RVM and improves the fits as well [58, 61, 85]. It is therefore natural to keep on exploring the RVM framework and perform the next step. Thus, in the present paper we address for the first time the running vacuum in the BD context, i.e. the BD-RVM scenario, and check if further improvement (relieving) of the tensions can be attained as compared to the BD-ΛCDM. We find indeed a significant improvement of the current status of the mentioned tensions.

The paper is organized as follows. In Sec. II, we present the models and the corresponding field equations within the context of Brans & Dicke theory. Section III is devoted to discuss the effective equation of state (EoS) of the extended BD models. Being the large scale structure (LSS) growing data indispensable for a proper discussion of the σ_8 tension, in Sec. IV we describe our treatment of the matter perturbations. In Section V we enumerate and briefly describe the different sources of observational data used in this paper and the method employed to constraint the free parameters of the models. At the same time we define the three datasets that will be used throughout our comparative model analysis. The results are discussed in detail in Section VI. Finally, in Section VII we summarize our findings and present the main conclusions of this work.

II. BD FIELD EQUATIONS WITH DYNAMICAL VACUUM

We consider a spatially flat homogeneous and isotropic universe described by the FLRW metric

$$ds^2 = -dt^2 + a^2(t)\left[dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2)\right], \quad (1)$$

where (r, θ, ϕ) are defined as the comoving coordinates, t denotes the cosmic time and $a(t)$ is the scale factor of the universe. Throughout this work we consider the natural units convention, namely: $\hbar = c = 1$. The gravity theory of Brans & Dicke [41], contains and additional *d.o.f.* with respect to General Relativity, which is encoded in the BD-field denoted by ϕ. The BD-action, written in the Jordan frame and in the presence of the vacuum energy density $\rho_{\rm vac}$ takes the following form:

$$S_{\rm BD} = \int d^4x\sqrt{-g}\Big[\frac{1}{16\pi}\left(\phi R - \frac{\omega_{\rm BD}}{\phi}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi\right) - \rho_{\rm vac} + \mathcal{L}_m\Big]. \quad (2)$$

The scalar field, which is minimally coupled to the curvature, has dimension 2 in natural units and it represents, at any time, the inverse value of the Newtonian gravitational coupling $\phi(t) = 1/G(t)$. The dimensionless parameter in front of the kinetic term, $\omega_{\rm BD}$, is the BD-parameter and $\mathcal{L}_m$ contains the contribution from the matter fields.

Variation of the action (2) with respect to the metric $g_{\mu\nu}$ and the BD-scalar field ϕ, respectively, yields the field equations:

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = \frac{8\pi}{\phi}\left(\tilde{T}_{\mu\nu} + T^{\rm BD}_{\mu\nu}\right), \quad (3)$$

and

$$\Box\phi = \frac{8\pi}{(2\omega_{\rm BD}+3)}\tilde{T}. \quad (4)$$

In the above equations,

$$\tilde{T}_{\mu\nu} = pg_{\mu\nu} + (\rho + p)U_\mu U_\nu \quad (5)$$

is the total energy-momentum tensor, with $\tilde{T} \equiv \tilde{T}^\mu_\mu$ its trace. In it, $\rho \equiv \rho_m + \rho_r + \rho_{\rm vac}$ and $p \equiv p_m + p_r + p_{\rm vac}$

represent the total (proper) density and pressure, whose matter and vacuum components satisfy the usual equations of state $p_m = 0$, $p_r = (1/3)\rho_r$ and $p_{\rm vac} = -\rho_{\rm vac}$. The term ρ_m involves the contribution from cold dark matter (CDM) and baryons, whereas ρ_r is the relativistic component from photons and massless neutrinos. Furthermore, in Eq.(3), the piece $T_{\mu\nu}^{\rm BD}$ stands for the remaining term that emerges from the metric functional variation of the action (2) and its structure depends entirely on ϕ and its derivatives. Such a piece defines the energy-momentum tensor of the BD-field:

$$T_{\mu\nu}^{\rm BD} = \frac{1}{8\pi}\left[\frac{\omega_{\rm BD}}{\phi}\left(\nabla_\mu\phi\nabla_\nu\phi - \frac{1}{2}g_{\mu\nu}\nabla_\alpha\phi\nabla^\alpha\phi\right) + (\nabla_\mu\nabla_\nu\phi - g_{\mu\nu}\nabla_\alpha\nabla^\alpha\phi)\right]. \tag{6}$$

For ϕ =const. the above energy-momentum tensor vanishes, and when such a constant is precisely $\phi = 1/G_N$, Eq. (3) boils down to the ordinary Einstein's field equations, i.e. we recover GR. Here $G_N = 1/m_{\rm Pl}^2$ is the measured value of Newton's constant, with $m_{\rm Pl} \simeq 1.2 \times 10^{19}$ GeV the ordinary Planck mass. Notice, however, that on implementing the GR limit it is mandatory to set also $\omega_{\rm BD} \to \infty$ such that the wave equation for the BD-field (4) consistently decouples at the same time. For variable $\phi(t)$, we must recover too the measured gravity value $1/\phi(t_0) = G(t_0) = G_N$ at the present cosmic time t_0, of course, but $G(t) \equiv 1/\phi(t)$ can be slightly different from G_N at other times. Such small difference, however, can impinge positively on our analysis.

In the framework of the FLRW metric (1), the cosmological equations that govern the dynamics of the universe within BD gravity can be written as follows:

$$3H^2 + 3H\frac{\dot\phi}{\phi} - \frac{\omega_{\rm BD}}{2}\frac{\dot\phi^2}{\phi^2} = \frac{8\pi}{\phi}\rho \tag{7}$$

$$2\dot H + 3H^2 + \frac{\ddot\phi}{\phi} + 2H\frac{\dot\phi}{\phi} + \frac{\omega_{\rm BD}}{2}\frac{\dot\phi^2}{\phi^2} = -\frac{8\pi}{\phi}p \tag{8}$$

$$\ddot\phi + 3H\dot\phi = \frac{8\pi}{(2\omega_{\rm BD}+3)}(\rho - 3p) \tag{9}$$

where the overdots denote derivatives with respect to the cosmic time t and $H = \dot a/a$ is the Hubble parameter. Once more we can check that when we set $\phi = 1/G_N$ *and* at the same time $\omega_{\rm BD} \to \infty$, we do in fact retrieve the GR limit, which in this case is obtained within the FLRW metric. Indeed, in the mentioned limit we find that equations (7) and (8) render the standard Friedmann and pressure equations respectively, whereas (9) consistently decouples.

No interaction between the BD-field and the ordinary matter and vacuum is considered, so from the Bianchi identity we get $\nabla^\mu \tilde T_{\mu\nu} = 0$. This can be explicitly verified with the above field equations. We assume that baryons and radiation (photons and relativistic neutrinos) are covariantly conserved separately, which leads to the usual local energy conservation equations, namely:

$$\dot\rho_b + 3H\rho_b = 0 \tag{10}$$

and

$$\dot\rho_r + 4H\rho_r = 0\,. \tag{11}$$

It should be mention at this point that despite neutrinos do not behave as pure radiation in the late universe, the ratio between the energy density of non-relativistic neutrinos to the total energy density of non-relativistic matter (baryons plus CDM) remains very small throughout the entire cosmic history up to our time (below 10^{-3}) [75], and therefore for our purposes in this paper we can safely treat neutrinos as if being relativistic all the time.

In contradistinction to baryons and radiation, cold dark matter need not be conserved in general and it may interact with vacuum, similarly as in the analysis of the GR-RVM performed in [82]. So for CDM we have, in general, the following modified conservation equation:

$$\dot\rho_{\rm cdm} + 3H\rho_{\rm cdm} = -\dot\rho_{\rm vac}\,. \tag{12}$$

Notice that in the case of the BD-ΛCDM the *r.h.s.* of Eq. (12) is zero and the CDM is conserved. Not so, of course, when $\rho_{\rm vac}$ changes with the cosmological expansion, as e.g. in the BD-RVM context. Both options will be considered separately in the next sections.

Unfortunately it is not possible to obtain analytical solutions for the system of differential equations (7)-(8)-(9). However, it is reasonable to assume that the evolution of the BD-field $\phi(a)$ should be close around the present value $\phi(a=1) = 1/G_N$, as otherwise the departures of $1/\phi(a)$ from G_N would be significant and would have been detected experimentally. Bearing this in mind, it is meaningful to search for solutions of the field equations in power-law form. We proceed as in [56, 57] with the following ansatz:

$$\phi(a) = \phi_0 a^{-\epsilon} = \frac{1}{G_N}a^{-\epsilon} \quad (|\epsilon| \ll 1)\,, \tag{13}$$

where ϵ is a constant whose value should be small in order to be consistent with the measurements of the Newtonian coupling. For $\epsilon > 0$, ϕ decreases (and hence Newton's coupling increases) with the expansion and this means

that it has an asymptotically free behavior in the past (where it lies the most energetic epoch of the universe) since it becomes smaller. For $\epsilon < 0$, the opposite occurs, that is to say, in such case the gravitational coupling decreases with the expansion.

If we insert (13) into the set of equations (7)-(8)-(9) we arrive at the equations

$$H^2 = \tfrac{8\pi}{3}\tfrac{1}{\beta\phi}(\rho_m + \rho_r + \rho_{\rm vac}) \tag{14}$$

$$\dot{H}(2-\epsilon) + H^2\left(3 - 2\epsilon + \epsilon^2 + (1/2)\omega_{\rm BD}\epsilon^2\right) \tag{15}$$
$$= -\tfrac{8\pi}{\phi}\left(\tfrac{\rho_r}{3} - \rho_{\rm vac}\right)$$

$$-\epsilon\dot{H} + (-3\epsilon + \epsilon^2)H^2 = \tfrac{8\pi}{2\omega_{\rm BD}+3}\tfrac{1}{\phi}(\rho_m + 4\rho_{\rm vac}). \tag{16}$$

We recall that $\rho_m = \rho_b + \rho_{\rm cdm}$. In the above equations we have introduced the following definitions

$$\beta \equiv 1 - \nu_{\rm eff}, \qquad \nu_{\rm eff} \equiv \epsilon\left(1 + \frac{\omega_{\rm BD}}{6}\epsilon\right). \tag{17}$$

As in [57] we assume that $|\omega_{\rm BD}\epsilon| \sim 1$, consequently we have $|\omega_{\rm BD}\epsilon| > |\epsilon| \gg |\epsilon^2|$ and $|\omega_{\rm BD}\epsilon^2| \sim |\epsilon|$. Taking into account this hierarchy of orders of magnitudes, it is possible to get a relation between $\omega_{\rm BD}$, ϵ and the current values of the non-relativistic and radiation energy densities, see [57] for the details:

$$\omega_{\rm BD}\epsilon = -\frac{12 - 9\Omega_m^0 - 12\Omega_r^0}{6 - 3\Omega_m^0 - 4\Omega_r^0} + \mathcal{O}(\epsilon). \tag{18}$$

From the above expression it is clear that $|\omega_{\rm BD}\epsilon| \sim \mathcal{O}(1)$. The final numerical values are to be determined when the cosmological models under study are tested against the observational data considered. Hereafter we consider two different cosmological models both in the theoretical framework of the Brans & Dicke modified gravity. Additionally, we also provide the results for the standard model of cosmology which in this paper is denoted by GR-ΛCDM model.

A. BD-ΛCDM (constant vacuum)
$\rho_{\rm vac}(a) = \rho_{\rm vac}^0 = $ **const.**

In this scenario, vacuum is not in interaction with cold dark matter and the value of its energy density remains constant. It is important to remark that in the original formulation of Brans & Dicke [41] the cosmological term was not included so we may consider the BD-ΛCDM model as the minimal extension of the BD model in order to predict the observed accelerated expansion of the universe. See [58] and [61] for a detailed study of the BD-ΛCDM model. Considering the usual definitions for the mass parameter of each one of the species at present time $\Omega_i^0 = \rho_i^0/\rho_c^0$, where i runs for $m, {\rm vac}, r$ and $\rho_c^0 = 3H_0^2/8\pi G_N$ is the current critical energy density we can express the normalized Hubble function $E(a) \equiv H(a)/H_0$ as:

$$E^2(a) = \frac{a^\epsilon}{\beta}\left[\Omega_m^0 a^{-3} + \Omega_r^0 a^{-4} + \Omega_{\rm vac}^0\right]. \tag{19}$$

As expected the cosmic sum rule for Ω_i^0 does not have the usual form due to the presence of ϵ and $\omega_{\rm BD}$

$$\Omega_m^0 + \Omega_r^0 + \Omega_{\rm vac}^0 = \beta = 1 - \nu_{\rm eff}. \tag{20}$$

The GR-ΛCDM limit, $E^2(a) = \Omega_m^0 a^{-3} + \Omega_r^0 a^{-4} + \Omega_{\rm vac}^0$, is recovered by setting $\epsilon \to 0$ and $\beta \to 1$.

B. BD-RVM (running vacuum)
$\rho_{\rm vac}(H) = \frac{3\phi}{8\pi}\left(c_0 + \nu H^2\right)$

In this alternative scenario the cosmological term is no longer constant. It is nevertheless assumed to be essentially constant up to a dynamical component which evolves as the square of the Hubble rate:

$$\Lambda(H) = 3c_0 + 3\nu H^2\,, \tag{21}$$

where the coefficient 3 is for convenience. Thus, the vacuum energy density is $\rho_{\rm vac}(H) = \Lambda(H)\phi/(8\pi)$ and since $\phi = 1/G \sim m_{\rm Pl}^2$ it evolves as $\delta\rho_{\rm vac} \sim 3\,\nu\, m_{\rm Pl}^2(H^2 - H_0^2)$ when we compare its value at H with its current value at H_0. This form is characteristic of the RVM, as discussed previously. The dimensionless parameter ν must be fixed by comparing with observations and we expect it to have a small value ($|\nu| \ll 1$), typically $\nu \sim 10^{-3}$. This value is compatible with recent BBN bounds on modified gravity models [86]. Because the vacuum is slightly time-dependent now, we shall assume that it is in interaction with cold dark matter (CDM) through the generalized local conservation law (12), where now we have the specific RVM form for $\rho_{\rm vac}$. We can fix the value of the constant c_0 (of dimension 2 in natural units) by imposing the boundary condition $\rho_{\rm vac}(a=1) = \rho_{\rm vac}^0$ which leads us to the following equality $c_0 = H_0^2(\Omega_{\rm vac}^0 - \nu)$, where $\Omega_{\rm vac}^0$ is $\rho_{\rm vac}^0$ in units of the current critical density ρ_c^0. Notice that the vacuum in Eq.(21) is dominated by the constant term c_0, whilst νH^2 is a subdominant component. The fact of having $c_0 \neq 0$ is essential in order to remain close to the concordance model and hence be able to fit the large scale structure observational data comparatively well. The dimensionless coefficient ν can be interpreted as the β-function of the running of the vacuum at low energies and its theoretical value can be estimated in the context of QFT in curved space times see [33–35] and references therein. However, here we determine its value phenomenologically by testing the theoretical predictions of the BD-RVM cosmological model against the latest cosmological data.

A close formula for the normalized Hubble function in the BD-RVM framework can be explicitly derived using the equations of the previous section, with the following result:

$$E^2(a) = 1 + \tfrac{1}{\beta-\nu}\Big[\Omega_m^0(a^{-C_1+\epsilon} - 1) + \alpha(1 - a^{-C_1+\epsilon}) + \Omega_r^0\left(\tfrac{\beta-\nu}{\beta+3\nu}a^{-4+\epsilon} + \tfrac{4\nu}{\beta+3\nu}a^{-C_1+\epsilon} - 1\right)\Big]. \tag{22}$$

where to lighten the formulae we have introduced the following definitions

$$C_1 \equiv 3\left(\frac{\beta-\nu}{\beta}\right) \qquad \alpha = -\epsilon\frac{\Omega_{\rm vac}^0-\nu}{C_1-\epsilon}. \tag{23}$$

In this case the GR-ΛCDM limit is recovered by setting $\epsilon, \nu \to 0$ and $\beta \to 1$, which implies $C_1 = 3$ and $\alpha = 0$.

We should also stress that the dynamical character of the vacuum energy density through the square of the Hubble rate $\sim H^2$, as implied by the RVM form (21), is highly convenient since it is compatible with the general covariance of the effective action, as demonstrated also from the explicit calculations of [33–36]. In contrast, a time-dependence of the VED through a linear term $\sim H$ is not compatible with general covariance. This does not a model of this sort a priori, though, since a linear term can be admitted on phenomenological grounds and may be associated with e.g. dissipative effects produced by bulk viscosity [87, 88]. However, even on phenomenological grounds such linear dependence on H is not favored, see e.g. [47] and especially the recent analysis [63]. In point of fact, the difficulties with the linear term $\sim H$ as part of the VED have also been studied at length in the GR context, see [89], where it has been shown that it leads to serious problems with the description of the LSS data. In contrast, as we shall see here, the $\sim H^2$ component in the BD-RVM framework fits extremely well the cosmological observations. This turns out to be one of the main results of the present work.

The convenience of the quadratic component in the Hubble rate is in correspondence with the good results obtained for the RVM in the GR context [75]. The BD framework adds up to it the dynamics of the gravitational coupling, and this feature proves very helpful to further optimize the phenomenological performance of the model. We shall substantiate this assertion in detail throughout the present analysis.

III. EFFECTIVE EOS OF THE DARK ENERGY IN THE EXTENDED BD-FRAMEWORK

Along the lines of [61], we now rewrite the Brans-Dicke cosmological equations in an effective GR-picture, which is useful to parameterize the departure, at the background level, of the BD theory from the standard GR-ΛCDM model. Using the field equations (7)-(8)-(9) of BD gravity, we find that the Friedmann and pressure equation of the effective GR theory describing them take the following form:

$$3H^2 = 8\pi G_N\left(\rho_m+\rho_r+\rho_{\rm vac}+\rho_\varphi\right) \tag{24}$$

$$2\dot{H}+3H^2 = -8\pi G_N\left(p_r+p_{\rm vac}+p_\varphi\right) \tag{25}$$

where we have introduced the dimensionless BD-field

$$\varphi \equiv G_N\phi = a^{-\epsilon} \tag{26}$$

and used Eq. (13). Notice that in the above generalized Friedmann's and pressure equations we have encoded the non-GR terms of the BD theory as if they were a new fluid (the BD-fluid) characterized by the following energy density and pressure:

$$\rho_\varphi \equiv \frac{3}{8\pi G_N}\left(H^2\Delta\varphi - H\dot{\varphi}+\frac{\omega_{\rm BD}}{6}\frac{\dot{\varphi}^2}{\varphi}\right) \tag{27}$$

$$p_\varphi \equiv \frac{1}{8\pi G_N}\left(-3H^2\Delta\varphi-2\dot{H}\Delta\varphi+\ddot{\varphi}+2H\dot{\varphi}+\frac{\omega_{\rm BD}}{2}\frac{\dot{\varphi}^2}{\varphi}\right) \tag{28}$$

where we have defined

$$\Delta\varphi \equiv 1-\varphi\,. \tag{29}$$

According to (26), $\Delta\varphi = 0$ only in the current universe. One can check from the field equations that the effective BD-fluid (27)-(28) fulfills the energy conservation equation

$$\dot{\rho}_\varphi + 3H\left(p_\varphi+\rho_\varphi\right) = 0. \tag{30}$$

Once we have defined the above quantities we can go a step further and compute the effective equation of state (EoS) of the combined BD-fluid and the vacuum energy density (VED), namely:

$$w_{\rm eff}(z) \equiv \frac{p_{\rm vac}+p_\varphi}{\rho_{\rm vac}+\rho_\varphi} = -1+\frac{p_\varphi+\rho_\varphi}{\rho_{\rm vac}+\rho_\varphi}\,, \tag{31}$$

where we used the fact that $p_{\rm vac} = -\rho_{\rm vac}$. With the help of equations (27) and (28) we find

$$w_{\rm eff}(z) = -1+\frac{-2\dot{H}\Delta\varphi+f_1(\varphi,\dot{\varphi},\ddot{\varphi})}{\Lambda(H)(1-\Delta\varphi)+3H^2\Delta\varphi+f_2(\varphi,\dot{\varphi})} \tag{32}$$

where

$$f_1(\varphi,\dot{\varphi},\ddot{\varphi}) = \ddot{\varphi}-H\dot{\varphi}+\omega_{\rm BD}\tfrac{\dot{\varphi}^2}{\varphi} \tag{33}$$

$$f_2(\varphi,\dot{\varphi}) = -3H\dot{\varphi}+\tfrac{\omega_{\rm BD}}{2}\tfrac{\dot{\varphi}^2}{\varphi}. \tag{34}$$

We note that for the BD-ΛCDM model we must replace $\Lambda(H)(1-\Delta\varphi) \Rightarrow \Lambda = $ const. in (32) since for this model, in the expression of $\rho_{\rm vac}$, the term Λ is constant and $\varphi = 1$. The explicit form of the effective EoS in terms of the cosmological redshift, z, can be obtained upon combining (32)-(34) with equations (22) and (26). We refrain from writing out this formula, but in order to better appreciate the behaviour of $w_{\rm eff}(z)$ close to the present time we provide the following analytic approximation in terms of z, valid up to $\mathcal{O}(z^3)$:

$$w_{\rm eff}(z) = -1+\frac{\epsilon}{1-\Omega_m^0}\left[\frac{1}{3}+\frac{\Omega_m^0}{2}+\frac{\omega_{\rm BD}\epsilon}{3}\right.$$

$$\left.+\Omega_m^0\left(\frac{3}{2}+\omega_{\rm BD}\epsilon\right)z+\Omega_m^0\omega_{\rm BD}\epsilon z^2\right.$$

Baseline

Parameter	**GR-ΛCDM**	**BD-ΛCDM**	**BD-RVM**
ω_b	0.02213 ± 0.00020	0.02207 ± 0.00020	0.02201 ± 0.00022
n_s	0.9665 ± 0.0042	0.9639 ± 0.0044	0.9620 ± 0.0050
$10^9 A_s$	2.054 ± 0.033	2.081 ± 0.036	2.085 ± 0.036
H_0[km/s/Mpc]	68.56 ± 0.41	66.73 ± 0.81	$68.0^{+1.4}_{-1.2}$
Ω_m^0	0.3010 ± 0.0052	0.3031 ± 0.0053	0.3041 ± 0.0054
$\sigma_8(0)$	0.7997 ± 0.0071	0.758 ± 0.018	0.751 ± 0.019
S_8	0.801 ± 0.011	0.761 ± 0.019	0.768 ± 0.018
ϵ	-	0.0039 ± 0.0015	$-0.0016^{+0.0037}_{-0.0049}$
ν	-	-	$0.0019^{+0.0016}_{-0.0013}$
$\chi^2_{\rm min}$	71.83	64.68	63.15
ΔAIC	-	4.76	3.81
ΔDIC	-	5.06	5.10

TABLE I. The mean-fit values and the 68% confidence limits for the characterizing parameters of the cosmological models under study upon consideration of our Baseline dataset, which contains data from the string SNIa+$H(z)$+BAO+LSS+CMB. We provide the values of the parameters common to all the cosmological models, to wit: the baryonic density parameter $\omega_b = \Omega_b^0 h^2$, the spectral tilt parameter n_s and the amplitude parameter A_s of the primordial power spectrum, the Hubble parameter H_0 and the normalized matter density parameter Ω_m^0. Note that we do not consider the optical depth parameter τ since it is not sensitive to our analysis. In addition to the before mentioned parameters the BD-ΛCDM model contains an extra *d.o.f* contained in ϵ and the BD-RVM model has two extra parameters, ϵ and ν. We also display for all the models the value of the derived parameters $\sigma_8(0)$ and $S_8 \equiv \sigma_8\sqrt{\Omega_m^0/0.3}$. $\chi^2_{\rm min}$ denotes the minimum value of the χ^2-function and finally ΔAIC and ΔDIC represent the differences of the AIC and DIC values with respect to the GR-ΛCDM model.

Baseline+S_8

Parameter	**GR-ΛCDM**	**BD-ΛCDM**	**BD-RVM**
ω_b	0.02215 ± 0.00020	0.02207 ± 0.00020	0.02200 ± 0.00021
n_s	0.9670 ± 0.0042	0.9639 ± 0.0044	0.9613 ± 0.0047
$10^9 A_s$	2.046 ± 0.034	2.080 ± 0.035	2.085 ± 0.036
H_0[km/s/Mpc]	68.71 ± 0.40	66.65 ± 0.80	68.35 ± 0.71
Ω_m^0	0.2990 ± 0.0050	0.3026 ± 0.0052	0.3044 ± 0.0054
$\sigma_8(0)$	0.7970 ± 0.0071	0.754 ± 0.016	0.751 ± 0.019
S_8	0.796 ± 0.011	0.757 ± 0.017	0.756 ± 0.020
ϵ	-	0.0042 ± 0.0014	-0.0031 ± 0.0019
ν	-	-	0.00237 ± 0.00080
$\chi^2_{\rm min}$	74.42	65.02	63.17
ΔAIC	-	7.01	6.39
ΔDIC	-	7.61	6.59

TABLE II. As in Table I but adding as an input the $S_8 \equiv \sigma_8\sqrt{\Omega_m^0/0.3} = 0.737^{+0.040}_{-0.036}$ weak-lensing measurement from [90].

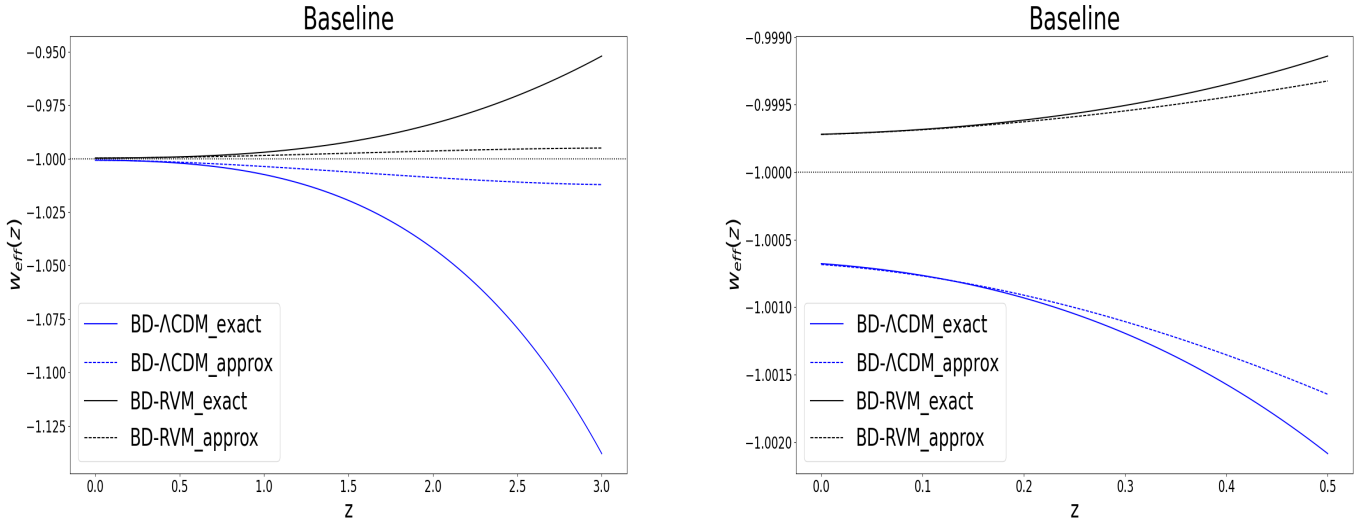
$$+\Omega_m^0\left(1+\frac{\omega_{\rm BD}\epsilon}{3}\right)z^3\Big] + \mathcal{O}\left(\epsilon^2, \epsilon\nu, z^4\right) . \qquad (35)$$

where we neglected the higher order powers of the parameters ϵ and ν beyond the linear order since they both satisfy $|\epsilon, \nu| \ll 1$. However, we keep the product $\omega_{\rm BD}\epsilon$ since it can be of order 1 for large enough $\omega_{\rm BD}$. Similarly $\omega_{\rm BD}\epsilon^2 \sim \epsilon$ is also retained as any term of $\mathcal{O}(\epsilon)$.

The effective EoS of the BD-fluid +VED evolves with the expansion, and when viewed from the GR standpoint it describes an effective dark energy (DE) inherent to the BD theory with cosmological term. This was studied in [58, 61] for the case of Λ =const., which is the situation with the model BD-ΛCDM being studied here, but in this work we extend the analysis also for the case BD-RVM,

	Baseline+H_0		
Parameter	**GR-ΛCDM**	**BD-ΛCDM**	**BD-RVM**
ω_b	0.02228 ± 0.00020	0.02228 ± 0.00020	0.02205 ± 0.00020
n_s	0.9693 ± 0.0042	0.9692 ± 0.0043	0.9619 ± 0.0045
$10^9 A_s$	2.058 ± 0.033	2.060 ± 0.035	2.078 ± 0.036
H_0[km/s/Mpc]	69.13 ± 0.38	69.02 ± 0.67	70.34 ± 0.63
Ω_m^0	0.2942 ± 0.0048	0.2942 ± 0.0048	0.2994 ± 0.0050
$\sigma_8(0)$	0.7970 ± 0.0072	0.795 ± 0.016	0.764 ± 0.018
S_8	0.789 ± 0.011	0.757 ± 0.017	0.763 ± 0.019
ϵ	-	0.0002 ± 0.0013	$-0.00767^{+0.00069}_{-0.00220}$
ν	-	-	$0.00332^{+0.00084}_{-0.00065}$
$\chi^2_{\rm min}$	87.97	87.95	70.68
ΔAIC	-	-2.37	12.43
ΔDIC	-	-2.46	12.10

TABLE III. The same as in Table I but considering the prior on $H_0 = 73.04 \pm 1.04$ km s^{-1} Mpc^{-1} from [91].



(a) (b)

FIG. 1. *(a) plot*: Effective equation of state parameter (32) for the BD-ΛCDM and the BD-RVM cosmological models (denoted by BD-ΛCDM$_{\rm exact}$ and BD-RVM$_{\rm exact}$ respectively) considering the results obtained with the Baseline set. For comparison we also depict the approximate expression (35) (denoted by BD-ΛCDM$_{\rm approx}$ and BD-RVM$_{\rm approx}$ respectively). *(b) plot*: Magnifies the region of *(a)* near our time.

which is when Λ is actually 'running' within the context of the RVM, see Eq. (21). The plots of the effective EoS (32) as a function of the redshift for the extended BD frameworks under study are depicted in Figs. 1 and 2.

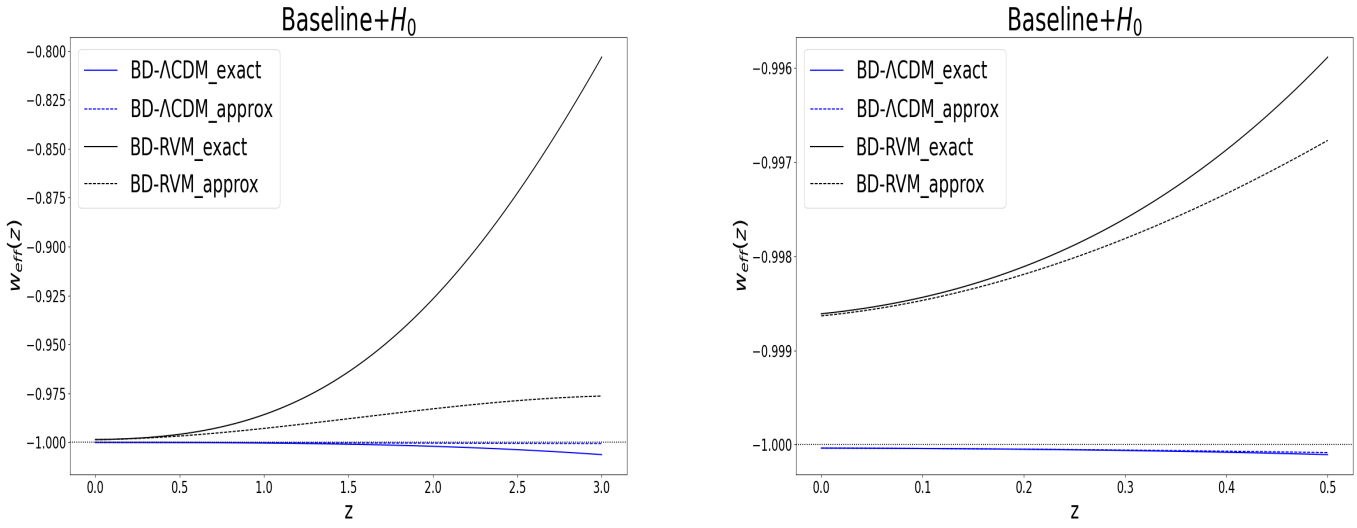
If we consider the above expression for today's universe ($z = 0$) and utilize the relation (18), without including the (currently very small) radiation contribution, we get

$$w_{\rm eff}(0) = -1 - \epsilon F(\Omega_m^0) + \mathcal{O}\left(\epsilon^2, \epsilon\nu\right) \tag{36}$$

with

$$F(\Omega_m^0) = \frac{4 - 10\,\Omega_m^0 + 3\Omega_m^{0^2}}{12 - 18\Omega_m^0 + 6\Omega_m^{0^2}}\,. \tag{37}$$

We point out the following interesting feature of the previous result: being $F(\Omega_m^0) > 0$ for the current values of the cosmological parameters, the behavior of the EoS at $z = 0$ turns out to be phantom-like or quintessence-like depending on the sign of ϵ. Specifically, we can see from



(a) (b)

FIG. 2. The same as in Fig. 1 but considering the results obtained with the Baseline+H_0 set.

Figs. 1 and 2 that the BD-RVM mimics quintessence (since $\epsilon < 0$ for it in all fitting scenarios considered here, see Tables I, II and III), whereas BD-LCDM mimics phantom-like behavior (since $\epsilon > 0$ for this model in the same fitting tables). It is worth noticing that quintessence and phantom like behaviors can effectively be mimicked by extensions of BD gravity which do not involve any sort of ad hoc scalar fields but just rigid or dynamical vacuum effects.

IV. STRUCTURE FORMATION

In order to properly test the theoretical predictions of the various cosmological models under study we must consider structure formation. The differential equation for the matter density contrast $\delta_m \equiv \delta\rho_m/\rho_m$ for the models considered here can be approximated as follows [57]:

$$\delta_m'' + \left(3 + a\frac{H'(a)}{H(a)}\right)\frac{\delta_m'}{a} - \frac{4\pi G(a)}{H^2(a)}\frac{\delta_m}{a^2} = 0\,, \qquad (38)$$

where $()' \equiv d/da$ denotes derivative with respect to the scale factor. The Hubble function is the corresponding one for each model. Notice also that $G(a) = 1/\phi(a) = G_N a^\epsilon$ is given by Eq. (13), and in particular $G(a) = G_N$ for the GR-ΛCDM model ($\epsilon = 0$). The above approximation for the structure formation equation turns out to be accurate since the main effects actually come from the different expressions of the Hubble function in each model.

We are interested in finding the required initial conditions deep into the matter dominated epoch, which means that we calculate the limit of the different terms in (38) when the scale factor a is small enough, see [57] for the details. In that limit the resultant differential equation admits power law solutions (we only consider the growing mode solution) $\delta_m \sim a^s$. For each model we find:

$$\mathbf{GR-\Lambda CDM}: \quad s = 1 \qquad (39)$$

$$\mathbf{BD-\Lambda CDM}: \quad s \simeq 1 - \frac{4}{5}\epsilon - \frac{1}{10}\epsilon^2\omega_{\rm BD} \qquad (40)$$

$$\mathbf{BD-RVM}: \quad s \simeq 1 - \frac{4}{5}\epsilon - \frac{6}{5}\nu - \frac{1}{10}\epsilon^2\omega_{\rm BD}, \qquad (41)$$

where for the BD-ΛCDM we have neglected the terms of order $\mathcal{O}(\epsilon^2)$ whereas for the BD-RVM we have neglected those of order $\mathcal{O}(\epsilon^2, \epsilon\nu, \nu^2)$. Notice that in order to have suppression of structure formation we need that the overall contribution from the ϵ, ν and $\epsilon^2\omega_{\rm BD}$ provides $s < 1$.

Comparing with the available structure formation data requires the computation of the standard LSS quantity $f(z)\sigma_8(z)$, where $f(z) = d\ln\delta_m(a)/d\ln a$ is usually called the growth factor and $\sigma_8(z)$ is the rms mass fluctuations on $R_8 = 8h^{-1}$ Mpc scales, whose expression is:

$$\sigma_8^2(z) = \delta_m^2(z)\int \frac{d^3k}{(2\pi)^3}P(k,\vec{p})W^2(kR_8). \qquad (42)$$

$P(k,\vec{p}) = P_0 k^{n_s}T^2(k,\vec{p})$ is the linear matter power spectrum, P_0 a normalization constant (see below), $T(k,\vec{p})$ represents the transfer function and, finally, $\vec{p}$ a parameter vector involving the characteristic free parameters of each model. Among the options available for the transfer

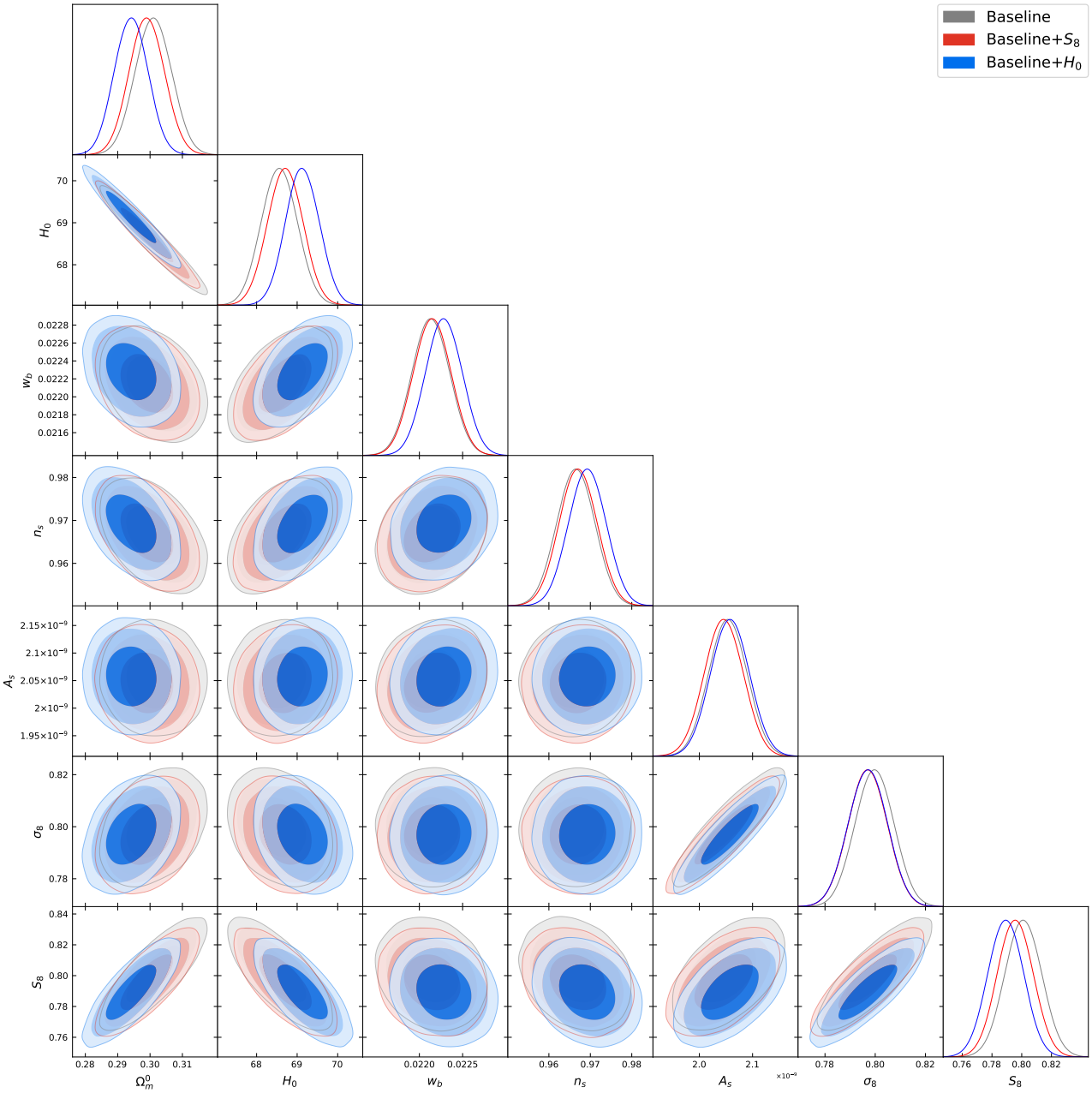


FIG. 3. Two-dimensional marginalized likelihood distributions, for each one of the parameters, in the GR-ΛCDM model at 1σ, 2σ and 3σ c.l., obtained considering the Baseline (gray contours), the Baseline+S_8 (red contours) and the Baseline+H_0 (blue contours) datasets. We also included the corresponding one-dimensional marginalized likelihood distributions. The H_0 parameter is expressed in km/s/Mpc.

function, we use here the BBKS one [92]:

$$T(x) = \frac{\ln(1+0.171x)}{0.171\,x}\Big[1 + 0.284x + (1.18x)^2 + \\ + (0.399x)^3 + (0.490x)^4\Big]^{-1/4} . \tag{43}$$

In the above expression $x \equiv k/\tilde{\Gamma}k_{\rm eq}$ where

$$k_{\rm eq} = a_{\rm eq}H(a_{\rm eq}) \tag{44}$$

is the value of the comoving wave number at the equality point between matter and radiation, defined through $\rho_r(a_{\rm eq}) = \rho_m(a_{\rm eq})$. The quantity $\tilde{\Gamma} = e^{-\Omega_b^0 - \sqrt{2h}\frac{\Omega_b^0}{\Omega_m^0}}$ is known as the modified shape parameter and it incorporates the effect of baryons in the transfer function. In (42) we have also introduced the top-hat smoothing function

$$W(kR_8) = \frac{3}{k^2R_8^2}\left(\frac{\sin(kR_8)}{kR_8} - \cos(kR_8)\right) , \tag{45}$$

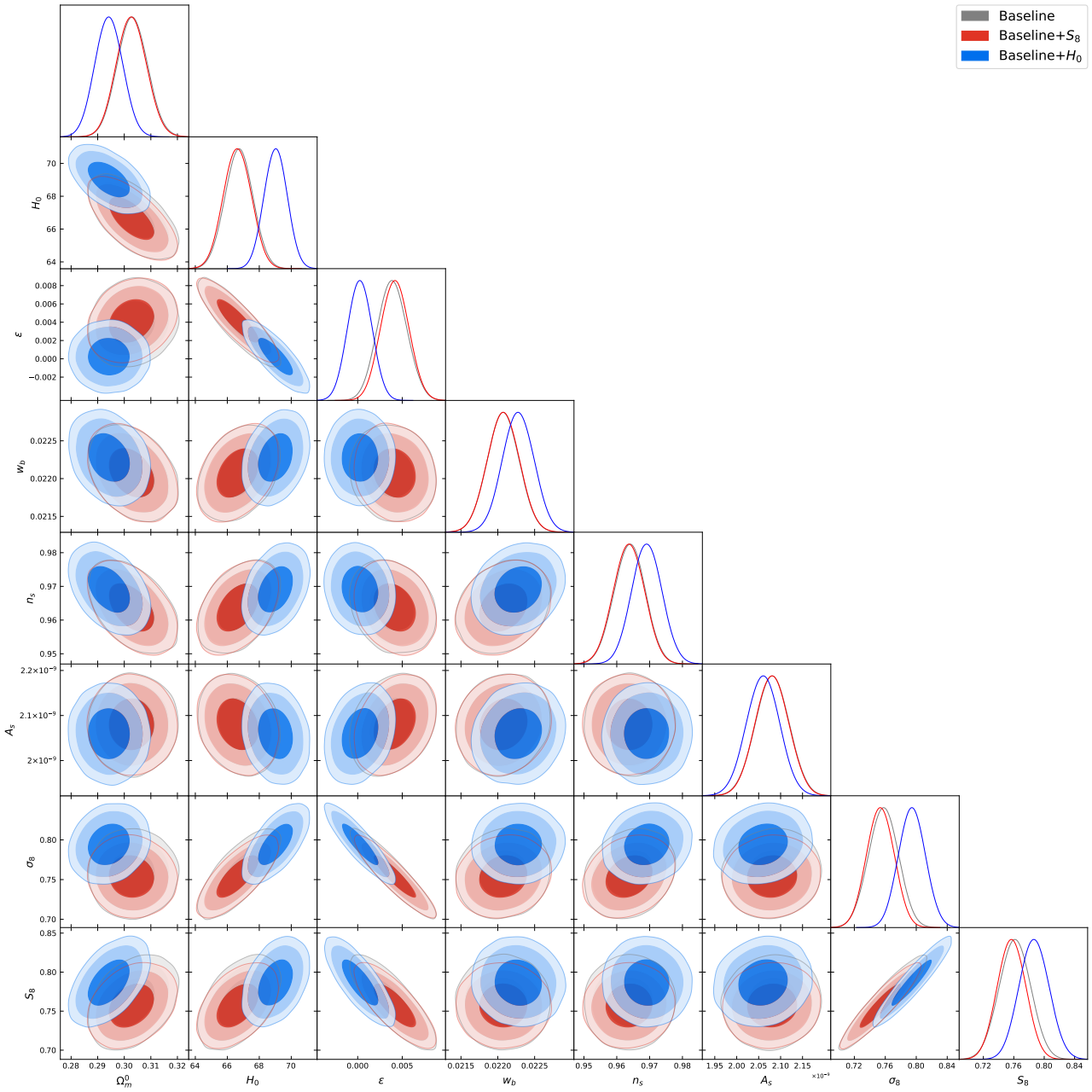


FIG. 4. Same as in Fig. 3 but for the BD-ΛCDM cosmological model.

with R_8 defined previously. There are several ways of dealing with the normalization factor P_0 of the matter power spectrum. In this work, the procedure followed takes advantage of using the connection between P_0 and the normalization factor of the primordial power spectrum, A_s, for which we have obsevational constraints. The relation between these quantities is [82]:

$$P_0 = A_s \frac{8\pi^2}{25} \frac{k_*^{1-n_s}}{(\Omega_m^0 h^2)^2 (100\varsigma)^4}\,, \tag{46}$$

where $k_* = 0.05\text{Mpc}^{-1}$ is the pivotal scale used by the Planck team and $\varsigma \equiv 1$ km/s/Mpc$= 2.1332 \times 10^{-44}$GeV. In contrast to [82], however, here we can use A_s as a fit parameter since we computed a compressed likelihood involving this parameter with the remaining distance priors, see (49) below. To ease the computation what we actually do is calculate the normalized quantity $\sigma_8(z)/\sigma_8^\Lambda(0)$, where $\sigma_8^\Lambda(0) = 0.8118$ is the value obtained by the Planck team for the analysis Planck 2018 TT+lowE [10] with the vanilla model. Therefore, taking into account the definitions of the different quantities involved we may compute the value of the $\sigma_8(z)$ function

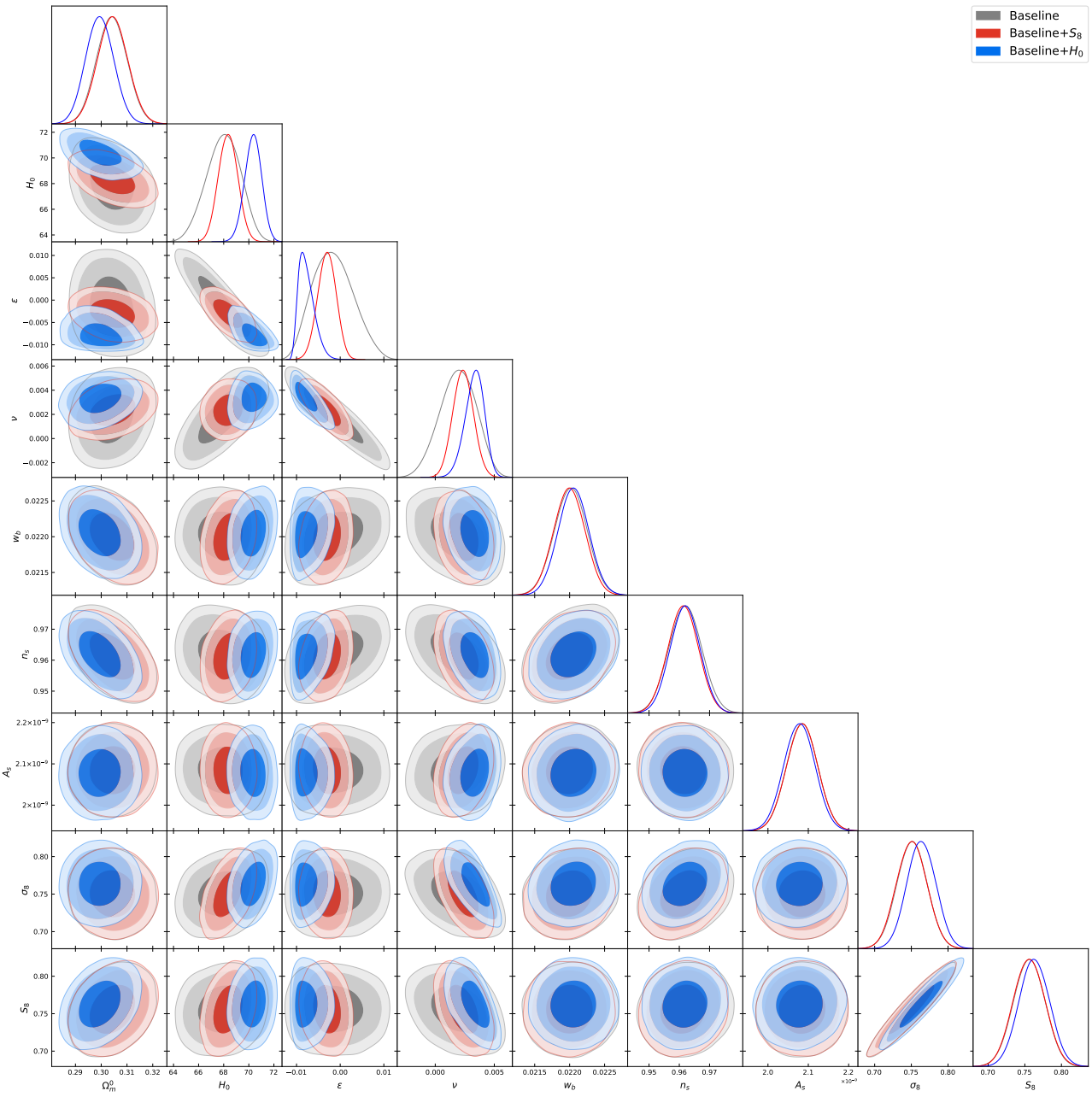


FIG. 5. Same as in Fig. 3 but for the BD-RVM cosmological model.

at different redshifts as follows:

$$\sigma_8(z) = \sigma_8^{\Lambda}(0)\sqrt{\frac{P_0}{P_0^{\Lambda}}}\frac{\delta_m(z)}{\delta_m^{\Lambda}(0)}\frac{\gamma}{\gamma_{\Lambda}} \tag{47}$$

where for the sake of simplicity we have defined

$$\gamma \equiv \left(\int_0^{\infty} dk k^{2+n_s} T^2(k,\vec{p}) W^2(kR_8)\right)^{1/2}. \tag{48}$$

In order to fix the normalization constant P_0 we have to choose a fiducial cosmology. In this work we choose the vanilla standard model of cosmology constrained with Planck 2018 TT+lowE data. So, in (47) those quantities labeled with Λ (superscript or subscript) correspond to either the fiducial values, of the cosmological parameters, found in the Planck 2018 TT+lowE analysis [10] or to the values of the functions ((43)-(46) and (48)) computed using these fiducial parameters. Similarly, the quantity denoted by $\delta_m^{\Lambda}(0)$ corresponds to the present value of the matter density contrast for the fiducial model obtained after solving the perturbations equation (38).

Thus, using (47) for $\sigma_8(z)$ and the expression for the growth factor $f(z) = d\ln\delta_m(a)/d\ln a$ which follows upon solving (38) for the density contrast $\delta_m(a)$ for each

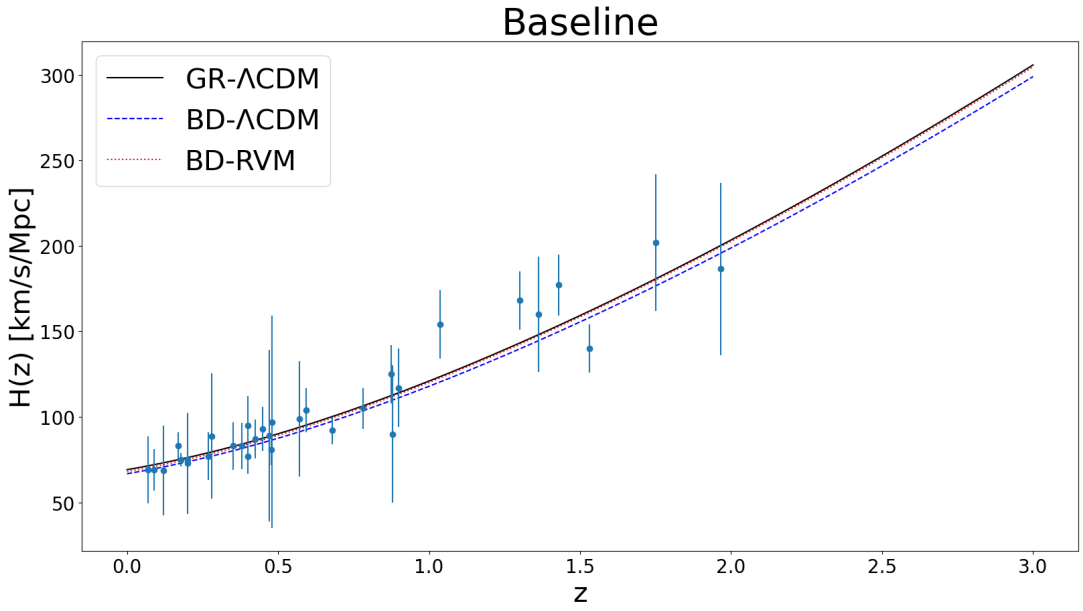


FIG. 6. Evolution of the Hubble parameter as a function of the redshfit z for each one of the cosmological models using the results obtained with the Baseline set. The 33 observational data points $H(z_i)$ are shown with the corresponding error bars. The $\chi^2_{H(z)}$ for the different models when the fitting values from Table I are used, are provided in Table IV. The differences among the models in fitting this particular data are not very significant as compared to other data sources.

model, we are in position to numerically compute the values of the observable $f(z)\sigma_8(z)$ at different redshifts.

V. DATA AND METHODOLOGY

We fit the GR-ΛCDM, the BD-ΛCDM and the BD-RVM cosmological models to a very complete and robust dataset compiled from: distant type Ia Supernovae (SNIa), baryonic acoustic oscillations (BAO), a compilation of measurements of the Hubble parameter at different redshifts ($H(z_i)$), large-scale structure formation data ($f(z_i)\sigma_8(z_i)$), CMB Planck 2018 TT+lowE data, a prior on the weak-lensing observable (S_8) and, finally, a prior on the H_0 parameter [93]. In what follows a brief description of each one of the datasets is provided, together with the corresponding references:

SNIa: 6 effective points on the inverse of the normalized Hubble function $1/E(z_i)$, with its corresponding covariance matrix, from the Pantheon+MCT sample [94, 95], which compresses the information from 1063 SNIa.

BAO: We employ data on both, isotropic and anisotropic BAO estimators [96–101]. See [61] and [75] in order to see the tables with the data points.

H(z): We use the 32 data points on the Hubble parameter $H(z_i)$ [102–111], all of them, obtained employing the differential age technique, and therefore defining a robust type of $H(z)$ data which are not correlated with BAO, the co-called cosmic chronometers.

LSS: 14 large scale structure (LSS) data points at different redshifts embodied in the observable $f(z_i)\sigma_8(z_i)$ [98, 101, 112–118].

CMB: From Planck 2018 TT+lowE data [10] we have computed a compressed likelihood for the distance priors $\mathcal{R}$ (shift parameter) and ℓ_a (acoustic length) together with the corresponding correlations for the parameters (ω_b, n_s, A_s). The mean values and the corresponding error bars for the different components are $\omega_b = 0.022038 \pm 0.00022$, $A_s = (2.0898 \pm 0.034) \times 10^{-9}$, $n_s = 0.9641 \pm 0.0058$, $\ell_a = 301.54 \pm 0.1319$ and $\mathcal{R} = 1.7469 \pm 0.0072$, whereas the matrix containing the correlations between the different elements takes the form:

$$C_{\rm CMB} = \begin{pmatrix} 1 & 0.1193 & 0.4926 & -0.3014 & -0.5546 \\ 0.1193 & 1 & 0.0262 & -0.0434 & 0.0240 \\ 0.4926 & 0.0262 & 1 & -0.3571 & -0.7776 \\ -0.3014 & -0.0434 & -0.3571 & 1 & 0.4207 \\ -0.5546 & 0.0240 & -0.7776 & 0.4207 & 1 \end{pmatrix} \tag{49}$$

Prior on S_8: A prior value on the observable $S_8 \equiv \sigma_8\sqrt{\Omega^0_m/0.3} = 0.737^{+0.040}_{-0.036}$ from the gravitational weak-lensing survey KiDS+VIKING-450 [90].

Prior on H_0: We consider the prior on the H_0 parameter provided by the SH0ES team [91] (obtained from the cosmic distance ladder method), namely

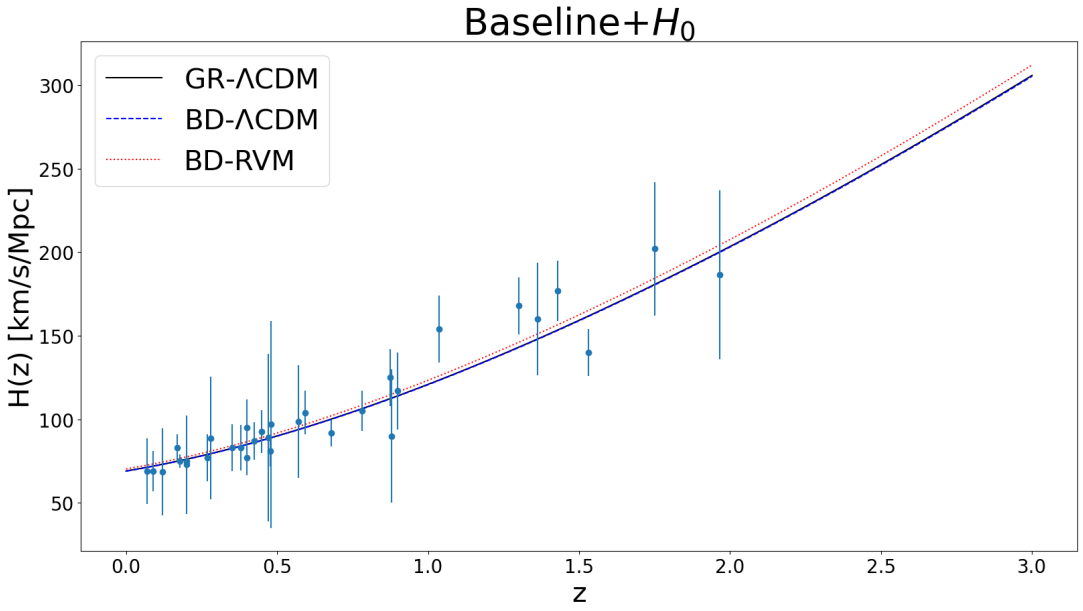


FIG. 7. Same as in Fig. 6 but with the Baseline+H_0 set. The $\chi^2_{H(z)}$ for the different models when the fitting values from Table III are used, are provided in Table V Again, the differences among the models in fitting the $H(z)$ data are not significant.

$H_0 = 73.04 \pm 1.04$ km s^{-1} Mpc^{-1}. This measurement exceeds the Planck+ΛCDM result by 5σ (one in 3.5 million), making it implausible to reconcile the two by chance.

We do not use simultaneously all of the data points given above to fit the models under study, but arrange for three different combinations of observational points. Specifically, our dataset scenarios are the following:

- **Baseline**: We consider as our Baseline dataset the one made of the following string of data: SNIa+$H(z)$+BAO+LSS+CMB. For this scenario, the SH0ES prior for H_0 is not included.
- **Baseline+S_8**: In addition to the data considered in the Baseline set, we add the prior on the S_8 parameter.
- **Baseline+H_0**: The Baseline data is in this case complemented with the prior on the H_0 parameter provided by the SH0ES team.

Using the above datasets, we proceed to analyze whether the two extensions of the Brans & Dicke cosmology being considered in this work, namely the BD-ΛCDM and BD-RVM models, are capable to fit the observations in a comparable way to the concordance model (GR-ΛCDM), or maybe even better, and in particular if they are able to cope with the σ_8-tension [71] and the H_0-tension [69].

In order to compare the BD-ΛCDM and the BD-RVM with the GR-ΛCDM, we define the joint χ^2-function, with Gaussian errors, as follows:

$$\chi^2_{\rm tot} = \chi^2_{\rm SNIa} + \chi^2_H + \chi^2_{\rm BAO} + \chi^2_{\rm LSS} + \chi^2_{\rm CMB}. \qquad (50)$$

The different contributions of $\chi^2_{\rm tot}$ are defined in the standard way. Only a few observations are in order. In some cases such as e.g. the CMB, see Eq. (49), the correlation matrix and hence the covariance matrix for $\chi^2_{\rm CMB}$, is non-diagonal. Notice also that the term χ^2_H may contain or not the H_0-prior from [91] depending on the dataset used for our analysis, and similarly the term $\chi^2_{\rm LSS}$ may contain or not the S_8-prior from [90].

We employ the aforementioned datasets to constrain the cosmological parameters of the models under consideration. We are also interested in determining which of the models is the one that performs better when it comes to reproducing the observational data. To do so, we utilize two statistical criteria, the Akaike information criteria (AIC) [119] and the deviance information criteria (DIC) [120], which take into account the presence of extra *d.o.f* by appropriately increasing the value of $\chi^2_{\rm min}$. See [121], for instance, for a summarized discussion on how to use and interpret these criteria in the cosmological context.

The formula for the AIC turns out to be

$$\mathrm{AIC} = \chi^2_{\rm min} + \frac{2nN}{N-n-1}\,, \qquad (51)$$

where $\chi^2_{\rm min}$ is the minimum value of the χ^2-function, n represents the number of independent parameters and N is the total number of data points. The second term on the *r.h.s.* obviously represents the penalty for the additional number of free parameters (*d.o.f*).

Baseline			
	GR-ΛCDM	**BD-ΛCDM**	**BD-RVM**
$\chi^2_{\rm SNIa}$	5.57	5.62	5.63
$\chi^2_{\rm BAO}$	13.33	14.26	14.49
$\chi^2_{H(z)}$	14.69	16.47	14.92
$\chi^2_{f\sigma_8}$	16.46	9.78	9.27
$\chi^2_{\rm CMB}$	1.78	0.19	0.43

TABLE IV. We display the contributions of some of the data employed to test the cosmological models, obtained considering the mean values provided in Table I for the Baseline data set.

Baseline+H_0			
	GR-ΛCDM	**BD-ΛCDM**	**BD-RVM**
$\chi^2_{\rm SNIa}$	5.53	5.53	5.53
$\chi^2_{\rm BAO}$	14.32	14.29	14.90
$\chi^2_{H(z)}$	14.73	14.75	15.23
$\chi^2_{f\sigma_8}$	14.45	14.24	9.27
$\chi^2_{\rm CMB}$	2.87	2.94	0.51
$\chi^2_{H_0}$	14.13	14.94	6.74

TABLE V. Same as in Table IV but for the Baseline+H_0 data set. Notice in particular the low values of $\chi^2_{f\sigma_8}$, $\chi^2_{\rm CMB}$ and $\chi^2_{H_0}$ in the case of the BD-RVM, which helps to understand why this model can alleviate the two tensions σ_8 and H_0 at a time. In contrast, the BD-ΛCDM model fails to reduce these tensions, despite it helps to lessen the σ_8 tension when the prior on H_0 is not included, see Table IV .

On the other hand, the DIC value can be computed from

$$\mathrm{DIC} = \chi^2(\bar{\theta}) + 2p_D\,, \tag{52}$$

with $p_D = \overline{\chi^2} - \chi^2(\bar{\theta})$ the effective number of parameters (compare with (51) in the limit of large N), $2p_D$ being the so-called 'model complexity', which is the quantity used in the DIC criterion to penalize the presence of extra *d.o.f.* The term $\overline{\chi^2}$ is the mean value of the χ^2-function obtained from the Markov chains and $\bar{\theta}$ represents the mean values of the fitting parameters.

We choose to define the AIC and DIC differences with respect to the vanilla model (GR-ΛCDM):

$$\Delta\mathrm{AIC} \equiv \mathrm{AIC}_{\text{GR-}\Lambda\mathrm{CDM}} - \mathrm{AIC}_{\rm X} \tag{53}$$

$$\Delta\mathrm{DIC} \equiv \mathrm{DIC}_{\text{GR-}\Lambda\mathrm{CDM}} - \mathrm{DIC}_{\rm X}\,. \tag{54}$$

Therefore positive differences mean that the new models (X) fare better than the vanilla model, whereas negative differences mean that they fare worse. In our case, X represents either the BD-ΛCDM or the BD-RVM models. For values $0 \leq \Delta\mathrm{AIC}, \Delta\mathrm{DIC} < 2$ the level of evidence in favor of the considered option beyond the standard model is deemed as *weak*. If, however, $2 \leq \Delta\mathrm{AIC}, \Delta\mathrm{DIC} < 6$ we speak of *positive* evidence, whereas if $6 \leq \Delta\mathrm{AIC}, \Delta\mathrm{DIC} < 10$ it is said that there is *strong* evidence supporting the non-standard model. Finally, if $\Delta\mathrm{AIC}, \Delta\mathrm{DIC} > 10$ one may licitly conclude(according to these information criteria) that there is *very strong* evidence in favor of the model under study relative to the GR-ΛCDM model. On the other hand, negative values of $\Delta\mathrm{AIC}, \Delta\mathrm{DIC}$ would indicate that the GR-ΛCDM is the model favored by the observational data.

VI. DISCUSSION OF THE RESULTS

We are now ready to discuss the results obtained in this analysis and hence to effectively assess if the considered extensions of the Brans-Dicke framework are competitive with the vanilla concordance model of cosmology (which we recall is denoted GR-ΛCDM in this work). The results of our analysis are collected in our tables and figures.

Specifically, in Tables I, II and III we present the fitted values of the cosmological parameters for each one of the models under study, namely the vanilla model and the two Brans-Dicke extensions, i) BD-ΛCDM and ii) BD-RVM, which have fixed and running cosmological term, respectively. The fitting values for each model scenario are obtained by considering the three different datasets described in Sec. V. Furthermore, in Tables IV and V we furnish the detailed breakdown of the different contributions to χ^2, defined in Eq. (50), from two of the dataset scenarios (viz. the Baseline and the Baseline+H_0 data sets). Additionally, in Figs. 3, 4 and 5 we provide the two-dimensional marginalized likelihood distributions for the cosmological parameters. Their mean values as well as the mentioned Figs. 3-5 have been obtained with the help of the `GetDist` code [122]. Finally, in Figs. 6 and 7 we confront theoretical predictions and data points for a particular source of data, as we shall comment.

Let us now discuss our results with some more detail. As it can be seen from the caption of Table I, the vanilla model has 5 free parameters $(\omega_b, n_s, A_s, H_0, \Omega_m^0)$, where we recall that we do not consider the optical depth parameter τ since it is not sensitive to our analysis. On the other hand, the BD-ΛCDM has 6 free parameters, the new one being ϵ from (13). Finally, the BD-RVM has 7 free parameters; the additional one with respect to the previous model is ν from (21). Owing to the different number of free parameters among the competing models being considered here, we are naturally led to use the statistical information criteria AIC and DIC defined in equations (51) and (52). Because these criteria penalize the extra number of parameters, model comparison will be much fairer than just using the minimum χ^2 test. For the BD-ΛCDM, and considering the Baseline dataset, we find (cf. Table I) ΔAIC (ΔDIC) = 4.76 (5.06), whereas for the BD-RVM we get ΔAIC (ΔDIC) = 3.81 (5.10), all of them being positive, which means, according to the definitions (53) and (54), that in both cases there is *positive* evidence in favor of the non-standard models, and hence supporting the extended BD frameworks under scrutiny. Notwithstanding, it should be clear that not all the data used in our analysis show a particular sensitivity in favor of these extensions. For example, in Figs. 6 and 7 the evolution of the Hubble function with the redshift is plotted against the observational data points and their corresponding errors. At naked eye, we can see that these data are fitted comparably well by the three models under consideration, and this is quantitatively confirmed by the corresponding $\chi^2_{H(z)}$ values in Tables IV and V.

In point of fact, the real improvement in the fitting performance of the models based on BD gravity with respect to the vanilla model comes from another direction and it goes hand in hand with the significantly smaller value quoted in Table IV regarding the $\chi^2_{f\sigma_8}$ contribution as compared to GR-ΛCDM within the Baseline dataset. This trend is confirmed with the inclusion of the S_8-prior (i.e. the Baseline+S_8 dataset), which increases the positive differences in the information criteria. Thus, for the BD-ΛCDM we find (cf. Table II) ΔAIC (ΔDIC) = 7.01 (7.61) and for the BD-RVM, ΔAIC (ΔDIC) = 6.39 (6.59). Both statistical criteria point consistently towards the same conclusion, which is that when the Baseline+S_8 dataset is considered we find *strong* evidence in favor of the two BD-gravity extensions as compared with the concordance model. This situation, however, does not replicate in exactly the same way when the alternative Baseline+H_0 dataset is analyzed. Here the BD model with running vacuum (the BD-RVM) stands out over the other two and becomes the front runner. Indeed, whereas for the BD-ΛCDM we obtain the negative differences ΔAIC (ΔDIC) = -2.37 (-2.46) (cf. Table III), for the BD-RVM we get, instead, the peak positive yields of our analysis: ΔAIC (ΔDIC) = 12.43 (12.10). These results illustrate, on the one hand, that the BD-ΛCDM is no better than the vanilla model (with respect to the dataset under consideration), and on the other hand show that the BD-RVM appears to be a model preferred to the vanilla model (with respect to the same dataset), this being confirmed with *very strong* evidence according to the jargon used by the information criteria. On inspecting Table V we confirm once more that the origin of such an improvement lies in the significantly lower value of $\chi^2_{f\sigma_8}$ and also, in this case, of $\chi^2_{H_0}$. In other words, its success stems from a better description of the LSS data, and, at the same time, of the H_0 prior, than the vanilla model. Such an accomplishment is unmatched by the BD model with constant Λ, i.e. the BD-ΛCDM.

The obtained results are a reflex of the changes in the values of the cosmological parameters depending on which dataset is analyzed. Let us focus on the change of the extra parameters with respect to the GR-ΛCDM model, namely ϵ for the BD-ΛCDM, and ϵ and ν for the BD-RVM. For the former within the Baseline dataset, we find $\epsilon = 0.0039 \pm 0.0015$ (2.60σ), whereas for the latter two we get $\epsilon = -0.0016^{+0.0037}_{-0.0049}$ (0.37σ) and $\nu = 0.0019^{+0.0016}_{-0.0013}$ (1.31σ). If we now move to the Baseline+S_8 dataset, for the BD-ΛCDM model we get $\epsilon = 0.0042 \pm 0.0015$ (2.80σ). The fact of finding a higher confidence level in favor of $\epsilon > 0$ in this dataset as compared to the Baseline one should not come as a surprise since this result goes hand in hand with the increase in the positive differences of AIC and DIC observed when we move from the Baseline to the Baseline+S_8 dataset.

The situation with the BD framework with running vacuum (the BD-RVM model) deserves a particular consideration since the two extra parameters ϵ and ν tend to present a high level of degeneracy depending on the dataset being considered. This can be immediately appraised by looking at the panel ϵ-ν in Fig. 5. There is a strong degeneracy between these two parameters when the Baseline dataset is considered and this is why we do not observe a comparable level of evidence in favor of $\epsilon \neq 0$ to that one we observe in the BD-ΛCDM model (namely 0.37σ in the former versus 2.60σ in the latter). This degeneracy, however, is broken once the S_8-prior enters the fit. In effect, using the Baseline+S_8 dataset

the BD-RVM model now renders $\epsilon = -0.0031 \pm 0.0019$ (1.63σ). At the same time we find $\nu = 0.00237 \pm 0.00080$ (2.96σ) and hence the nonvanishing character of ν becomes now also reinforced (from 1.31σ to 2.96σ). The corresponding contour plot for these parameters in Fig. 5 also reflects this situation, as it has clearly shrunk as compared to the one of the Baseline dataset.

Furthermore, as could be expected, the inclusion of the prior $H_0 = 73.04 \pm 1.04$ km/s/Mpc [91] has a dramatic impact on our analysis. Within the context of the BD-ΛCDM model under the Baseline+H_0 dataset, we find a rather inconspicuous output for the extra parameter of the model, as it is found to be compatible with zero: $\epsilon = 0.0002 \pm 0.0013$ (0.15σ). As a result, the evidence in favor of a time evolving Newtonian gravitational coupling has subsided in this setup. In actual fact, the differences among the fitting values of BD-ΛCDM and GR-ΛCDM are now virtually absent (cf. Table III) and since the BD-ΛCDM has one additional parameter, this automatically entails a penalty from the information criteria which is reflected in the negative values of ΔAIC and ΔDIC in that table. To recap, we may say that the inability of the BD-ΛCDM model to accommodate the H_0-prior causes the suppression of the dynamics of the BD scalar field $\phi = 1/G(a)$. In stark contrast to this meager situation with fixed Λ when the H_0 prior is included, the conclusions reached in the case with dynamical vacuum are in the opposite pole. Indeed, for the the BD-RVM model we find (cf. Table III): $\epsilon = -0.00767^{+0.00069}_{-0.00220}$ (5.31σ) and $\nu = 0.00332^{+0.00084}_{-0.00065}$ (4.46σ). The high confidence levels obtained for the nonvanishing values of both parameters ϵ and ν are now in the rather conspicuous range $\sim 4.5 - 5.3\sigma$, which is in consonance with the outstanding positive values of the information criteria in Table III: ΔAIC = 12.43 and ΔDIC = 12.10 Altogether, the output of the BD-RVM for the dataset involving the H_0 prior brings that model to the forefront position with respect to the vanilla cosmological model and the BD-ΛCDM. In obtaining these results, we point out that the analysis does not show a especial sensitivity to the value and sign of the BD parameter $\omega_{\rm BD}$, provided it is sufficiently large. Let us quote central values only. For instance, in the context of the BD-ΛCDM model we find $\omega_{\rm BD} \simeq -467$ when the Baseline dataset is considered, $\omega_{\rm BD} \simeq -434$ for the Baseline+S_8 and $\omega_{\rm BD} \simeq -9138$ for the Baseline+H_0. In all these cases $\epsilon > 0$, implying that the effective gravitational coupling $G(a)$ increases with the expansion (i.e. for increasing a). On the other hand, ϵ proves negative for the BD-RVM within the three datasets, and hence $G(a)$ decreases with the expansion. This is favorable since the structure formation is suppressed. For this model we find, instead, positive values of the BD parameter, to wit: $\omega_{\rm BD} \simeq 1138$, $\omega_{\rm BD} \simeq 587$ and $\omega_{\rm BD} \simeq 238$ for the Baseline, Baseline+S_8 and Baseline+H_0 datasets, respectively.

It is to be emphasized that in spite of the fact that the two generalized BD models under study can cope with the σ_8-tension for certain datasets, it turns out that only the BD-RVM model can ease the H_0-tension along with the σ_8-tension, and this occurs within the Baseline+H_0 dataset. This is manifest from the results displayed in Table III. Here, the vanilla (GR-ΛCDM) model yields $H_0 = 69.13 \pm 0.38$ km/s/Mpc, the BD-ΛCDM gives $H_0 = 69.02 \pm 0.67$ km/s/Mpc and, finally, the BD-RVM renders the highest fitting value: $H_0 = 70.34 \pm 0.63$ km/s/Mpc. The residual tensions of these fitting results with respect to the latest local measurement $H_0 = 73.4 \pm 1.04$ km/s/Mpc [91] are 3.53σ, 3.25σ and 2.22σ respectively. That the BD-RVM can lower the H_0-tension without worsening the σ_8-tension within the Baseline+H_0 dataset is evinced from the fitting value of σ_8 in Table III, which reads $\sigma_8 = 0.764 \pm 0.018$. This is a very good result, which can be rephrased in terms of the fitting value of the related LSS observable $S_8 = \sigma_8 \sqrt{\Omega_m^0/0.3}$ quoted in the same table: $S_8 = 0.763 \pm 0.019$. If we compare it with the weak-lensing measurement of this same observable from [90], namely $S_8 = 0.737^{+0.040}_{-0.036}$, we find that there is no residual tension left to speak of at this point, for the mismatch is less than 1σ (specifically $\sim 0.61\sigma$).

Our analysis also shows that a very useful way to parametrize the background deviations of the two BD extensions with respect to the vanilla model is to consider that the BD field and the cosmological term induce an effective form of dark energy (DE) when viewed from the GR standpoint. The deviation from -1 of the corresponding effective equation of state (EoS) parameter and also its dynamical character (see Figs. 1 and 2), could be the appropriate smoking gun of the effect we are searching for. To be more precise, we find that the EoS signature of the BD-extension with fixed Λ (the BD-ΛCDM model) is its phantom-like behavior, whereas the BD model with running Λ (BD-RVM) mimics quintessence. No phantom nor quintessence fundamental fields are around, of course.

We end up this section by emphasizing that the current study on extensions of the BD framework assume the FLRW metric with three-dimensional flat geometry. However, many recent studies in the literature have exploited the properties of the non spatially flat FLRW spacetime, which can be of interest in order to tackle various conflicting aspects of the modern cosmological observations, including the σ_8 and H_0 tensions that are presently besetting the vanilla ΛCDM or concordance model of cosmology, see for instance [123–132] and references therein.

VII. CONCLUSIONS

In this work we have found intriguing signs that a mild (logarithmic) time-variation of the Newtonian coupling in combination with a slow cosmic evolution of the vacuum energy density (VED) can be essential ingredients to improve the global quality fit to the modern cosmological data and relieve the ΛCDM tensions. The framework that we have used to realize this scenario is the time-honored Brans & Dicke (BD) gravitational theory [41],

where the time variation of the Newtonian coupling is implemented through the slowly varying BD field ϕ coupled to curvature. In contrast to the original BD theory, however, we have admitted the presence of a cosmological term, which we have have treated either as a rigid VED $\rho_{\rm vac}$ = const. or as a function $\rho_{\rm vac}(H)$ evolving with the Hubble rate H. The rigid option leads to what we have called the BD-ΛCDM model, which is the BD counterpart of the usual concordance ΛCDM model in GR (denoted here GR-ΛCDM for obvious reasons). As for the dynamical option, in which the VED is evolving as a function of the Hubble rate H, $\rho_{\rm vac}(H)$, we have chosen the very specific form predicted within the context of the running vacuum model (RVM), which has been recently derived from first principles, namely within QFT in FLRW spacetime [33–36], see also the recent review [15]. This vacuum dynamics in combination with the BD gravity framework leads to what we have called the BD-RVM model, in which $\rho_{\rm vac}$ changes quadratically with H through a small coefficient $|\nu| \sim 10^{-3}$ associated with the β-function of the running VED in QFT in cosmological spacetime. That coefficient is precisely computable (analytically) from the quantum effects associated with the effective action of gravity with quantized matter fields, as shown in the aforementioned references. However, its ultimate numerical value must be determined by comparison with the observational data.

The results that we have obtained for the BD-RVM show that we can improve significantly the performance of the standard model of cosmology, or GR-ΛCDM model, when it comes to fitting three different datasets involving the latest SNIa+$H(z)$+BAO+LSS+CMB observations. This is true also in some cases for the BD model with rigid cosmological constant (i.e. the BD-ΛCDM). The version of the BD framework with running vacuum, however, improves the efficiency with respect to the three datasets used. This does not occur for the BD model with fixed Λ. If we attend to the information criteria that we have used (the AIC and DIC statistics), the BD-RVM is by far the best model capable of fitting the data used in our analysis, especially when the model is tested against the larger set of data (which involves also the H_0 prior). According to the statistical criteria AIC and DIC, and in the context of what we have called the Baseline dataset, we find that the BD-ΛCDM and the BD-RVM models are *positively* favored over the GR-ΛCDM model. When we add weak-lensing observations through the S_8 parameter (defining the Baseline+S_8 dataset), the evidence for both models increases and reaches the level of being *strong* in the language of the information criteria.

As indicated above, of especial significance is when the H_0-prior is also included in the fit, i.e. the Baseline+H_0 dataset. Then the most competitive model among those considered here is by far the BD-RVM, i.e. the BD framework incorporating the RVM dynamics of the vacuum in the form $\delta\rho_{\rm vac} \propto \nu\, m_{\rm Pl}^2(H^2 - H_0^2)$. In this scenario, we find fitting values for the model parameters which attain very high confidence levels: $\epsilon = -0.00767^{+0.00069}_{-0.00220}$ (5.31σ) and $\nu = 0.00332^{+0.00084}_{-0.00065}$ (4.46σ). This result attest that the extra (beyond the standard model) contributions induced by these parameters are definitely welcome; in other words, they are highly preferred to be non-vanishing rather than vanishing. In full consistency with these fitting results, the statistical information criteria provide large positive values ΔAIC (ΔDIC) = +12.43 (+12.10) to the BD-RVM (as compared to the vanilla model), cf. Table III. The overall performance of the BD-RVM, if we should take seriously these standard statistical tools for model comparison, proves clearly superior to that of the conventional GR-ΛCDM. In addition, the BD-RVM model is capable of alleviating the H_0-tension and the σ_8-tension at a time. This is not possible for the BD-ΛCDM, despite it has a positive bearing on the σ_8 tension. Specifically, within the BD-RVM there remains essentially no σ_8 tension at all ($\sigma_8 = 0.764 \pm 0.018$) while the usual $\sim 5\sigma$ c.l. H_0 tension narrows down to only 2.2σ ($H_0 = 70.34 \pm 0.63$ km/s/Mpc). Overall, in the light of our analysis it is fair to say that within the BD-RVM there is potential enough for a significant reduction of both tensions and for providing a global fit to the cosmological data which is highly competitive with the current concordance model.

ACKNOWLEDGMENTS

JSP is funded by projects PID2019-105614GB-C21 (MINECO), 2017-SGR-929 (Generalitat de Catalunya) and CEX2019-000918-M (ICCUB). He also acknowledges the participation in the COST Association Action CA18108 "Quantum Gravity Phenomenology in the Multimessenger Approach (QG-MM)". JSP and JdCP acknowledge as well their participation in the COST Association Action CA21136 "Addressing observational tensions in cosmology with systematics and fundamental physics" (CosmoVerse). JdCP was supported by a FPI fellowship associated to the project FPA2016-76005-C2-1-P (MINECO).

[1] P. J. E. Peebles, *Principles of physical cosmology*. Princeton University Press, 1993.

[2] M. S. Turner, "The Road to Precision Cosmology," 2201.04741 [astro-ph.CO] (to appear in Annu. Rev. Nucl. Part. Sci. 2022). arXiv:2201.04741.

[3] A. G. Riess *et al.*, "Observational evidence from supernovae for an accelerating universe and a cosmological constant," *Astron. J.*, vol. 116, pp. 1009–1038, 1998.

astro-ph/9805201.

[4] S. Perlmutter *et al.*, "Measurements of Ω and Λ from 42 high redshift supernovae," *Astrophys. J.*, vol. 517, pp. 565–586, 1999. astro-ph/9812133.

[5] P. de Bernardis *et al.*, "A Flat universe from high resolution maps of the cosmic microwave background radiation," *Nature*, vol. 404, pp. 955–959, 2000. astro-ph/0004404.

[6] M. Tegmark *et al.*, "Cosmological Constraints from the SDSS Luminous Red Galaxies," *Phys. Rev. D*, vol. 74, p. 123507, 2006. astro-ph/0608632.

[7] E. Komatsu *et al.*, "Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation," *Astrophys. J. Suppl.*, vol. 180, pp. 330–376, 2009. arXiv:0803.0547.

[8] A. G. Riess, L. Macri, S. Casertano, H. Lampeitl, H. C. Ferguson, A. V. Filippenko, S. W. Jha, W. Li, and R. Chornock, "A 3% Solution: Determination of the Hubble Constant with the Hubble Space Telescope and Wide Field Camera 3," *Astrophys. J.*, vol. 730, p. 119, 2011. [Erratum: Astrophys. J.732,129(2011)].

[9] P. A. R. Ade *et al.*, "Planck 2015 results. XIII. Cosmological parameters," *Astron. Astrophys.*, vol. 594, p. A13, 2016. arXiv:1502.01589.

[10] N. Aghanim *et al.*, "Planck 2018 results. VI. Cosmological parameters," *Astron. Astrophys.*, vol. 641, p. A6, 2020. arXiv:1807.06209.

[11] P. J. E. Peebles and B. Ratra, "The Cosmological Constant and Dark Energy," *Rev. Mod. Phys.*, vol. 75, pp. 559–606, 2003. astro-ph/0207347.

[12] T. Padmanabhan, "Cosmological constant: The Weight of the vacuum," *Phys. Rept.*, vol. 380, pp. 235–320, 2003. hep-th/0212290.

[13] S. Weinberg, "The Cosmological Constant Problem," *Rev. Mod. Phys.*, vol. 61, pp. 1–23, 1989.

[14] J. Solà, "Cosmological constant and vacuum energy: old and new ideas," *J. Phys. Conf. Ser.*, vol. 453, p. 012015, 2013. arXiv:1306.1527.

[15] J. Solà Peracaula, "The cosmological constant problem and running vacuum in the expanding universe," *Phil. Trans. Roy. Soc. Lond. A*, vol. 380, p. 20210182, 2022. arXiv:2203.13757.

[16] O. Bertolami, "Time dependent cosmological term," *Nuovo Cim. B*, vol. 93, pp. 36–42, 1986.

[17] M. Ozer and M. Taha, "A Solution to the Main Cosmological Problems," *Phys. Lett. B*, vol. 171, pp. 363–365, 1986.

[18] P. J. E. Peebles and B. Ratra, "Cosmology with a Time Variable Cosmological Constant," *Astrophys. J.*, vol. 325, p. L17, 1988.

[19] J. Carvalho, J. Lima, and I. Waga, "On the cosmological consequences of a time dependent lambda term," *Phys. Rev. D*, vol. 46, pp. 2404–2407, 1992.

[20] C. Espana-Bonet et al, "Testing the running of the cosmological constant with type Ia supernovae at high z," *JCAP*, vol. 02, p. 006, 2004. hep-ph/0311171.

[21] P. Wang and X.-H. Meng, "Can vacuum decay in our universe?," *Class. Quant. Grav.*, vol. 22, pp. 283–294, 2005. astro-ph/0408495.

[22] H. A. Borges and S. Carneiro, "Friedmann cosmology with decaying vacuum density," *Gen. Rel. Grav.*, vol. 37, pp. 1385–1394, 2005. gr-qc/0503037.

[23] J. Solà and H. Stefancic, "Effective equation of state for dark energy: Mimicking quintessence and phantom energy through a variable lambda," *Phys. Lett. B*, vol. 624, pp. 147–157, 2005. astro-ph/0505133.

[24] J. Solà and H. Stefancic, "Dynamical dark energy or variable cosmological parameters?," *Mod. Phys. Lett. A*, vol. 21, pp. 479–494, 2006. astro-ph/0507110.

[25] J. S. Alcaniz and J. A. S. Lima, "Interpreting cosmological vacuum decay," *Phys. Rev. D*, vol. 72, p. 063516, 2005. astro-ph/0507372.

[26] J. Grande, J. Solà, and H. Stefancic, "LXCDM: A Cosmon model solution to the cosmological coincidence problem?," *JCAP*, vol. 08, p. 011, 2006. gr-qc/0604057.

[27] J. D. Barrow and T. Clifton, "Cosmologies with energy exchange," *Phys. Rev. D*, vol. 73, p. 103520, 2006. astro-ph/0604063.

[28] F. E. M. Costa and J. S. Alcaniz, "Cosmological consequences of a possible Λ-dark matter interaction," *Phys. Rev. D*, vol. 81, p. 043506, 2010. arXiv:0908.4251.

[29] J. Grande, J. Solà, S. Basilakos, and M. Plionis, "Hubble expansion and structure formation in the running FLRW model of the cosmic evolution," *JCAP*, vol. 08, p. 007, 2011. arXiv:1103.4632.

[30] D. Bessada and O. D. Miranda, "Probing a cosmological model with a $\Lambda = \Lambda_0 + 3\beta H^2$ decaying vacuum," *Phys. Rev. D*, vol. 88, no. 8, p. 083530, 2013. arXiv:1310.8571.

[31] A. Gómez-Valent, E. Karimkhani, and J. Solà, "Background history and cosmic perturbations for a general system of self-conserved dynamical dark energy and matter," *JCAP*, vol. 12, p. 048, 2015. arXiv:1509.03298.

[32] J. Overduin and F. Cooperstock, "Evolution of the scale factor with a variable cosmological term," *Phys. Rev. D*, vol. 58, p. 043506, 1998. astro-ph/9805260.

[33] C. Moreno-Pulido and J. Solà, "Running vacuum in quantum field theory in curved spacetime: renormalizing ρ_{vac} without $\sim m^4$ terms," *Eur. Phys. J. C*, vol. 80, no. 8, p. 692, 2020. arXiv:2005.03164.

[34] C. Moreno-Pulido and J. Solà Peracaula, "Renormalizing the vacuum energy in cosmological spacetime: implications for the cosmological constant problem," *Eur. Phys. J. C*, vol. 82, no. 6, p. 551, 2022. arXiv:2201.05827.

[35] C. Moreno-Pulido and J. Solà Peracaula, "Equation of state of the running vacuum," *Eur. Phys. J. C*, vol. 82, no. 12, p. 1137, 2022. arXiv:2207.07111.

[36] C. Moreno-Pulido, J. Solà Peracaula, and S. Cheraghchi, "Running vacuum in QFT in FLRW spacetime: The dynamics of $\rho_{\rm vac}(H)$ from the quantized matter fields," 2023. arXiv:2301.05205.

[37] S. Basilakos, N. E. Mavromatos, and J. Solà Peracaula, "Gravitational and Chiral Anomalies in the Running Vacuum Universe and Matter-Antimatter Asymmetry," *Phys. Rev.*, vol. D101, no. 4, p. 045001, 2020. arXiv:1907.04890.

[38] S. Basilakos, N. E. Mavromatos, and J. Solà Peracaula, "Quantum Anomalies in String-Inspired Running Vacuum Universe: Inflation and Axion Dark Matter," *Phys. Lett.*, vol. B803, p. 135342, 2020. arXiv:2001.03465.

[39] N. E. Mavromatos and J. Solà Peracaula, "Stringy-running-vacuum-model inflation: from primordial gravitational waves and stiff axion matter to dynamical dark energy," *Eur. Phys. J. ST*, vol. 230, no. 9, pp. 2077–2110, 2021. arXiv:2012.07971.

[40] N. E. Mavromatos and J. Solà Peracaula, "Inflationary physics and trans-Planckian conjecture in the stringy running vacuum model: from the phantom vacuum to

the true vacuum," *Eur. Phys. J. Plus*, vol. 136, no. 11, p. 1152, 2021. arXiv:2105.02659.
[41] C. Brans and R. Dicke, "Mach's principle and a relativistic theory of gravitation," *Phys. Rev*, vol. 124, p. 925, 1961.
[42] K. Uehara and C. W. Kim, "Brans-dicke Cosmology With the Cosmological Constant," *Phys. Rev. D*, vol. 26, p. 2575, 1982.
[43] P. G. O. Freund, "Kaluza-Klein Cosmologies," *Nucl. Phys. B*, vol. 209, p. 146, 1982.
[44] N. Banerjee and D. Pavon, "A Quintessence scalar field in Brans-Dicke theory," *Class. Quant. Grav.*, vol. 18, p. 593, 2001. gr-qc/0012098.
[45] H. Kim, "Brans-Dicke theory as an unified model for dark matter - dark energy," *Mon. Not. Roy. Astron. Soc.*, vol. 364, pp. 813–822, 2005. astro-ph/0408577.
[46] T. Clifton and J. D. Barrow, "The Power of General Relativity," *Phys. Rev. D*, vol. 72, no. 10, p. 103005, 2005. gr-qc/0509059.
[47] A. E. Montenegro, Jr. and S. Carneiro, "Exact solutions of Brans-Dicke cosmology with decaying vacuum density," *Class. Quant. Grav.*, vol. 24, pp. 313–327, 2007. gr-qc/0609059.
[48] N. Banerjee and D. Pavon, "Holographic dark energy in Brans-Dicke theory," *Phys. Lett. B*, vol. 647, pp. 477–481, 2007. gr-qc/0702110.
[49] G. Olivares, F. Atrio-Barandela, and D. Pavon, "Dynamics of Interacting Quintessence Models: Observational Constraints," *Phys. Rev. D*, vol. 77, p. 063513, 2008. arXiv:0706.3860.
[50] A. Sheykhi, "Interacting new agegraphic dark energy in nonflat Brans-Dicke cosmology," *Phys. Rev. D*, vol. 81, p. 023525, 2010. arXiv:0908.0606.
[51] C. P. Singh, "FRW models with particle creation in Brans-Dicke theory," *Astrophys. Space Sci.*, vol. 338, pp. 411–419, 2012.
[52] P. Kumar and C. P. Singh, "New agegraphic dark energy model in Brans-Dicke theory with logarithmic form of scalar field," *Astrophys. Space Sci.*, vol. 362, no. 3, p. 52, 2017. arXiv:1609.02751.
[53] C. P. Singh and P. Kumar, "Holographic dark energy in Brans-Dicke theory with logarithmic form of scalar field," *Int. J. Theor. Phys.*, vol. 56, no. 10, pp. 3297–3310, 2017. arXiv:1609.01477.
[54] S. Ghaffari, H. Moradpour, I. P. Lobo, J. P. Morais Graça, and V. B. Bezerra, "Tsallis holographic dark energy in the Brans–Dicke cosmology," *Eur. Phys. J. C*, vol. 78, no. 9, p. 706, 2018. arXiv:1807.04637.
[55] P. Mukherjee and S. Chakrabarti, "Exact solutions and accelerating universe in modified Brans–Dicke theories," *Eur. Phys. J. C*, vol. 79, no. 8, p. 681, 2019. arXiv:1908.01564.
[56] J. Solà Peracaula, "Brans–Dicke gravity: From Higgs physics to (dynamical) dark energy," *Int. J. Mod. Phys. D*, vol. 27, no. 14, p. 1847029, 2018. arXiv:1805.09810.
[57] J. de Cruz Pérez and J. Solà Peracaula, "Brans–Dicke cosmology mimicking running vacuum," *Mod. Phys. Lett. A*, vol. 33, no. 38, p. 1850228, 2018. arXiv:1809.03329.
[58] J. Solà Peracaula, A. Gomez-Valent, J. de Cruz Pérez, and C. Moreno-Pulido, "Brans–Dicke Gravity with a Cosmological Constant Smoothes Out ΛCDM Tensions," *Astrophys. J. Lett.*, vol. 886, no. 1, p. L6, 2019. arXiv:1909.02554.
[59] C. P. Singh and S. Kaur, "Matter creation cosmology in Brans-Dicke theory: Observational tests and thermodynamic analysis," *Phys. Rev. D*, vol. 100, no. 8, p. 084057, 2019.
[60] C. P. Singh and S. Kaur, "Probing bulk viscous matter-dominated model in Brans-Dicke theory," *Astrophys. Space Sci.*, vol. 365, no. 1, p. 2, 2020.
[61] J. Solà Peracaula, A. Gómez-Valent, J. de Cruz Pérez, and C. Moreno-Pulido, "Brans–Dicke cosmology with a Λ-term: a possible solution to ΛCDM tensions," *Class. Quant. Grav.*, vol. 37, no. 24, p. 245003, 2020. arXiv:2006.04273.
[62] S. Kaur and C. P. Singh, "Constraints on holographic dark energy model with matter creation in Brans–Dicke theory and thermodynamic analysis," *Phys. Dark Univ.*, vol. 33, p. 100869, 2021.
[63] C. P. Singh and J. Solà Peracaula, "Friedmann cosmology with decaying vacuum density in Brans–Dicke theory," *Eur. Phys. J. C*, vol. 81, no. 10, p. 960, 2021. arXiv:2110.12411.
[64] L. Amendola and S. Tsujikawa, *Dark Energy: Theory and Observations*. Cambridge University Press, 2015.
[65] T. Clifton, P. G. Ferreira, A. Padilla, and C. Skordis, "Modified Gravity and Cosmology," *Phys. Rept.*, vol. 513, pp. 1–189, 2012. arXiv:1106.2476.
[66] A. Avilez and C. Skordis, "Cosmological constraints on Brans-Dicke theory," *Phys. Rev. Lett.*, vol. 113, no. 1, p. 011101, 2014. arXiv:1303.4330.
[67] A. Gómez-Valent and P. Hassan Puttasiddappa, "Difficulties in reconciling non-negligible differences between the local and cosmological values of the gravitational coupling in extended Brans-Dicke theories," *JCAP*, vol. 09, p. 040, 2021. arXiv:2105.14819.
[68] We note that this analysis is different from that presented in in [63] based on a dynamical Λ-term in BD-theory, in which the dependence on H was linear rather than quadratic. The linear model turned out to be not particularly favored.
[69] E. Di Valentino *et al.*, "Snowmass2021 - Letter of interest cosmology intertwined II: The hubble constant tension," *Astropart. Phys.*, vol. 131, p. 102605, 2021. arXiv:2008.11284.
[70] M. G. Dainotti, B. De Simone, T. Schiavone, G. Montani, E. Rinaldi, and G. Lambiase, "On the Hubble constant tension in the SNe Ia Pantheon sample," *Astrophys. J.*, vol. 912, no. 2, p. 150, 2021. arXiv:2103.02117.
[71] E. Di Valentino *et al.*, "Cosmology Intertwined III: $f\sigma_8$ and S_8," *Astropart. Phys.*, vol. 131, p. 102604, 2021. arXiv:2008.11285.
[72] L. Perivolaropoulos and F. Skara, "Challenges for ΛCDM: An update," *New Astron. Rev.*, vol. 95, p. 101659, 2022. arXiv:2105.05208.
[73] E. Abdalla *et al.*, "Cosmology intertwined: A review of the particle physics, astrophysics, and cosmology associated with the cosmological tensions and anomalies," *JHEAp*, vol. 34, pp. 49–211, 2022. arXiv:2203.06142.
[74] M. Dainotti, B. De Simone, G. Montani, T. Schiavone, and G. Lambiase, "The Hubble constant tension: current status and future perspectives through new cosmological probes," 2023. arXiv:2301.10572.
[75] J. Solà Peracaula, A. Gómez-Valent, J. de Cruz Perez, and C. Moreno-Pulido, "Running vacuum against the H_0 and σ_8 tensions," *EPL*, vol. 134, no. 1, p. 19001, 2021. arXiv:2102.12758.

[76] J. Solà Peracaula, A. Gómez-Valent, and J. de Cruz Pérez, "The H_0 tension in light of vacuum dynamics in the Universe," *Phys. Lett.*, vol. B774, pp. 317–324, 2017. arXiv:1705.06723.

[77] A. Gómez-Valent and J. Solà, "Relaxing the σ_8-tension through running vacuum in the Universe," *EPL*, vol. 120, no. 3, p. 39001, 2017. arXiv:1711.00692.

[78] A. Gómez-Valent and J. Solà Peracaula, "Density perturbations for running vacuum: a successful approach to structure formation and to the σ_8-tension," *Mon. Not. Roy. Astron. Soc.*, vol. 478, no. 1, pp. 126–145, 2018. arXiv:1801.08501.

[79] J. Solà Peracaula, J. de Cruz Pérez, and A. Gómez-Valent, "Dynamical dark energy vs. Λ = const in light of observations," *EPL*, vol. 121, no. 3, p. 39001, 2018. arXiv:1606.00450.

[80] J. Solà, A. Gómez-Valent, and J. de Cruz Pérez, "Hints of dynamical vacuum energy in the expanding Universe," *Astrophys. J.*, vol. 811, p. L14, 2015. arXiv:1506.05793.

[81] J. Solà, A. Gómez-Valent, and J. de Cruz Pérez, "First evidence of running cosmic vacuum: challenging the concordance model," *Astrophys. J.*, vol. 836, no. 1, p. 43, 2017. arXiv:1602.02103.

[82] J. Solà Peracaula, J. de Cruz Pérez, and A. Gómez-Valent, "Possible signals of vacuum dynamics in the Universe," *Mon. Not. Roy. Astron. Soc.*, vol. 478, no. 4, pp. 4357–4373, 2018. arXiv:1703.08218.

[83] J. Solà Peracaula, "Running Vacuum and the ΛCDM tensions," *PoS*, vol. CORFU2021, p. 106, 2022, doi:10.22323/1.406.0106.

[84] J. Solà Peracaula, "Tensions in the ΛCDM and vacuum dynamics," *Int. J. Mod. Phys.*, vol. A33, no. 31, p. 1844009, 2018.

[85] J. Solà, A. Gómez-Valent, J. de Cruz Pérez, and C. Moreno-Pulido, "Running vacuum against the H_0 and σ_8 tensions," *EPL*, vol. 134, no. 1, p. 19001, 2021. arXiv:2102.12758.

[86] P. Asimakis, S. Basilakos, N. E. Mavromatos, and E. N. Saridakis, "Big bang nucleosynthesis constraints on higher-order modified gravities," *Phys. Rev. D*, vol. 105, no. 8, p. 084010, 2022. arXiv:2112.10863.

[87] J. Ren and X.-H. Meng, "Dark viscous fluid described by a unified equation of state in cosmology," *Int. J. Mod. Phys. D*, vol. 16, pp. 1341–1348, 2007. astro-ph/0605010.

[88] N. Komatsu and S. Kimura, "Entropic cosmology for a generalized black-hole entropy," *Phys. Rev. D*, vol. 88, p. 083534, 2013. arXiv:1307.5949.

[89] A. Gómez-Valent and J. Solà, "Vacuum models with a linear and a quadratic term in H: structure formation and number counts analysis," *Mon. Not. Roy. Astron. Soc.*, vol. 448, pp. 2810–2821, 2015. arXiv:1412.3785.

[90] H. Hildebrandt *et al.*, "KiDS+VIKING-450: Cosmic shear tomography with optical+infrared data," *Astron. Astrophys.*, vol. 633, p. A69, 2020. arXiv:1812.06076.

[91] A. G. Riess *et al.*, "A Comprehensive Measurement of the Local Value of the Hubble Constant with 1 km/s/Mpc Uncertainty from the Hubble Space Telescope and the SH0ES Team," *Astrophys. J. Lett.*, vol. 934, no. 1, p. L7, 2022. arXiv:2112.04510.

[92] J. M. Bardeen, J. Bond, N. Kaiser, and A. Szalay, "The Statistics of Peaks of Gaussian Random Fields," *Astrophys. J.*, vol. 304, pp. 15–61, 1986.

[93] The dataset employed in this work is very similar to the one used in references [61] and [75] check them for more details.

[94] D. M. Scolnic *et al.*, "The Complete Light-curve Sample of Spectroscopically Confirmed SNe Ia from Pan-STARRS1 and Cosmological Constraints from the Combined Pantheon Sample," *Astrophys. J.*, vol. 859, no. 2, p. 101, 2018. arXiv:110.00845.

[95] A. G. Riess *et al.*, "Type Ia Supernova Distances at Redshift > 1.5 from the Hubble Space Telescope Multi-cycle Treasury Programs: The Early Expansion Rate," *Astrophys. J.*, vol. 853, no. 2, p. 126, 2018. arXiv:1710.00844.

[96] P. Carter, F. Beutler, W. J. Percival, C. Blake, J. Koda, and A. J. Ross, "Low Redshift Baryon Acoustic Oscillation Measurement from the Reconstructed 6-degree Field Galaxy Survey," *Mon. Not. Roy. Astron. Soc.*, vol. 481, no. 2, pp. 2371–2383, 2018. arXiv:1803.01746.

[97] E. A. Kazin *et al.*, "The WiggleZ Dark Energy Survey: improved distance measurements to z = 1 with reconstruction of the baryonic acoustic feature," *Mon. Not. Roy. Astron. Soc.*, vol. 441, no. 4, pp. 3524–3542, 2014. arXiv:1401.0358.

[98] H. Gil-Marín, W. J. Percival, L. Verde, J. R. Brownstein, C.-H. Chuang, F.-S. Kitaura, S. A. Rodríguez-Torres, and M. D. Olmstead, "The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: RSD measurement from the power spectrum and bispectrum of the DR12 BOSS galaxies," *Mon. Not. Roy. Astron. Soc.*, vol. 465, no. 2, pp. 1757–1788, 2017. arXiv:1606.00439.

[99] H. du Mas des Bourboux *et al.*, "The Completed SDSS-IV Extended Baryon Oscillation Spectroscopic Survey: Baryon Acoustic Oscillations with Lyα Forests," *Astrophys. J.*, vol. 901, no. 2, p. 153, 2020. arXiv:2007.08995.

[100] T. M. C. Abbott *et al.*, "Dark Energy Survey Year 1 Results: Measurement of the Baryon Acoustic Oscillation scale in the distribution of galaxies to redshift 1," *Mon. Not. Roy. Astron. Soc.*, vol. 483, no. 4, pp. 4866–4883, 2019. arXiv:1712.06209.

[101] R. Neveux *et al.*, "The completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: BAO and RSD measurements from the anisotropic power spectrum of the quasar sample between redshift 0.8 and 2.2," *Mon. Not. Roy. Astron. Soc.*, vol. 499, no. 1, pp. 210–229, 2020. arXiv:2007.08999.

[102] R. Jiménez, L. Verde, T. Treu, and D. Stern, "Constraints on the equation of state of dark energy and the Hubble constant from stellar ages and the CMB," *Astrophys. J.*, vol. 593, pp. 622–629, 2003. astro-ph/0302560.

[103] J. Simon, L. Verde, and R. Jiménez, "Constraints on the redshift dependence of the dark energy potential," *Phys. Rev.*, vol. D71, p. 123001, 2005. astro-ph/0412269.

[104] D. Stern, R. Jiménez, L. Verde, M. Kamionkowski, and S. A. Stanford, "Cosmic Chronometers: Constraining the Equation of State of Dark Energy. I: H(z) Measurements," *JCAP*, vol. 1002, p. 008, 2010. arXiv:0907.3149.

[105] M. Moresco *et al.*, "Improved constraints on the expansion rate of the Universe up to z 1.1 from the spectroscopic evolution of cosmic chronometers," *JCAP*, vol. 1208, p. 006, 2012. arXiv:1201.3609.

[106] C. Zhang, H. Zhang, S. Yuan, T.-J. Zhang, and Y.-C. Sun, "Four new observational $H(z)$ data from luminous red galaxies in the Sloan Digital Sky Survey data release seven," *Res. Astron. Astrophys.*, vol. 14, no. 10,

pp. 1221–1233, 2014. arXiv:1207.4541.

[107] M. Moresco, "Raising the bar: new constraints on the Hubble parameter with cosmic chronometers at $z \sim 2$," *Mon. Not. Roy. Astron. Soc.*, vol. 450, no. 1, pp. L16–L20, 2015. arXiv:1503.01116.

[108] M. Moresco, L. Pozzetti, A. Cimatti, R. Jiménez, C. Maraston, L. Verde, D. Thomas, A. Citro, R. Tojeiro, and D. Wilkinson, "A 6% measurement of the Hubble parameter at $z \sim 0.45$: direct evidence of the epoch of cosmic re-acceleration," *JCAP*, vol. 1605, no. 05, p. 014, 2016. arXiv:1601.01701.

[109] A. Ratsimbazafy, S. Loubser, S. Crawford, C. Cress, B. Bassett, R. Nichol, and P. Väisänen, "Age-dating Luminous Red Galaxies observed with the Southern African Large Telescope," *Mon. Not. Roy. Astron. Soc.*, vol. 467, no. 3, pp. 3239–3254, 2017. arXiv:1702.00418.

[110] N. Borghi, M. Moresco, and A. Cimatti, "Toward a Better Understanding of Cosmic Chronometers: A New Measurement of H(z) at z $\sim$ 0.7," *Astrophys. J. Lett.*, vol. 928, no. 1, p. L4, 2022. arXiv:2110.04304.

[111] A. Favale, A. Gómez-Valent, and M. Migliaccio, "Cosmic chronometers to calibrate the ladders and measure the curvature of the Universe. A model-independent study," 2023. arXiv:2301.09591.

[112] K. Said, M. Colless, C. Magoulas, J. R. Lucey, and M. J. Hudson, "Joint analysis of 6dFGS and SDSS peculiar velocities for the growth rate of cosmic structure and tests of gravity," *Mon. Not. Roy. Astron. Soc.*, vol. 497, no. 1, pp. 1275–1293, 2020. arXiv:2007.04993.

[113] C. Blake *et al.*, "Galaxy And Mass Assembly (GAMA): improved cosmic growth measurements using multiple tracers of large-scale structure," *Mon. Not. Roy. Astron. Soc.*, vol. 436, p. 3089, 2013. arXiv:1309.5556.

[114] F. Simpson, C. Blake, J. A. Peacock, I. Baldry, J. Bland-Hawthorn, A. Heavens, C. Heymans, J. Loveday, and P. Norberg, "Galaxy and mass assembly: Redshift space distortions from the clipped galaxy field," *Phys. Rev.*, vol. D93, no. 2, p. 023525, 2016. arXiv:1505.03865.

[115] C. Blake *et al.*, "The WiggleZ Dark Energy Survey: the growth rate of cosmic structure since redshift z=0.9," *Mon. Not. Roy. Astron. Soc.*, vol. 415, p. 2876, 2011. arXiv:1104.2948.

[116] F. G. Mohammad *et al.*, "The VIMOS Public Extragalactic Redshift Survey (VIPERS): Unbiased clustering estimate with VIPERS slit assignment," *Astron. Astrophys.*, vol. 619, p. A17, 2018. arXiv:1807.05999.

[117] L. Guzzo *et al.*, "A test of the nature of cosmic acceleration using galaxy redshift distortions," *Nature*, vol. 451, pp. 541–545, 2008. arXiv:0802.1944.

[118] Y.-S. Song and W. J. Percival, "Reconstructing the history of structure formation using Redshift Distortions," *JCAP*, vol. 0910, p. 004, 2009. arXiv:0807.0810.

[119] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Autom. Control*, vol. 19, pp. 716–723, 1974.

[120] D. J. Spiegelhalter, N. G. Best, B. P. Carlin, and A. van der Linde, "Bayesian measures of model complexity and fit," *J. Roy. Stat. Soc.*, vol. 64, p. 583, 2002.

[121] A. R. Liddle, "Information criteria for astrophysical model selection," *Mon. Not. Roy. Astron. Soc.*, vol. 377, pp. L74–L78, 2007. astro-ph/0701113.

[122] A. Lewis, "GetDist: a Python package for analysing Monte Carlo samples," 2019. arXiv:1910.13970.

[123] J. de Cruz Pérez, C.-G. Park, and B. Ratra, "Current data are consistent with flat spatial hypersurfaces in the ΛCDM cosmological model but favor more lensing than the model predicts," 11 2022. arXiv:2211.04268.

[124] B. Ratra, "Tilted spatially nonflat inflation," *Phys. Rev. D*, vol. 106, no. 12, p. 123524, 2022. arXiv:2211.02999.

[125] S. Cao and B. Ratra, "Using lower redshift, non-CMB, data to constrain the Hubble constant and other cosmological parameters," *Mon. Not. Roy. Astron. Soc.*, vol. 513, no. 4, pp. 5686–5700, 2022. arXiv:2203.10825.

[126] S. Cao, J. Ryan, and B. Ratra, "Using Pantheon and DES supernova, baryon acoustic oscillation, and Hubble parameter data to constrain the Hubble constant, dark energy dynamics, and spatial curvature," *Mon. Not. Roy. Astron. Soc.*, vol. 504, no. 1, pp. 300–310, 2021. arXiv:2101.08817.

[127] N. Khadka and B. Ratra, "Determining the range of validity of quasar X-ray and UV flux measurements for constraining cosmological model parameters," *Mon. Not. Roy. Astron. Soc.*, vol. 502, no. 4, pp. 6140–6156, 2021. arXiv:2012.09291.

[128] S. Cao, J. Ryan, N. Khadka, and B. Ratra, "Cosmological constraints from higher redshift gamma-ray burst, H ii starburst galaxy, and quasar (and other) data," *Mon. Not. Roy. Astron. Soc.*, vol. 501, no. 1, pp. 1520–1538, 2021. arXiv:2009.12953.

[129] N. Khadka and B. Ratra, "Using quasar X-ray and UV flux measurements to constrain cosmological model parameters," *Mon. Not. Roy. Astron. Soc.*, vol. 497, no. 1, pp. 263–278, 2020. arXiv:2004.09979.

[130] C.-G. Park and B. Ratra, "Observational constraints on the tilted flat-XCDM and the untilted nonflat XCDM dynamical dark energy inflation parameterizations," *Astrophys. Space Sci.*, vol. 364, no. 5, p. 82, 2019. arXiv:1803.05522.

[131] C.-G. Park and B. Ratra, "Observational constraints on the tilted spatially-flat and the untilted nonflat ϕCDM dynamical dark energy inflation models," *Astrophys. J.*, vol. 868, no. 2, p. 83, 2018. arXiv:1807.07421.

[132] C.-G. Park and B. Ratra, "Using the tilted flat-ΛCDM and the untilted non-flat ΛCDM inflation models to measure cosmological parameters from a compilation of observational data," *Astrophys. J.*, vol. 882, p. 158, 2019. arXiv:1801.00213.

Photochemical & Photobiological Sciences
https://doi.org/10.1007/s43630-023-00393-6

ORIGINAL PAPERS

Selectively probing ferric ions in aqueous environments using protonated and neutral forms of 7-azaindole as a multiparametric chemosensor

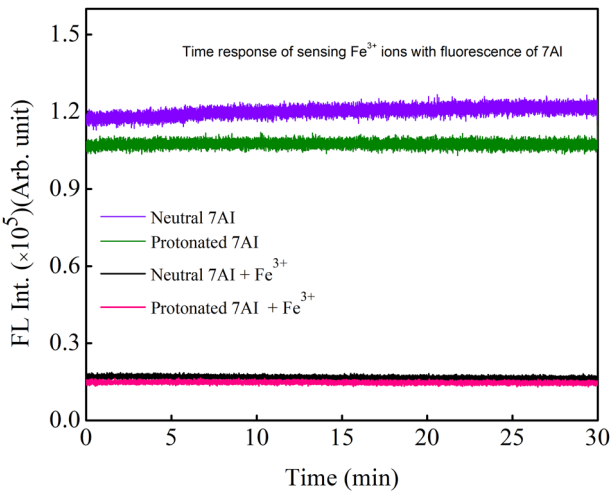
Mohan Singh Mehata[1] · **Aneesha**[1]

Received: 1 October 2022 / Accepted: 5 February 2023
© The Author(s), under exclusive licence to European Photochemistry Association, European Society for Photobiology 2023

Abstract

7-azaindole (7AI) dimer is a model molecule for DNA study and understanding the mutagenic behavior based on the excited-state proton transfer process in hydrogen-bonded networks. The neutral and protonated forms of 7AI monomer with significant fluorescence (FL) intensity fit the fluorescent sensor strategy to recognize selective metal ions. Out of several metal ions (Fe^{3+}, Al^{3+}, Fe^{2+}, Pb^{2+}, Ba^{2+}, Ni^{2+}, Zn^{2+}, Mg^{2+}, Ca^{2+}, Cu^{2+}, Hg^{2+} and Cd^{2+}), the absorption, fluorescence and fluorescence lifetime of 7AI in the aqueous medium are selectively sensitive to the ferric (Fe^{3+}) ions. The absolute value of absorption intensity increases linearly with concentration of a particular metal ions. FL intensity of both the forms of 7AI decreases gradually with Fe^{3+} ions and trails the linear Stern–Volmer relation. The formation of non-fluorescent complexes was confirmed with Benesi-Hildebrand and Job plots, along with FL and FL decays. The FL lifetime of the protonated form of 7AI, which is 0.83 ± 0.01 ns, is nearly constant with Fe^{3+} ions concentrations, confirming the static quenching mechanism. The limit of detection (LoD) of Fe^{3+} ions over the long range of 16–363 μM for the neutral and protonated forms of 7AI is 0.46 ± 0.02 and 0.49 ± 0.02 μM, respectively, estimated using FL spectra. Additionally, the linear plot of absorbance with Fe^{3+} ions of both the forms of 7AI can also act as a calibration curve with very close LoDs, as obtained by FL spectra. Thus, the multi-parameters-based probe for detecting the Fe^{3+} ions over long-range in real aqueous environments with operational, high sensitivity, fast response (< 5 s), and good selectivity (over 12 metal ions) is undoubtedly a superior approach over other methods.

Graphical abstract



Extended author information available on the last page of the article

Published online: 18 February 2023

Keywords Neutral and protonated forms of 7AI · Absorption · Fluorescence · Fluorescence lifetimes · Metal ions · Aqueous environment

1 Introduction

7-Azaindole (7AI), which is a model molecule, is a chromophoric fraction of amino acid 7-azatryptophan (7AT) and became an alternative for the study of proteins; hence the deoxyribonucleic acid (DNA) [1–6]. This is because of the similarity of the DNA base pairs, guanine–cytosine (G–C) and adenine–thymine (A–T) and 7AI dimer. A double hydrogen bond network is formed between two molecules of 7AI, denoted as $7(AI)_2$, which is identical to A–T and G–C pairs. Understanding the change in behavior of intramolecular H-bonded networks following the photoexcitation of $7(AI)_2$ and A–T/G–C pairs could be crucial in diagnosing mutagenic behavior in genetics [6]. The H-bonding formed between the functional groups is a primary and essential driving force for the electronically excited-state double proton transfer. There has been a long debate in understanding the excited-state double proton transfer of $7(AI)_2$ [6]. Photoinduced excited-state double proton jump, i.e., tautomerization, may result from the concerted double proton transfer [5, 7] or the formation of intermediate photoproducts during the process [8–10]. $7(AI)_2$ has a weak fluorescence (FL) at room temperature but has a strong FL at a cryogenic temperature [11]. The optical properties of 7AI are distinguishable concerning the 7AT; it shows higher quantum yield, longer lifetime and spectral shifts with solvents [12]. The change in dipole, i.e., the charge transfer following photoexcitation, is quite significant in the excited state and sensitive to the solvents [13]. In alcoholic solvents, 7AI exhibits well-separated dual FL bands corresponding to the usual species and tautomer formed by the excited-state double proton transfer proposed by Kasha et al. [1, 11]. 7AI and its derivatives were studied by several groups using different environmental conditions [14–19]. The excited state proton transfer reaction, which usually occurs in a nanosecond to the picosecond time scale, is considered one of the important photoinduced processes of molecular systems for both fundamental and application aspects [20–28].

In addition to understanding mutation in DNA following the excited-state proton transfer process, the derivatives of azaindole are considered biocompatible, biodegradable, and less harmful (to the human body) probe to monitor cyanide and iron (III) metal ions [18, 19]. Iron is regarded as the fourth most abundant transition metal on earth in a living organism, including humans as micronutrients. Iron (ferric and ferrous forms) is one of the essential metals for electron transportation in the respiratory chain and enzymatic reactions, as a significant contribution (~ 70%) of iron within RBCs helps in oxygen transportation for tissues. The average amount of iron in the adult is about 3–5 g depending on the body weight, i.e., 50–60 mg of iron/kg, and cellular concentration is about 50–100 mM. Consuming iron in excess and insufficient amounts harms cellular functioning and creates health issues in the form of several diseases, including cancer, hepatitis, hemochromatosis and diarrhea [29–31]. For example, excessive amounts of ferric (Fe^{3+}) ions are responsible for the dysfunction of vital organs [30, 31].

Therefore, developing and finding fluorescent chemosensors for targeted metal ions with high selectivity, sensitivity, fast response, nondestructive, and well-controlled have recently received much attention. Various chemosensors, including organic molecules, quantum dots and 2D materials, have been proposed as a probe for detecting numerous metal ions (Cu^{2+}, Al^{3+}, Cr^{3+}, Pb^{2+}, Fe^{3+}, etc.), explosives (TNT) and halide ions (Cl^-, Br^- and I^-) based on acceleration and deceleration of energy transfer, electron transfer and proton transfer processes along with dynamic/static quenching behavior [32–42]. Here, we have proposed a commercially available and well-known 7AI molecule as a fluorescent probe for detecting (with fast response) selective metal ions, Fe^{3+} ions, in an aqueous environment with high selectivity and sensitivity. The selected probe molecule has a quantum yield of ~ 6% when it is neutral and ~ 3% when it is protonated.

2 Materials and methods

All chemicals used were of analytical grade and used without further purification. The probe molecule 7AI (98%) and the metal ions (99.8%), namely iron (III) perchlorate hexahydrate, calcium (II) perchloride tetrahydrate, lead (II) perchlorate trihydrate, aluminium (III) perchlorate nonahydrate, magnesium (II) perchloride hexahydrate, cadmium (II) chloride hydrate, iron (II) perchlorate hydrate, cobalt (II) perchlorate hexahydrate, zinc (II) perchlorate hexahydrate, nickel (II) perchlorate hexahydrate, barium (II) perchlorate anhydrous and mercury (II) perchlorate hydrate used in the experiment were purchased from Sigma-Aldrich. Deionized/ultrapure water (18.2 MΩ•cm) and real/tap water were used in the experiments. Samples were prepared at neutral pH (~ 7.0) and acidic pH (~ 2.4). The probe 7AI concentration was optimized to 50 µM. The concentration of metal ions in water was prepared and varied from 0.0 to 363 µM.

The steady-state absorption spectra were collected with a double beam spectrometer Lambda 750 (Perkin Elmer) and FL spectra with Fluorolog-3 (Horiba Scientific Inc.) mounted with double monochromators for excitation/emission and 450 W xenon lamp. In addition, 10 × 10 mm quartz cell was used to contain liquid samples for absorption and FL measurements. The slit width for absorption was kept at 2.0 nm for all measurements, whereas for FL measurement, the excitation and emission slit widths were 1.5 and 2.0 nm, respectively. The data were recorded and analyzed with the Origin software supplied with the systems. To measure FL decays, a TCSPC (time-correlated single-photon counting) system (model DeltaFlex-01-DD) of Horiba Jobin Yvon, IBH Ltd., armed with a picosecond photon detection module (PPD 850) and excitation source as DD (Delta Diode) of 320 nm with a repetition rate of 8 MHz was used. The observed FL decay curves were analyzed with the least square fitting method. All the measurements, i.e., absorption, FL, and FL decays, were obtained at the same concentration for 7AI.

3 Results and discussion

The absorption spectrum of 7AI in water exhibits bands at 288 nm, as shown in Fig. 1a, b. The longer wavelength band is assigned to the π–π^* transition of the azomethine group. In contrast, the shorter wavelength band, which arises at 215 nm, is regarded as having higher electronic transitions of benzene rings. The longer wavelength band, which appeared between 250 and 300 nm, comprises two overlapping well-defined electronic states, ^1L_a and ^1L_b bands, which are well separated in commonly used non-polar solvents [13]. The lower energy ^1L_a state takes parts in the reaction more strongly with solvent than the higher energy ^1L_b state. Accordingly, 7AI behaves differently in the ground

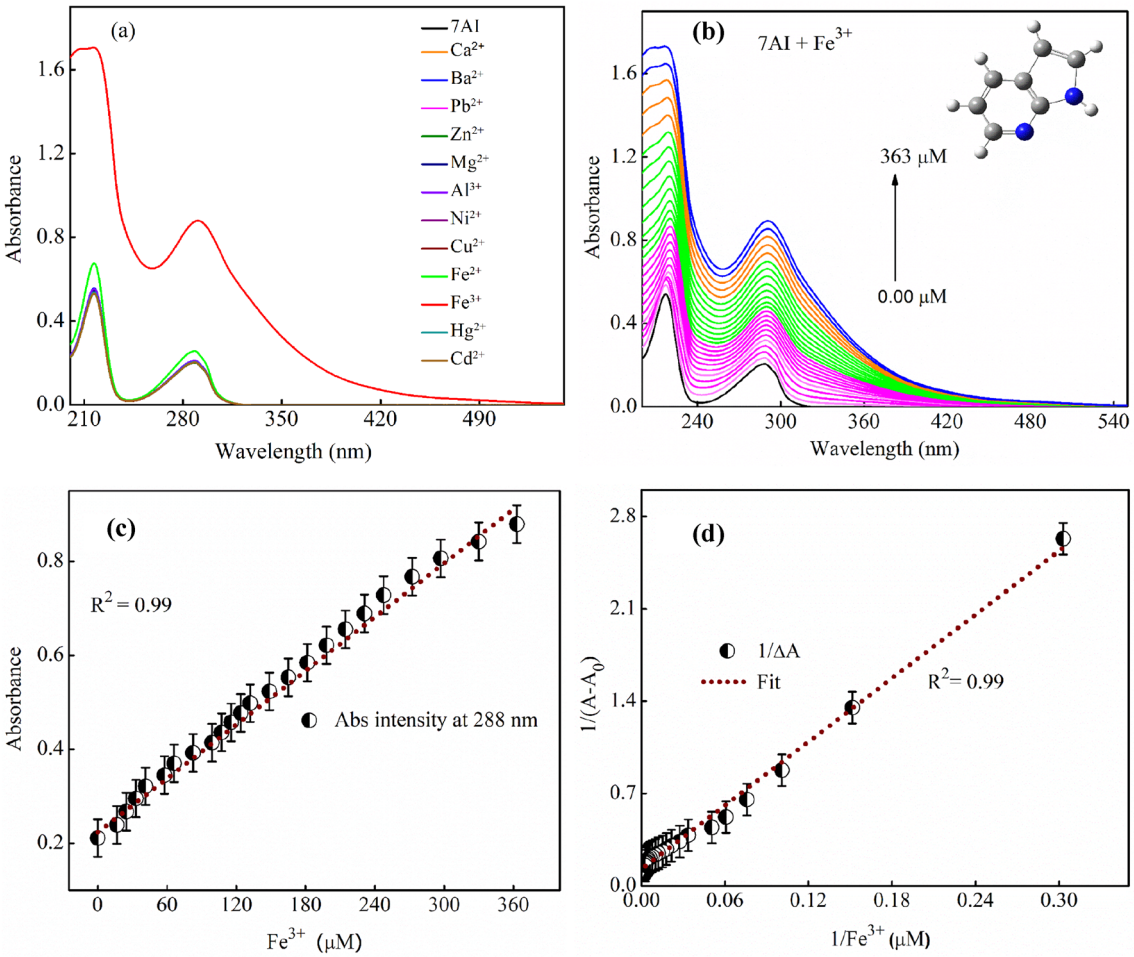


Fig. 1 Absorption spectra of the neutral form of 7AI (50 µM) in an aqueous medium in the absence and presence of various metal ions (363 µM) (**a**), in the presence of different concentrations of Fe^{3+} ions (**b**), a plot between absorption intensity and Fe^{3+} ion concentrations (**c**) and the B–H plot of 1/ΔA vs. 1/ Fe^{3+} ion concentrations (**d**). The dotted red lines of **c** and **d** represent the linear fit. The inset of (**b**) shows the chemical structure of 7AI

and excited states with solvents; hence a change in dipole occurs [5, 13]. The chemical structure of 7AI is shown in the inset of Fig. 1b.

The absorption spectra of the neutral form of 7AI with numerous metal ions (Fe^{3+}, Al^{3+}, Fe^{2+}, Pb^{2+}, Ba^{2+}, Ni^{2+}, Zn^{2+}, Mg^{2+}, Ca^{2+}, Cu^{2+}, Hg^{2+} and Cd^{2+}) were recorded, and are shown in Fig. 1a, which alters significantly with the selective metal ions. The absorption intensity increases considerably in the presence of Fe^{3+} ions, demonstrating that probe 7AI is selectively sensitive to the Fe^{3+} ions. However, the absorption spectra of 7AI are nearly the same in the presence of various other metal ions (Fig. 1a), or the change in the peak position and intensity is minimal. Accordingly, the titration experiments were carried out for the selective metal ions, as shown in Fig. 1b. The absorption intensity increases upon adding a small amount (16.5 µM) of metal ions, particularly Fe^{3+} ions. Further, upon adding the metal ions, the absorption intensity of the longer and shorter wavelength bands increases monotonically until the concentration reaches 363 µM. The increase in absorption intensity of 7AI with Fe^{3+} ions exhibits a nearly linear relation, as shown in Fig. 1c. Thus, the increases in absorption intensity in metal ions solution lead to the rise in transition probability. A plot of $1/(A-A_0)$ vs. $1/[Fe^{3+}]$ [22] following the Benesi–Hildebrand (B–H) relation (Eq. 1) is shown in Fig. 1d.

$$\frac{1}{A-A_0}=\frac{1}{A-A_{\max}}+\left[\frac{1}{K\left(Fe^{3+}\right)}+1\right] \tag{1}$$

A_0 is the absorbance of the neutral form of 7AI without metal ions and A is with different Fe^{3+} ion concentrations. A_{max} is the absorbance with a maximum Fe^{3+} ions concentration. The term K is the binding constant estimated from the intercept/slope of the plot and comes out to be 1.59×10^4 M^{-1}. The linear B–H curve and binding constant indicate a 1:1 complexation between the neutral form of 7AI and Fe^{3+} ions with a significant binding constant. The formation of 1:1 complexation is further supported by the Job plot with an intersect (peak) value of 0.55, as shown in Fig. S1 of supplementary information (SI).

Figure 2 shows the FL of 7AI with and without Fe^{3+} ions in an aqueous medium. In an aqueous medium without metal ions, the FL band maximum appeared at 392 nm and is independent of the excitation wavelength. The FL is assigned to the usual FL of the neutral form of 7AI. The FL of 7AI in water is relatively redshifted compared to the non-polar solvents [13]. Unlike in water, in an alcoholic solvent, in addition to the usual FL, a large Stokes shifted band is also observed corresponding to the excited state proton transfer due to a strong H-bonding network formed between the two functional groups of 7AI and the alcohol molecules. The intensity of the unusual band is relatively small. The additional FL band is not observed in water or is relatively small,

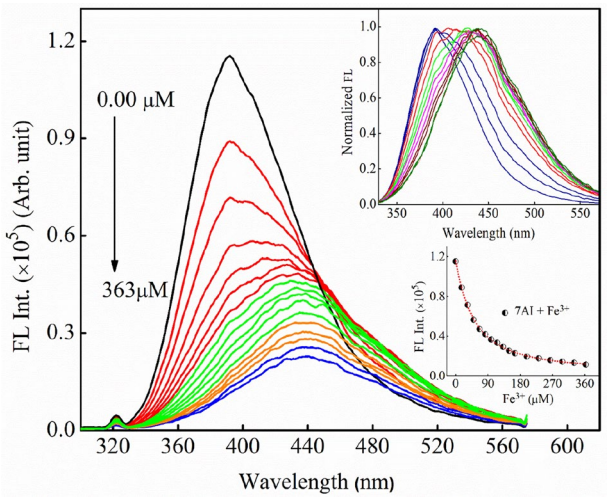


Fig. 2 Fluorescence spectra of the neutral form of 7AI in an aqueous medium in the presence and absence of different concentrations of Fe^{3+} ions. The insets show peak normalized FL spectra and a plot of FL intensity as a function of Fe^{3+} ions concentrations (0.0–363 µM). The excitation wavelength was 320 nm

though the H-bonded proton transfer cannot be neglected. The FL quantum yield (QY) of 7AI calculated from Eq. 2 comes out to be ~ 6% in water.

$$QY_S=QY_R\times\frac{FL(\text{Area})_S}{FL(\text{Area})_R}\times\frac{Abs_R}{Abs_S}\times\left(\frac{\eta_S}{\eta_R}\right)^2 \tag{2}$$

QY_R and QY_S denote the QY of reference taken as quinine sulfate (0.546) and sample 7AI, respectively. FL and Abs represent the FL and absorbance and η is the refractive indices for the solvents used for the reference and the 7AI.

Keeping in mind the observed change in the absorption spectrum/intensity of 7AI with metal ions, fluorescence spectra of 7AI were measured in the absence and presence of various metal ions, e.g., Fe^{3+}, Al^{3+}, Fe^{2+}, Pb^{2+}, Ba^{2+}, Ni^{2+}, Zn^{2+}, Mg^{2+}, Ca^{2+}, Cu^{2+}, Hg^{2+} and Cd^{2+}, and obtained results are depicted in Fig. S2 of (SI), which offers a drastic change in the FL intensity in the presence of Fe^{3+} ions. The FL intensity decreases significantly upon adding a small amount (16.5 µM) of Fe^{3+} ions in FL titration. Further adding Fe^{3+} ions and increasing the concentration upto 363 µM, the FL intensity reduced monotonically (nearly 90%) and became minimal. The FL intensity decreases exponentially with Fe^{3+} ions concentration, as shown in the insets of Fig. 2. With a decrease in FL intensity, the band maximum shifted gradually towards a longer wavelength region from 392 to 439 nm due to the protonation/deprotonation of 7AI. The normalized FL spectra of 7AI are shown in the inset of Fig. 2. The spectral shift as a function of concentration (Fig. S3 of SI) could also be a parameter for detecting Fe^{3+} ions in aqueous media.

Further, steady-state measurements of the protonated form of 7AI were performed in the same manner as for the neutral form of 7AI with different metal ions. The absorption spectra of the protonated form of 7AI with numerous metal ions (Fe^{3+}, Al^{3+}, Fe^{2+}, Pb^{2+}, Ba^{2+}, Ni^{2+}, Zn^{2+}, Mg^{2+}, Ca^{2+}, Cu^{2+}, Hg^{2+} and Cd^{2+}) are shown in Fig. 3a, which alters significantly with the selective metal ions. The absorption titration of the protonated form of 7AI with Fe^{3+} ions concentration is shown in Fig. 3b. In the absence of metal ions, the absorption bands appeared at 215 and 288 nm, nearly the same as observed for the neutral form of 7AI. In the presence of Fe^{3+} ions, the intensity of both absorption bands increases monotonically along with a slight shift (2 nm) towards a longer wavelength region. The plot of absorption intensity at 288 nm with Fe^{3+} ions is approximately linear, as shown in Fig. 3c. Thus, a linear plot of the absolute parameter (absorbance) with metal ions can act as a calibration curve for monitoring the Fe^{3+} ions in an aqueous medium with high accuracy using both the neutral and protonated forms of 7AI. The detection limit (LoD) for Fe^{3+} ions for neutral and protonated forms of 7AI estimated at 3.3 σ/K is 9.20 ± 0.05 and 9.18 ± 0.05 μM, respectively. Here, σ and K signify the standard deviation (which were estimated from the measurement of 10 blank samples of absorption intensity) and the slope of the plots (see Figs. 1c and 3c), respectively. The B–H curve $1/(\Delta A)$ vs. $1/[Fe^{3+}]$ shown in Fig. 3d is nearly linear, indicating 1:1 complexation between 7AI and Fe^{3+} ions, in the same manner, as for the neutral form of 7AI. This suggests the formation of complexes of 7AI with metal ions is quite strong. The estimated binding constant (K) is 1.49×10^4 M^{-1}. The Job plot (with a intersect value of 0.55) also confirmed the formation of 1:1 complexation between the protonated form of 7AI and metal ions, as shown in Fig. S4 of SI. Thus, the binding of the protonated 7AI with the metal ion has almost the same value as of neutral form, indicating that the bond strength is nearly in the same order of magnitude.

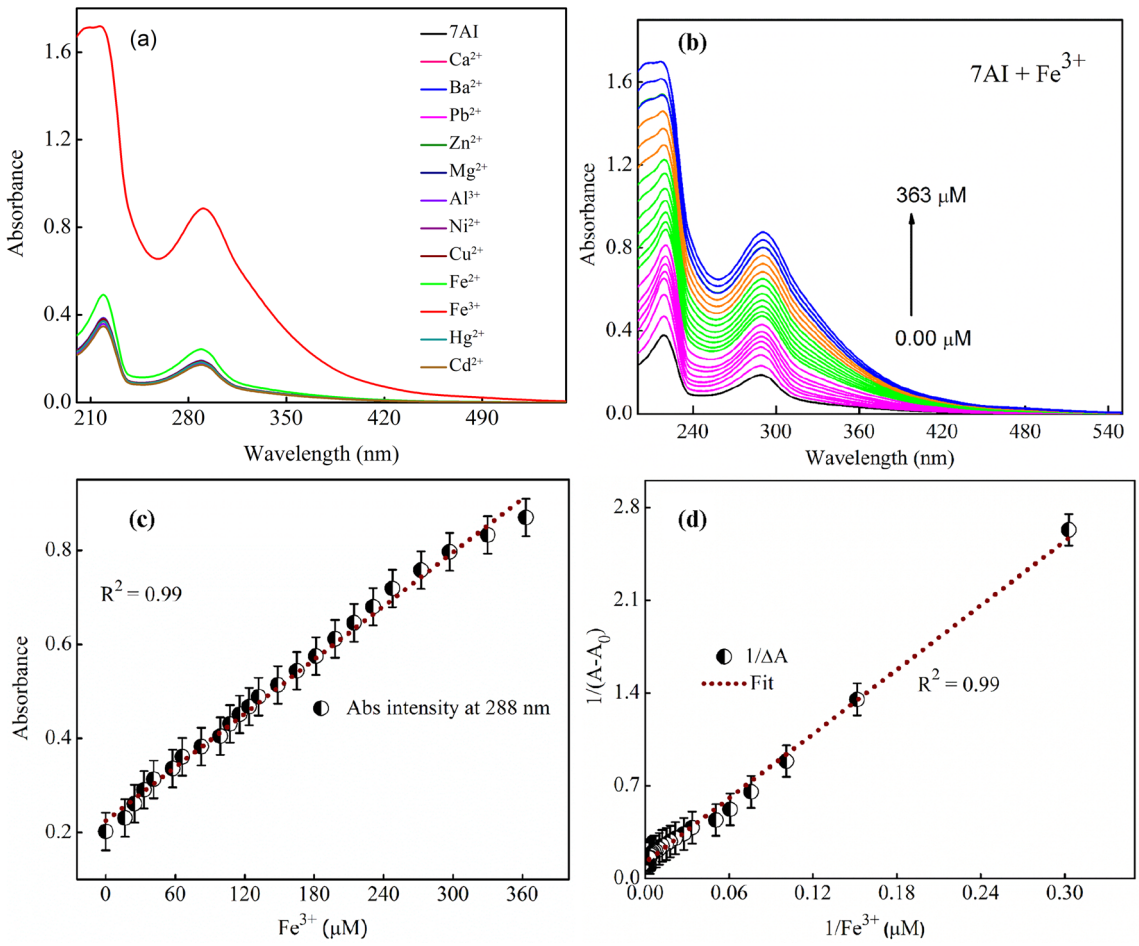


Fig. 3 Absorption spectra of the protonated form of 7AI (50 μM) in an aqueous medium in the absence and presence of various metal ions (363 μM) (**a**), in the presence of different concentrations of Fe^{3+} ions (**b**), a plot between absorption intensity and Fe^{3+} ion concentrations (**c**) and the B–H plot of 1/ΔA *vs.* 1/ Fe^{3+} ion concentrations (**d**). The dotted lines of *c* and *d* represent the linear fit

Figure 4 shows the FL spectra of the protonated form of 7AI in water and peaking at 442 nm, which has nearly 3% QY. The observed FL spectrum is the same irrespective of the excitation wavelength. The large Stokes shifted FL band is assigned to the cationic form of 7AI and originated due to the proton transfer reaction, i.e., the formation of protonated form. The FL spectra were also recorded with various metal ions and the results are presented in Fig. S5 of SI, which offers a dramatic change in the presence of Fe^{3+} ions. The FL titration was also conducted for various Fe^{3+} ion concentrations in the aqueous medium, considering 7AI as a fluorescent probe. Upon adding a minor amount of Fe^{3+} ions, the FL intensity was quenched significantly. The FL intensity decreases exponentially, as shown in the inset of Fig. 4. With increasing Fe^{3+} ions concentration in the aqueous solution of 7AI, the FL intensity diminishes monotonically and quenched nearly 84% with 363 µM of Fe^{3+} ions. Interestingly, the FL band position does not show any considerable shift, unlike that observed for the neutral form of 7AI. Such FL turn-off is explained due to the formation of a non-fluorescent 1:1 complex of 7AI and Fe^{3+} ions, as confirmed by the B-H and Job plots. In addition, a sharp band observed around 361, which is assigned to the water Raman peak, also decreases monotonically with increasing concentration of metal ions and can be considered an additional parameter to sense the metal ions.

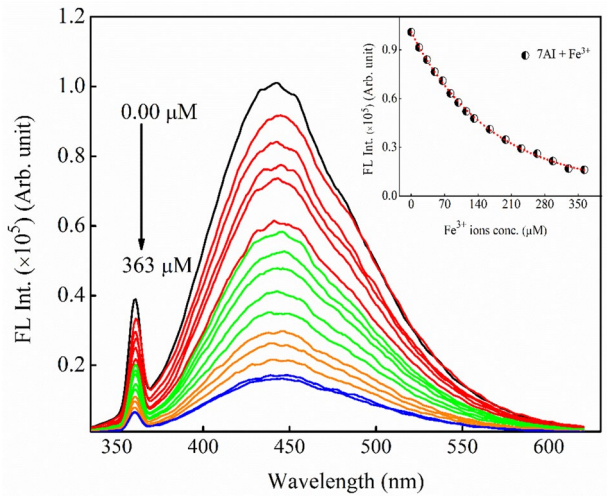


Fig. 4 Fluorescence spectra of the protonated form of 7AI in an aqueous medium in the absence and presence of different concentrations of Fe^{3+} ions. The inset shows the plot of fluorescence intensity as a function of Fe^{3+} ion concentrations (0.0–363 µM). The excitation wavelength was 320 nm

To obtain the limit of detection of the probe and to understand the quenching mechanism, the S–V (Stern–Volmer) relation given in Eq. (3) was plotted for FL intensity quenching of 7AI with Fe^{3+} ion concentrations and the obtained plots for neutral and cationic forms of 7AI are shown in Fig. 5(a & b).

$$F_0/F = 1 + K_{SV}[Q] \tag{3}$$

F and F_0 are the FL intensities of 7AI with or without quencher (Fe^{3+}), respectively. The Fe^{3+} ions molar concentration is denoted by Q and S–V constant by K_{SV} (quenching constant). The S–V plots for both forms of 7AI are linear for the measured range of concentrations. The S-V constant calculated from the linear relationship shown in Fig. 5a, b was obtained to be 2.4×10^4 M^{-1} and 1.6×10^4 M^{-1}, respectively.

The detection limit (LoD) for selectively detected Fe^{3+} ions for neutral and protonated forms of 7AI estimated with 3σ/K is 0.46 ± 0.02 and 0.49 ± 0.02 µM, respectively. Therefore, this claim might be more accurate than the reported

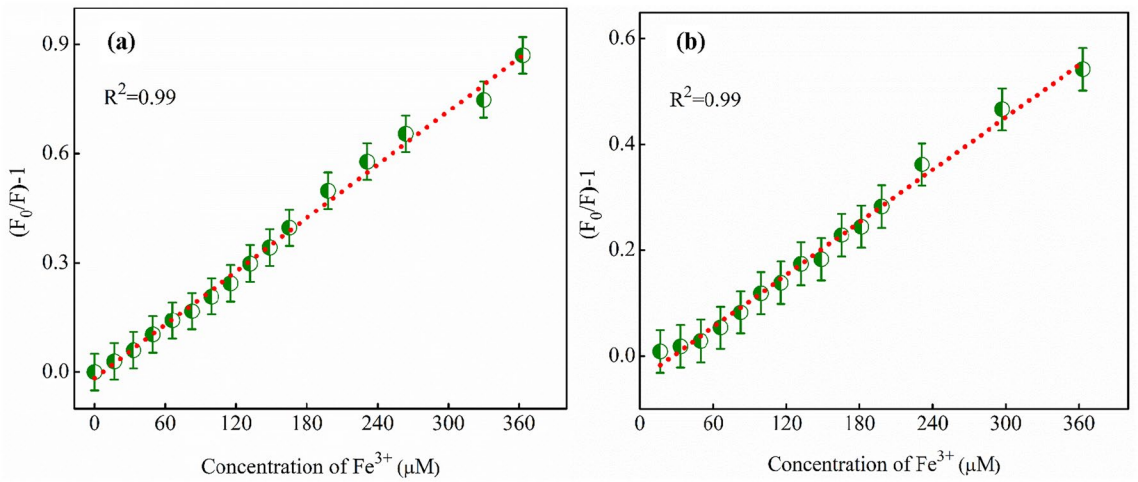


Fig. 5 Stern–Volmer plots between $(F/F_0) - 1$ *vs.* Fe^{3+} ion concentrations of neutral (**a**) and protonated (**b**) forms of 7AI in the aqueous medium. The red dotted line in *a* and *b* represents the linear fit

for various metal ions sensors [19, 34, 35]. The observed concentration ranges of the metal ions, which is from 17 to 363 µM, are quite extensive as compared to the 1–100 µM with LoD of 0.96 µM [35] and 2–25 µM with LoD of 56 µM [19]. If the sample preparation method can improve, the measured concentration range may be further lowered than 17 µM and increased above 363 µM with higher sensitivity.

Time-dependent changes in absorption and fluorescence intensities were examined after adding Fe^{3+} (363 µM) to the solution of 7AI. The obtained plots are shown in Fig. 6a, b. It is to be mentioned that the FL intensity diminished, as demonstrated in Figs. 2 and 4, just after adding Fe^{3+} ions. There is no time lag (in min) between completing the reaction and lowering FL intensity with Fe^{3+} ions. Thus, the response time of the decrease of FL intensity is considered < 5 s, as the mixing was done manually (Fig. 6b). Similarly, the absorption intensity increases upon adding (363 µM) Fe^{3+} ions (see Figs. 1b and 3b), with a time response of < 5 s; after that, the intensity remains constant with time (Fig. 6a). Thus, the response time of the sensor is considered to be very fast.

The change in absorption and FL intensity with varying pH from acidic to basic (pH ~ 2–11) were also investigated. The obtained results are shown in Fig. 7, which indicates the responsiveness of the probe with the pH of the solution, i.e., variation in FL intensity. The variation in the FL intensity at a particular wavelength (at 390 nm) is observed due to protonation and deprotonation during the change in pH.

Time-resolved FL decays were measured to examine photoexcitation dynamics of the neutral and protonated forms of 7AI with and without Fe^{3+} ions at 300 nm excitation by monitoring FL maxima. The decay curves of the neutral

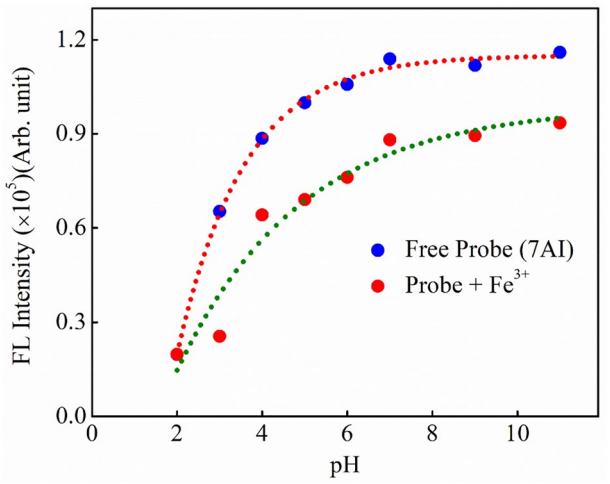


Fig. 7 Change of fluorescence intensity at 390 nm of 7AI as a function of pH in the absence and presence of Fe^{3+} ions (99 µM) in an aqueous solution

form of 7AI in water follow a bi-exponential fit (Eq. 4), as a single exponential fit does not give a good fitting and χ^2 value.

$$I(t) = \sum_{i=1}^{2} a_i \exp(-t/\tau_i) \quad (4)$$

τ_i denotes the decay times and a_i the corresponding amplitude. The obtained decay parameters are summarized in Tables 1 and 2. In the bi-exponential fitted decays, the shorter lifetime component of 0.66 ns has the highest contribution of 80%, whereas the longer lifetime component of 1.06 ns has only 20%. The shorter lifetime is assigned to the

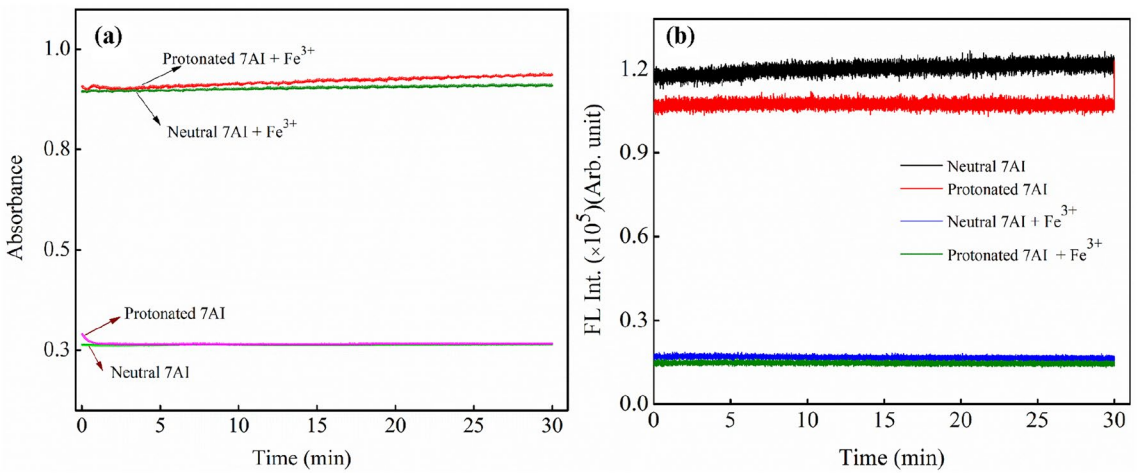


Fig. 6 Time-dependent absorption and fluorescence intensities of neutral and protonated forms of 7AI (50 µM) in the presence of Fe^{3+} ions (363 µM) in an aqueous medium. The change in absorption intensity at 288 nm (**a**) and FL intensity at 390 nm (neutral form) and 442 nm (for protonated form) with a function of time upon adding Fe^{3+} ions (**b**) show the sensing response time of the probe to metal ions

Table 1 Fluorescence lifetime and pre-exponential factors of the neutral form of 7AI in the absence and presence of different concentrations of Fe^{3+} ions. Excitation and emission wavelengths were 300 and 390 nm, respectively

7AI (μM)	τ_1 (ns)	τ_2 (ns)	a_1	a_2	τ_{av} (ns)	χ^2
00	0.66 ± 0.01	1.06 ± 0.02	80.12	19.88	0.77	1.43
16	0.71 ± 0.01	3.62 ± 0.18	97.99	2.01	0.98	1.49
33	0.74 ± 0.01	5.48 ± 0.15	97.76	2.24	1.43	0.96
66	0.78 ± 0.01	6.59 ± 0.17	97.08	2.92	1.96	1.13
99	0.80 ± 0.01	6.74 ± 0.14	96.05	3.95	2.33	1.07
165	0.82 ± 0.01	7.55 ± 0.15	95.53	4.47	2.97	1.09
297	0.83 ± 0.02	7.07 ± 0.15	94.95	5.05	2.78	1.17

Table 2 Fluorescence lifetime and pre-exponential factors of the protonated form of 7AI in the absence and presence of different concentrations of Fe^{3+} ions. Excitation and emission wavelengths were 300 and 420 nm, respectively

7AI (μM)	τ_1 (ns)	τ_2 (ns)	a_1	a_2	τ_{av} (ns)	χ^2
00	0.83 ± 0.01	3.87 ± 0.13	97.10	2.90	1.20	1.36
16	0.85 ± 0.01	4.26 ± 0.14	97.28	2.72	1.27	1.24
33	0.85 ± 0.01	4.41 ± 0.20	97.16	2.84	1.32	1.54
66	0.84 ± 0.01	2.18 ± 0.11	93.27	6.73	1.39	1.59
99	0.85 ± 0.01	2.60 ± 0.25	94.61	5.39	1.41	1.55
165	0.84 ± 0.01	2.25 ± 0.08	92.42	7.58	1.94	1.68
297	0.88 ± 0.01	2.38 ± 0.60	91.96	8.04	1.17	1.63

normal FL of 7AI. The bi-exponential decay has been considered to improve the fitting in the same manner as reported [12]. The second-lifetime component is not very clear in water but is assumed to originate from the proton transfer species.

Once a small amount of Fe^{3+} ions concentrations are added, both the lifetime components increase, as given in Table 1. The main lifetime increases from 0.66 ns to 0.83 ns, whereas the second lifetime increases from 1.06 ns to 7.55 ns. Thus, lifetimes increase monotonically with increasing Fe^{3+} ion concentrations until it approaches 165 μM. The average lifetime calculated using Eq. 5 increases gradually with Fe^{3+} ions concentration (Table 1).

$$\tau_{av} = \sum_{i=1}^{2} (\alpha_i \tau_i^{\,2}) / \sum_{i=1}^{2} (\alpha_i \tau_i) \tag{5}$$

The average lifetime of 0.77 ns in the absence of metal ions became 2.97 ns in the presence of 165 μM of Fe^{3+} ions concentration. At the concentration of 297 μM, the main lifetime component approaches 0.83 ± 0.02 ns. The longer lifetime decrease slightly results in a decrement in the average lifetime. When measured at 450 nm, the decay again shows a bi-exponential fit with a lifetime of 0.85 ± 0.01 ns (89.64%) and 1.49 ± 0.05 ns (10.36%).

The main component of the observed lifetime is nearly the same as of the protonated form of 7AI. Similarly, the decay times of protonated form of 7AI are also measured in the absence and presence of Fe^{3+} ion concentrations (Table 2). The observed decays are well-fitted with a bi-exponential fit. In the absence of metal ions, the lifetimes are 0.83 and 3.87 ns, and the contribution of the longer component is only 1–3%. Thus, the lifetime of the protonated form of 7AI is considered to be 0.83 ns. This varies from 0.83 to 0.88 ns in the presence of Fe^{3+} ions concentration (Table 2). Thus, the main lifetime component of the protonated form of 7AI is considered to be nearly constant with an error of ± 0.03 ns in the presence of metal ions (varies from 0 to 297 μM). However, the average lifetime varies from 1.20 to 1.94 ns due to variations in the longer lifetime component and the increase of a_2 [12]. The variation in τ_1, τ_2 and τ_{av} of both the forms of 7AI with ferric ion concentrations are demonstrated in Fig. S6a, b of SI. It is to be mentioned that the lifetime supports the steady-state results of forming a non-emitting complex between the probe and the quencher. Thus, both the steady-state and time-resolved results exhibit that the quenching mechanism is prominently static rather than dynamic. However, dynamic quenching cannot preclude completely.

Typical fitted overlapped decay curves of the neutral and protonated forms of 7AI in the absence and presence of Fe^{3+} ion concentrations are shown in Fig. 8 and Fig. S7 of SI, respectively. It is to be mentioned that small baseline signals (about 10–20 counts) in different decay profiles differ because of the experimental uncertainty. The K_q, the quenching rate constant estimated with Eq. 6, comes out to be 3.1×10^{13} M^{-1} s^{-1} and 1.4×10^{13} M^{-1} s^{-1} for neutral and protonated forms of 7AI, respectively.

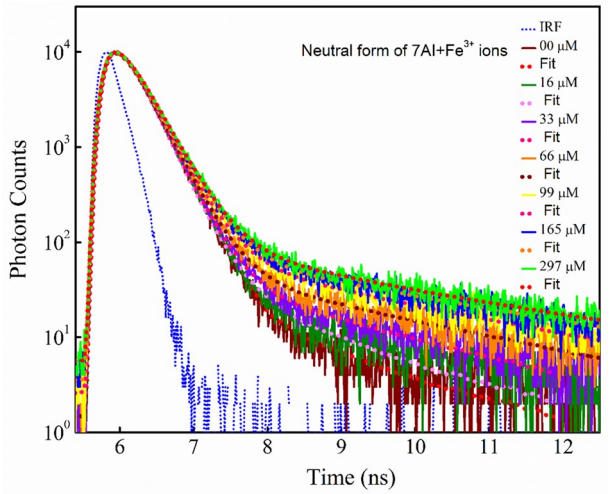


Fig. 8 Fitted fluorescence decay curves of the neutral form of 7AI in the absence and presence of different Fe^{3+} ions concentrations in the aqueous medium. The excitation wavelength was 300 nm and the emission wavelength was 390 nm. The dotted blue line shows the instrument response function (IRF)

$$K_{SV} = K_q \tau_0 \tag{6}$$

here, τ_0 is the FL lifetime in the absence of a quencher taken from Tables 1 and 2 and the value of S–V constant (K_{SV}) is used as discussed in the preceding sections. Thus, the value of K_q exhibits that the probe and quencher come closer at a fast rate, decreasing the FL intensity.

Usually, it has been assumed that if the lifetime of the probe molecule (7AI) does not alter with the reduction of FL intensity by the quencher, accordingly, the K_{sv} from equation $\tau_0/\tau = 1 + K_{sv}[Q]$ will be zero, and will represent only the static quenching in which the produced complex is non-fluorescent. A collision between the quencher and excited state molecule occurs within the probe lifetime in the case of dynamic quenching. When both processes occur, the S–V plots deviate towards positive from linearly and become quadratic with quencher concentrations. Thus, the lifetime of the protonated form of 7AI is nearly constant, which indicates the formation of a non-fluorescent complex. The B–H curves and Job plots exhibit the 1:1 complexation between 7AI and Fe^{3+} ions. Thus, the main lifetime component and the average lifetime of the protonated form of the 7AI confirm the quenching mechanism as static quenching. For a neutral form of 7AI, the change in absorption and formation of ground state complex with Fe^{3+} ions again support the static quenching as a predominant mechanism. However, the shift in FL band maximum and increase of the main lifetime component (τ_1) and corresponding pre-exponential factor with Fe^{3+} ions, as the value of τ_1 approach to the lifetime of the protonated form

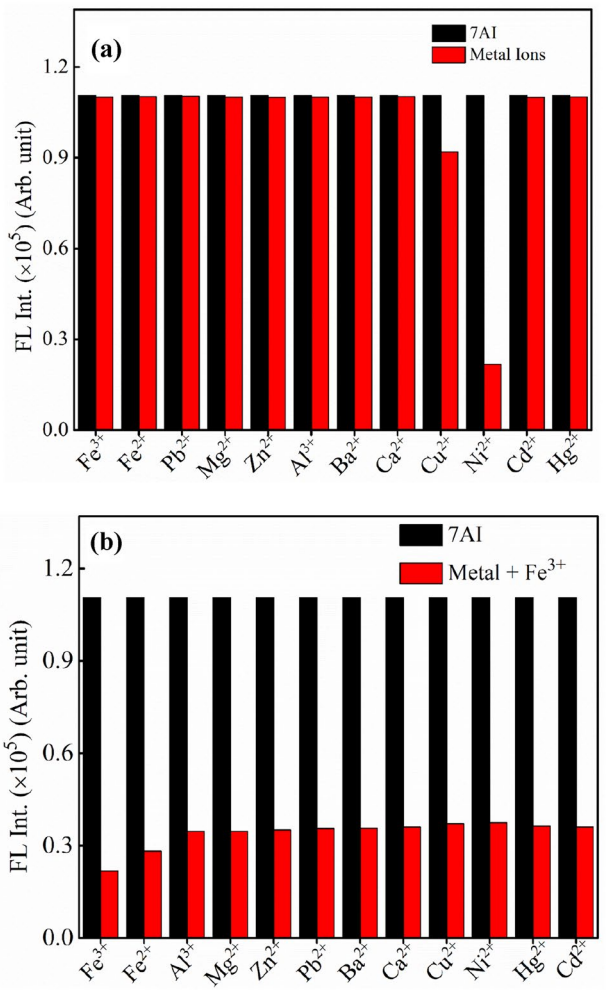


Fig. 9 The bar diagram of the fluorescence intensity of protonated form of 7AI (50 μM). The red bar shows the change in FL intensity in the presence (99 μM) of various metal ions (**a**) and different metal ions along with Fe^{3+} ions (**b**), whereas the black bars of *a* and *b* show the FL intensity in the absence of metal ions. There is nearly no interference effect of other metals while sensing Fe^{3+} ions (**b**)

of 7AI demonstrated the protonation/deprotonation along with static quenching. Usually, upon photoexcitation, the partially filled *d* orbital of metal ion readily receives electrons from the delocalized electron system of the probe in the complex, assisting electron transfer from the probe to the metal ions [34], which is commonly known as photoinduced electron transfer (PET). This process can be considered if there is no change in the absorption and no shift in the FL band maximum with Fe^{3+} ions concentration.

The effect of other metal ions (Al^{3+}, Fe^{2+}, Pb^{2+}, Ba^{2+}, Ni^{2+}, Zn^{2+}, Mg^{2+}, Ca^{2+}, Cu^{2+}, Hg^{2+} and Cd^{2+}) were also examined to find the selectivity of the probe 7AI. Interestingly, none of the metal ions show significant quenching in FL intensity except slightly Fe^{2+} ions, which reduced the overall FL intensity maximum by about < 20%. The

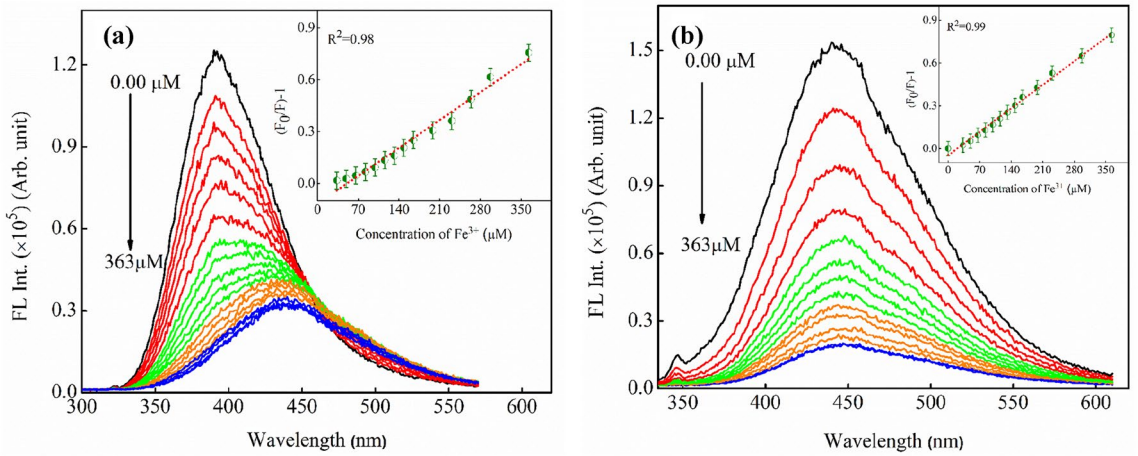


Fig. 10 Fluorescence spectra of the neutral (**a**) and protonated (**b**) forms of 7AI in tap water in the absence and presence of different concentrations of Fe^{3+} ions. Inset shows S–V plots between $(F/F_0) - 1$ vs. Fe^{3+} ions concentration of neutral and protonated forms of 7AI in tap water. The excitation wavelength was 320 nm

results are demonstrated with bar graphs in Fig. 9a and Fig. S8 of SI for the protonated and neutral forms of 7AI, respectively. To examine the interference effect of various other metal ions in the presence of Fe^{3+} ions, the FL of 7AI was also recorded with a certain amount of all the metal ions and in the existence of the same amount of the Fe^{3+} ions. The quenching of FL intensity is nearly the same as the Fe^{3+} ions, even in various other metal ions. The comparison is demonstrated with the help of the bar graph, as shown in Fig. 9b, with or without Fe^{3+} ions. Thus, it is to be mentioned that the presence of different metal ions in the aqueous environments does not show considerable interference or other metal ions could not outperform like Fe^{3+} ions.

Furthermore, the experiments were also carried out in actual samples using tap water. Similar results, as observed in ultrapure water, were obtained for neutral and protonated forms of 7AI. The FL intensity decreases monotonically in the presence of metal ions. For the neutral form of 7AI, the FL intensity falls entirely with the same amount of Fe^{3+} ions and, at the same time, shifts towards a longer wavelength region. Also, for the protonated form of 7AI, the FL intensity decreases in the same fashion as observed in ultrapure water. The S–V plots obtained from the intensity ratios are linear. The FL spectra and corresponding S–V plots (inset) for the neutral and protonated form of 7AI are shown in Fig. 10a, b. The S–V constant for the neutral and protonated form of 7AI are 2.3×10^4 M^{-1} and 2.4×10^4 M^{-1}, respectively. Accordingly, the LoDs are 1.22 ± 0.05 and 0.49 ± 0.05 μM, respectively. Thus, it has been concluded that the probe behaves in real samples like in ultrapure water. The concentration range and the order of $K_{SV,}$ as well as LoDs, are nearly the same. This reflects that the probe that shows multiparametric sensing can selectively detect ferric ions in water. Thus, the multiparametric probe undoubtedly covers a wider concentration range with high sensitivity and selectivity.

4 Conclusion

The first and systematic investigation of 7AI as a highly selective chemosensor application for metal ions based on multiple optical parameters has been proposed. The intense Fe^{3+} ions induced FL quenching of the neutral (~90%) and protonated (84%) forms of 7AI in the aqueous environment with low interference effects of other metal ions, thus acting as an efficient sensing probe. From the linearity of FL intensity with concentrations of the Fe^{3+} ions, the detection limit comes out to be 0.46 ± 0.02 and 0.49 ± 0.02 μM for the neutral and protonated form of 7AI, respectively. The linear plot of absorbance with metal ion concentrations of both the neutral and protonated form of 7AI can also act as a calibration curve with LoDs of 9.20 ± 0.05 and 9.18 ± 0.05 μM, respectively, for detecting the selective metal ions in the given range of concentrations. Thus, multiple parameters-based responses of 7AI with excellent photophysical properties, stable and fast sensing behavior, low interference, and commercial availability make the probe applicable to the optical sensing field for environmental and biological applications.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s43630-023-00393-6.

Acknowledgements The work is financially supported by the Science and Engineering Research Board (EMR/2016/001110), Department of Science and Technology (DST), Govt. of India.

Authors contributions MSM: Conception and design of the study, resources, funding acquisition, supervision, drafting, revising and editing of the manuscript. A: Acquisition and analysis of steady-state results.

Data availability All data generated or analysed during this study are included in this published article and its supplementary information file.

Declarations

Conflict of interest The authors declare that they have no conflicts of interest.

Consent to publish All authors have approved the paper and agree with its publication.

Ethical approval This article does not contain any studies with human participants or animals, clinical trial registration, or plant reproducibility performed by any authors.

References

1. Taylor, C. A., El-Bayoumi, M. A., & Kasha, M. (1969). Excited-state two-proton tautomerism in hydrogen-bonded *N*-heterocyclic base pairs. *PNAS, 63*, 253–260.
2. Douhal, A., Kim, S. K., & Zewail, A. H. (1995). Femtosecond molecular dynamics of tautomerization in model base pairs. *Nature, 378*, 260.
3. Chen, Y., Rich, R. L., Gai, F., & Petrich, J. W. (1993). Fluorescent species of 7-azaindole and 7-azatryptophan in water. *Journal of Physical Chemistry, 97*, 1770–1780. https://doi.org/10.1021/j100111a011
4. Smirnov, A. V., English, D. S., Rich, R. L., Lane, J., Teyton, L., Schwabacher, A. W., Luo, S., Thornburg, R. W., & Petrich, J. W. (1997). Photophysics and biological applications of 7-azaindole and its analogs. *The Journal of Physical Chemistry B, 101*, 2758–2769. https://doi.org/10.1021/jp9630232
5. Takeuchi, S., & Tahara, T. (1998). Femtosecond ultraviolet-visible fluorescence study of the excited-state proton-transfer reaction of 7-azaindole dimer. *Journal of Physical Chemistry A, 102*, 7740–7753. https://doi.org/10.1021/jp982522v
6. Sekiya, H., & Sakota, K. (2008). Excited-state double-proton transfer in a model DNA base pair: Resolution for stepwise and concerted mechanism controversy in the 7-azaindole dimer revealed by frequency- and time-resolved spectroscopy. *Journal of Photochemistry and Photobiology C: Photochemistry Reviews, 9*, 81–91. https://doi.org/10.1016/j.jphotochemrev.2008.04.001
7. Joshi, H. C., & Antonov, L. (2021). Excited-state intramolecular proton transfer: A short introductory review. *Molecules, 26*, 1475. https://doi.org/10.3390/molecules26051475
8. Chou, P.-T., Liu, Y.-I., Wu, G.-R., Shiao, M.-Y., Yu, W.-S., Cheng, C.-C., & Chang, C.-P. (2001). Proton-transfer tautomerism of β-carbolines mediated by hydrogen-bonded complexes. *The Journal of Physical Chemistry B, 105*, 10674–10683. https://doi.org/10.1021/jp011031z
9. Hung, F.-T., Hu, W.-P., & Chou, P.-T. (2001). The ground- and excited-state ($^1n\pi^*$ and $^1\pi\pi^*$) carboxylic acid-catalyzed proton (hydrogen atom)-transfer energy surfaces in 3-formyl-7-azaindole. *Journal of Physical Chemistry A, 105*, 10475–10482. https://doi.org/10.1021/jp012663o
10. Fujino, T., Arzhantsev, S. Y., & Tahara, T. (2001). Femtosecond time-resolved fluorescence study of photoisomerization of trans-azobenzene. *Journal of Physical Chemistry A, 105*, 8123–8129. https://doi.org/10.1021/jp0110713
11. Catalán, J., Pérez, P., del Valle, J. C., de Paz, J. L. G., & Kasha, M. (2004). H-bonded N-heterocyclic base-pair phototautomerizational potential barrier and mechanism: The 7-azaindole dimer. *PNAS, 101*, 419–422.
12. Chapman, C. F., & Maroncelli, M. (1992). Excited-state tautomerization of 7-azaindole in water. *Journal of Physical Chemistry, 96*, 8430–8441.
13. Mehata, M. S., Singh, A. K., & Sinha, R. K. (2017). Investigation of charge-separation/change in dipole moment of 7-azaindole: Quantitative measurement using solvatochromic shifts and computational approaches. *Journal of Molecular Liquids, 231*, 39–44.
14. Nakajima, A., Hirano, M., Hasumi, R., Kaya, K., Watanabe, H., Carter, C. C., Williamson, J. M., & Miller, T. A. (1997). High-resolution laser-induced fluorescence spectra of 7-azaindole-water complexes and 7-azaindole dimer. *Journal of Physical Chemistry A, 101*, 392–398. https://doi.org/10.1021/jp9614411
15. Guha Ray, J., & Sengupta, P. K. (1994). Luminescence behavior of 7-azaindole in AOT reverse micelles. *Chemical Physics Letters., 230*, 75–81. https://doi.org/10.1016/0009-2614(94)01154-0
16. Kwon, O. H., & Jang, D. J. (2005). Proton transfer of excited 7-azaindole in reverse-micellar methanol nanopools: Even faster than in bulk methanol. *The Journal of Physical Chemistry B, 109*, 8049–8052. https://doi.org/10.1021/jp050743c
17. Wu, Y.-S., Huang, H.-C., Shen, J.-Y., Tseng, H.-W., Ho, J.-W., Chen, Y.-H., & Chou, P.-T. (2015). Water-catalyzed, excited-State proton-transfer reactions in 7-azaindole and its analogues. *The Journal of Physical Chemistry B, 119*, 2302–2309.
18. Chen, K.-Y., & Lin, W.-C. (2015). A simple 7-azaindole-based ratiometric fluorescent sensor for detection of cyanide in aqueous media. *Dyes and Pigments, 123*, 1–7.
19. Kaur, K., Chaudhary, S., Singh, S., & Mehta, S. K. (2016). azaindole modified imine moiety as fluorescent probe for highly sensitive detection of Fe^{3+} ions. *Sensors and Actuators B, 232*, 396–401.
20. Mehata, M. S., Joshi, H. C., & Tripathi, H. B. (2002). Excited-state intermolecular proton transfer reaction of 6-hydroxyquinoline in protic polar medium. *Chemical Physics Letters., 359*, 314–320. https://doi.org/10.1016/S0009-2614(02)00716-9
21. Mehata, M. S., Joshi, H. C., & Tripathi, H. B. (2002). Complexation of 6-hydroxyquinoline with trimethylamine in polar and non-polar solvents. *Chemical Physics Letters, 366*, 628. https://doi.org/10.1016/S0009-2614(02)01579-8
22. Mehata, M. S. (2008). Proton translocation and electronic relaxation along a hydrogen-bonded molecular wire in a 6-hydroxyquinoline/acetic acid complex. *The Journal of Physical Chemistry B, 112*, 8383–8386. https://doi.org/10.1021/jp801811e. ibid, Photoinduced excited state proton rearrangement of 6-hydroxyquinoline along a hydrogen-bonded acetic acid wire. Chem. Phys. Lett. 436 (2007) 357–361. (doi.org/10.1016/j.cplett.2007.01.060).
23. Liu, Y. H., Mehata, M. S., & Liu, J. Y. (2011). Excited-state proton transfer via hydrogen-bonded acetic acid (AcOH) wire for 6-hydroxyquinoline. *Journal of Physical Chemistry A, 115*, 19–24. https://doi.org/10.1021/jp1101626
24. Liu, Z.-Y., Hu, J.-W., Chen, C.-L., Chen, Y.-A., Chen, K.-Y., & Chou, P.-T. (2018). Correlation among hydrogen bond, excited-state intramolecular proton-transfer kinetics and thermodynamics for −OH type proton-donor molecules. *Journal of Physical Chemistry C, 122*, 21833–21840. https://doi.org/10.1021/acs.jpcc.8b07433
25. Jacquemina, D., Khelladib, M., De Nicolab, A., & Ulrichb, G. (2019). Turning ESIPT-Based triazine fluorophores into dual

emitters: From theory to experiment. *Dyes and Pigments, 163*, 475–482. https://doi.org/10.1016/j.dyepig.2018.12.023
26. Liu, X., Liu, X., Shen, Y., & Gu, B. (2020). A simple water-soluble ESIPT fluorescent probe for fluoride ion with large Stokes shift in living cells. *ACS Omega, 5*, 21684–21688. https://doi.org/10.1021/acsomega.0c02589
27. Liu, Z.-Y., Hu, J.-W., Haung, T.-H., Chen, K.-Y., & Chou, P.-T. (2020). Excited-state intramolecular proton transfer in the kinetic control regime. *Physical Chemistry Chemical Physics: PCCP, 22*, 22271–22278. https://doi.org/10.1039/D0CP03408H
28. Wang, C.-H., Liu, Z.-Y., Huang, C.-H., Chen, C.-T., Meng, F.-Y., Liao, Y.-C., Liu, Y.-H., Chang, C.-C., Li, E. Y., & Chou, P.-T. (2021). Chapter open for the excited-state intramolecular thiol proton transfer in the room-temperature solution. *Journal of the American Chemical Society, 143*, 12715–12724. https://doi.org/10.1021/jacs.1c05602
29. Yiannikourides, A., & Dada, G. O. L. (2019). A short review of iron metabolism and pathophysiology of iron disorders. *Medicines, 6*, 85. https://doi.org/10.3390/medicines6030085
30. Simcox, J. A., & McClain, D. A. (2013). Iron and diabetes risk. *Cell Metabolism, 17*, 329–341. https://doi.org/10.1016/j.cmet.2013.02.007
31. Torti, S. V., & Torti, F. M. (2013). Iron and cancer: More ore to be mined. *Nature Reviews Cancer, 13*(2013), 342–355. https://doi.org/10.1038/nrc3495
32. Mehata, M. S. (2021). An efficient excited-state proton transfer fluorescence quenching based probe (7-hydroxyquinoline) for sensing trivalent cations in aqueous environment. *Journal of Molecular Liquids, 326*, 115379. https://doi.org/10.1016/j.molliq.2021.115379
33. Mehata, M. S., Joshi, H. C., & Tripathi, H. B. (2002). Fluorescence quenching of 6-methoxyquinoline: An indicator for sensing chloride ion in aqueous media. *Journal of Luminescence, 99*, 47–52. https://doi.org/10.1016/S0022-2313(02)00294-6
34. Desai, N. K., Kolekar, G. B., & Patil, S. R. (2014). Off–on fluorescent polyanthracene for recognition of ferric and fluoride ions in aqueous acidic media: Application in pharmaceutical and environmental analysis. *New Journal of Chemistry, 38*, 4394. https://doi.org/10.1039/c4nj00675e
35. Edison, T. N. J. I., Atchudan, R., Shim, J.-J., Kalimuthu, S., Ahn, B.-C., & Lee, Y. R. (2016). Turn-off fluorescence sensor for the detection of ferric ion in water using green synthesized N-doped carbon dots and its bio-imaging. *Journal of Photochemistry and Photobiology B: Biology, 158*, 235–242. https://doi.org/10.1016/j.jphotobiol.2016.03.010
36. Pandey, N., Mehata, M. S., Fatma, N., & Pant, S. (2020). Modulation of fluorescence properties of 5-aminoquinoline by Ag^{+} in aqueous media via charge transfer. *Journal of Photochemistry and Photobiology A: Chemistry., 396*, 112549. https://doi.org/10.1016/j.jphotochem.2020.112549
37. Fatma, N., Mehata, M. S., Pandey, N., & Pant, S. (2020). Flavones fluorescence-based dual response chemosensor for metal ions in aqueous media and fluorescence recovery. *Journal of Fluorescence, 30*, 759–772. https://doi.org/10.1007/s10895-020-02540-z
38. Lashgari, N., Badiei, A., & Ziarani, G. M. (2016). A Fluorescent sensor for Al(III) and colorimetric sensor for Fe(III) and Fe(II) based on a novel 8-hydroxyquinoline derivative. *Journal of Fluorescence, 26*, 1885–1894. https://doi.org/10.1007/s10895-016-1883-3
39. Sharma, V., & Mehata, M. S. (2021). Synthesis of photoactivated highly fluorescent Mn^{2+}-doped ZnSe quantum dots as effective lead sensor in drinking water. *Materials Research Bulletin., 134*, 111121. https://doi.org/10.1016/j.materresbull.2020.111121
40. Sharma, P., & Mehata, M. S. (2020). Rapid sensing of lead metal ions in an aqueous medium by MoS_2 quantum dots fluorescence turn-off. *Materials Research Bulletin., 131*, 110978. https://doi.org/10.1016/j.materresbull.2020.110978
41. Sharma, P., & Mehata, M. S. (2020). Colloidal MoS_2 quantum dots based optical sensor for detection of 2,4,6-TNP explosive in an aqueous medium. *Optical Materials., 100*, 109646. https://doi.org/10.1016/j.optmat.2019.109646
42. Sharma, V., & Mehata, M. S. (2021). Rapid optical sensor for recognition of explosive 2,4,6-TNP traces in water through fluorescent ZnSe quantum dots. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 260*, 119937. https://doi.org/10.1016/j.saa.2021.119937

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Authors and Affiliations

Mohan Singh Mehata[1] · **Aneesha**[1]

✉ Mohan Singh Mehata
msmehata@gmail.com

[1] Laser Spectroscopy Laboratory, Department of Applied Physics, Delhi Technological University, Bawana Road, Delhi 110042, India

mathematics

MDPI

Article

Snail: An Eco-Friendly Autonomous Steam Mopping Robot for Cleaning and Disinfection of Floors

Sathian Pookkuttath [1], Mohan Rajesh Elara [1], Madan Mohan Rayguru [2], Zaki Saptari Saldi [3], Vinu Sivanantham [1] and Balakrishnan Ramalingam [1,*]

[1] Engineering Product Development Pillar, Singapore University of Technology and Design (SUTD), Singapore 487372, Singapore
[2] Electrical Engineering, Delhi Technological University, New Delhi 110042, India
[3] Product Design Engineering, Universitas Prasetiya Mulya, Tangerang 15339, Indonesia
* Correspondence: balakrishnan@sutd.edu.sg

Abstract: Steam mopping is an eco-friendly solution for cleaning and disinfecting floors, avoiding harsh chemicals or harmful UV lights. Currently, commercial steam mopping systems are manually operated, which is laborious and unsafe. This work presents the design methodology of a novel, eco-friendly autonomous steam mopping robot named 'Snail' for cleaning and disinfecting typical indoor floors, namely tile, carpet, concrete, vinyl, sealed wood, and rugs. Our work is mainly focused on (1) the mechanical system design, including the locomotion, steam dispenser, and mopping unit, (2) the hardware and software architecture, and (3) the motion control system's design for smooth maneuverability and mitigating friction forces. A high-gain friction estimator is proposed, whose output is exploited by the motion controller to follow the desired path and compensate for the traction-based locomotion disturbances due to extensive mop contact with a rough-textured floor, such as carpet. The proposed autonomous steam mopping robot's efficiency is validated with different floors with zig-zagging cleaning patterns, setting a minimum of 100 °C as the disinfecting steam criterion. Using the proposed motion control algorithm, we validate the smooth maneuverability, assuring the robot traverses with a given speed and follows the trajectory even on a rough carpeted floor. The cleaning efficiency is tested by removing hard stains, consuming less water, and faster wet floor drying. The robot used only one liter of water to clean a 5 m^2 floor area, and the drying time was found to be four times faster than normal water mopping. The proposed steam mopping robot design is ideal for cleaning and disinfecting hospitals and eldercare centers, where hygiene is crucial, and promoting going green, avoiding harsh chemicals, and mitigating current cleaner labor concerns, including the risk of infection.

Keywords: eco-friendly cleaning; autonomous steam mopping robot; floor disinfection; robot system architecture; motion controller design; steam modeling

MSC: 93B51

check for updates

Citation: Pookkuttath, S.; Elara, M.R.; Mohan Rayguru, M.; Saldi, Z.S.; Sivanantham, V.; Ramalingam, B. Snail: An Eco-Friendly Autonomous Steam Mopping Robot for Cleaning and Disinfection of Floors. *Mathematics* **2023**, *11*, 1086. https://doi.org/10.3390/math11051086

Academic Editors: António Lopes, Xiaowei Huang, Wenjie Ruan and Xingyu Zhao

Received: 11 January 2023
Revised: 8 February 2023
Accepted: 13 February 2023
Published: 22 February 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Two emerging concepts, green cleaning and designing for the environment, are essential today to sustaining human life and environmental quality. Viral outbreaks such as COVID-19 demand frequent cleaning and disinfection. People move away from cleaning jobs not only due to their dull and repetitive nature but also due to the risk of infection. At the same time, the high growth of building infrastructure demands more cleaning forces to keep the environment eco-friendly. Unfortunately, the availability of the cleaning workforce has become tougher, considering cost, quality, and efficiency [1]. Floors contribute to a significant part of the indoor environment and become an indirect mode of transmission for viruses or other diseases [2]. This makes cleaning tasks more challenging, especially in crowded places such as malls, restaurants, airports, hospitals, educational institutes,

eldercare centers, and recreational centers. Even today, many cleaning service industries rely on harsh chemicals for cleaning and disinfecting floors, which are costly, require more caution, and are harmful to the cleaners and the environment [3]. Alternatively, replacing chemicals with the ultraviolet light range of 100–280 nm is effective for disinfection but harmful to human eyes and skin unless the utmost care and protection are considered, as discussed in [4–8]. Hence, a safe and eco-friendly automated disinfection method is essential to clean and disinfect floors, especially in large human-shared workplaces.

Steam mopping is an effective solution for adequate disinfection, avoiding harsh chemicals and harmful UV lights. Steam mopping offers better hygiene and superior cleaning performance to conventional manual cleaning methods with detergents. As steam mopping uses only clean water, steam cleaners have become a favorite cleaning method for many people with allergies, asthma, and chemical sensitivities. When used correctly, steam can quickly kill 99.99% of germs and bacteria, making it a safe, healthy, and natural way to clean the environment. Various clinical tests have been conducted in the literature, proving the disinfection efficacy of steam and advantages over chemicals, especially in [9–17]. Based on these published works, a minimum steam temperature of 100 °C is mandated for effective disinfection. Other than disinfection efficiency, the water consumption can be reduced by 90% through using steam, as briefed in [15]. In addition, as steam is used in a superheated state, it helps to remove hard stains, and the floor will dry around 3–6 times faster [18] than usual water mopping. Hence, steam is an effective medium for floor cleaning and disinfection tasks.

Steam mopping has been rising in popularity over the past few years. Many manually operated pieces of steam mopping equipment are commercially available [19–22]. However, it is hard to ensure the steam disinfection temperature and uniform dissipation throughout the mopping pad in a manual system. Other concerns with manual steam mops are burning hazards from the pad, handling long, tedious power cables, and heated long handles. Hence, an autonomous steam mopping device is a viable solution to overcome this shortcoming. Currently, several vacuum- or water-based mopping and scrubbing autonomous floor-cleaning robots at the domestic [23,24] and industrial grades [25,26] are available in the market. Additionally, many research works have been written on developing cleaning robots to improve the performance of floor cleaning in various aspects. Prabakaran et al. presented hTetro, a reconfigurable floor-cleaning robot that maximizes the coverage area, in [27]. Based on the environment, this robot can change its morphology to seven shapes, and a performance comparison was carried out with other fixed-shape platforms. Tun et al. presented enhanced maneuverability of the hTetro robot through a four-wheel steering and driving mechanism in [28] with a kinematic controller design. Similarly, various studies developed cleaning robots for other building domains, such as staircases [29,30], windows [31], walls [32,33], door handles [34], and tables [35].

The commercially available floor-cleaning robots and state-of-the-art cleaning performance explained above were developed based on vacuum- or water-based technologies for cleaning, mopping, and scrubbing. To extend floor disinfection application, the current practices using UV- and chemical-based solutions have their risk factors for usage and are not eco-friendly. Hence, a steam mopping autonomous robot is an optimal solution and essential today to mitigate the cleaning labor issues and viral outbreaks and strive for the green cleaning concept. However, for a steam-based mopping and disinfection robot, effective steam dissipation assuring a disinfection temperature with minimum energy consumption is crucial. In addition, while increasing the mopping pad size for better area coverage, the extensive friction raised at the mopping pad–floor interface (especially rough carpeted floors) affects the robot's maneuverability and ability to follow a given cleaning pattern. As per our knowledge, no studies have been conducted considering these factors (research gaps) to develop a steam mopping autonomous robot suitable for large workplaces (i.e., more toward industrial-grade than domestic use). Hence, this work presents an autonomous steam mopping robot design and addresses the design challenges, including developing a low-power steam dissipation system, assuring a minimum steam

dissipation temperature, designing a suitable motion controller for smooth maneuverability of the robot, and overcoming the extensive friction from the mopping pad while cleaning rough floors. The main research contributions of this work are the following:

- Develop an autonomous mopping robot with steam as a cleaning payload for cleaning and disinfecting typical indoor floors as a safe and eco-friendly solution;
- Design a steam mopping unit with a low power-rated compact steam generation unit and maximum mopping pad size to reduce energy consumption and increase area coverage while keeping the disinfection parameters;
- Develop a novel motion controller design for smooth maneuverability to mitigate the friction forces, especially while mopping rough floors such as carpets;
- Validate the uniform steam dissipation, maintaining the dissipation temperature, maneuverability following a given trajectory, and cleaning efficiency by testing with a zig-zagging cleaning pattern for real-time application, especially where hygiene is critical.

The outline of this paper is as follows. Section 2 elaborates on the design methodology, including the mechanical, system architecture, and motion controller design. Section 3 presents the steam modeling and experimental set-up with the results. Finally, Section 4 concludes with a summary of the paper.

2. Design Methodology of the Autonomous Steam Mopping Robot (Snail)

The outlook of the in-house developed, eco-friendly autonomous steam mopping robot, named 'Snail', is shown in Figure 1. The three design phases involved in realizing this robot were mechanical system design, hardware and software architecture design, and motion controller design. The detail of each design phase and its integration are described below.

Figure 1. Autonomous steam mopping and disinfection robot 'Snail'.

2.1. Mechanical System Design

The exploded view and subassembly CAD model of the steam mopping robot, showing the structure and indicating the key components, is shown in Figure 2. The robot's overall size is 75 × 60 × 40 cm, and the total weight is 20 kg. The main assembly of the robot is segmented into three layers: bottom, middle, and top. Here, the bottom layer is the mopping unit assembly. The middle layer is a ruggedized skeleton made of a 3 mm aluminum base plate and extruded profiles where the wheels, motors, controllers, sensors, and steam generation units are assembled. The top layer carries a battery, inverter, LiDAR sensor, and water tank. The robot locomotion was designed with a differential drive mechanism, and it uses wheels 150 mm in diameter made of solid rubber with an aluminum die-casted rim and two supporting caster wheels. Two 24 volt brushless DC geared motors with 1.9 N-m of torque were used to power the wheels separately with a closed-loop control function. For proper navigation of the robot, an inertial measurement unit (IMU) and time of flight (TOF) sensors were fastened to the middle layer base plate of the robot. Here, the

IMU is placed at the center portion of the two-wheel axis, and the TOF sensors are placed facing toward the ground. Furthermore, the electronic components include an NVIDIA Jetson AGX Xavier onboard processor, Arduino microcontroller, and voltage regulators, which are also mounted to the middle layer. Two waterproof sonars are mounted on both sides of the robot cover toward the rear. The slit designed in the snail cover enables LiDAR coverage of 210°. Furthermore, the lids of the robot cover allow easy access to the water tank and electronics control. The outer Snail shape's protective cover was 3D printed using ABS resin, and it can be easily removed to replace the mop. Here, the overall size of the robot is bulky compared with domestically used mopping robots, mainly due to the higher-capacity battery and inverter required for a tetherless steam-generating system. In addition, the outlook shape we designed as a snail took up more space. Here, we opted for a snail outlook because the robot prefers to run slower, allowing sufficient exposure per unit area (mopping pad size) for adequate disinfection. Additionally, it temporarily leaves a water trail mark while traveling, resembling a snail. This size and outlook will help highlight the robot's presence in a large, crowded workplace, avoiding accidents and being aesthetically pleasing. The mechanical design of the cleaning payload is detailed below.

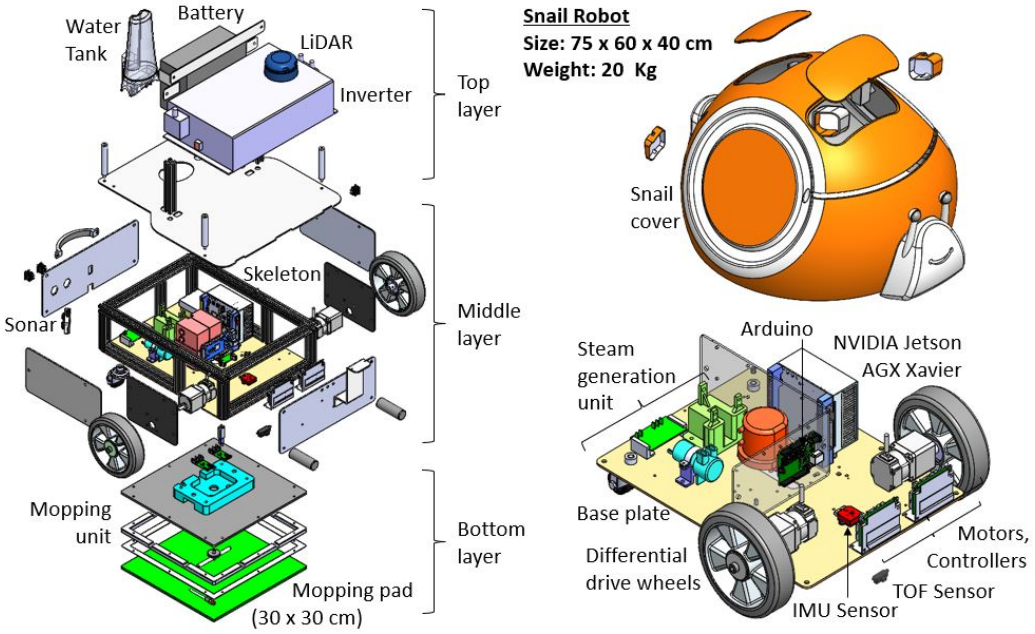


Figure 2. Snail robot's CAD mechanical design.

2.1.1. Steam Generation Unit

The steam generation unit is designed to provide continuous steam of 100 °C at the mop–floor interface for effective disinfection. A lower power of 1300 W and two bar absolute pressure-rated compact boiler were selected to balance energy consumption and meet the disinfection temperature requirements. The boiler system also includes a thermostat, thermal fuse, and an insulated casing for safety. A 0.5 L water tank, 30 mL/min discharge-rated peristaltic water pump, and a non-return valve were used to regulate the water supply using silicon rubber tubes. The boiler and pump operations, including the water level, are controlled by an Arduino mega microcontroller. The water level sensor will trigger an alarm once the water level falls below 20% of the tank capacity. A 24 V, 40 Ah Li-ion rechargeable battery is used to power the heater unit, which is capable of discharging 1300 W of power using a 2000 W inverter. The same battery was used for powering the rest of the other systems. The steam generation and distribution process flow diagram is shown in Figure 3.

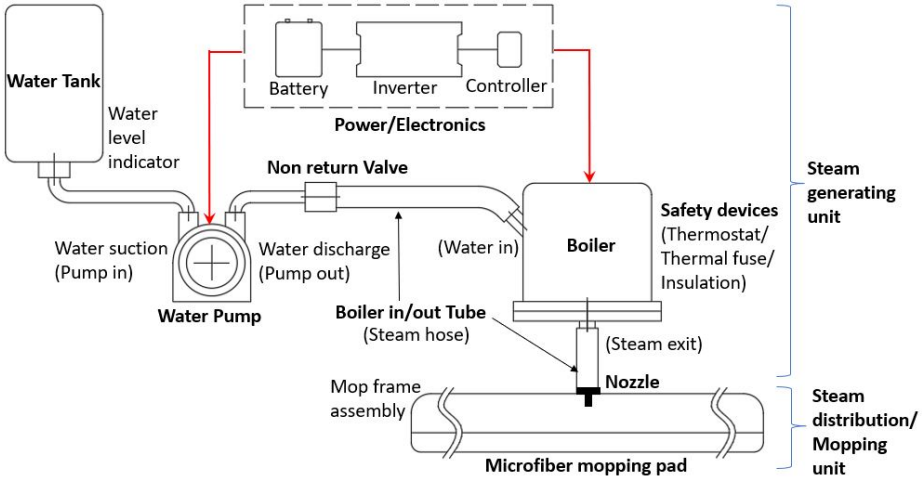


Figure 3. Steam generation and distribution process flow diagram.

2.1.2. Steam Distribution Unit

A microfiber material is used for the mopping pad, with a designed size of 30 × 30 cm. The microfiber material has high heat retention [36,37] and durability. The mopping pad was connected with the boiler unit using a synthetic braid flexible butyl rubber hose capable of holding around 6 bar of steam pressure and a temperature of around 165 °C. Furthermore, a twin-nozzle module was designed with two separate arms and a heat-resistant nylon material nozzle bracket for effective steam dissipation. The two arms of the nozzle are joined with a nozzle bracket sandwiched between a mop bracket and a mop plate. The mop bracket, plate, and frame were fabricated using heat-resistant polypropylene material. The proposed twin-nozzle design is shown in Figure 4.

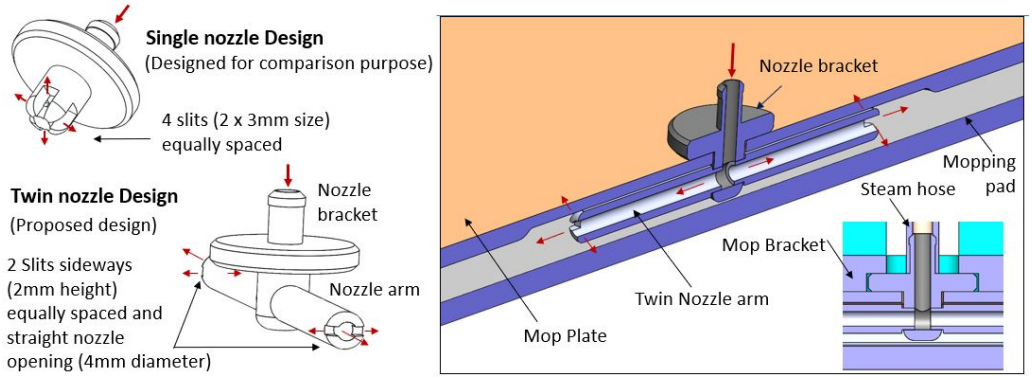


Figure 4. Nozzle designs and proposed twin-nozzle assembly.

The mopping unit is equipped with a steam temperature-sensing system to assure a uniform disinfection temperature of 100 °C at the mop. Two thermocouple sensors, sensing a range of 600 °C, were used to monitor the steam temperature. The sensors are housed in the mop bracket, and the probes are mounted at the two diagonal corners of the mop plate toward the mop. This helps with monitoring the target steam dissipation temperature across the mop. The sensors are connected to an Arduino UNO microcontroller. Furthermore, an auto-adjustable mop–floor contact-sensing system is integrated to maintain the mop's pressure with various floors, uneven surfaces, or any tiny floor material. Figure 5 depicts these two systems. In addition, a TOF sensor (range: 4–30 cm), mounted at the base plate and facing the floor, reads and stores the distance to the particular floor once the initial mop–floor contact set-up is complete. Furthermore, suppose there is any change in this initial TOF sensor reading due to an uneven surface or ground material. In this case, a micro linear actuator with feedback will trigger to auto-adjust the mop level. This linear actuator (22 N and 30 mm stroke) is mounted with a base plate and mopping unit. A

roboclaw motor controller is used to update the actuator motion with the Arduino UNO microcontroller. The whole mechanism helps to adjust the mopping unit height up to 9 mm, aided with guide bushes and pins.

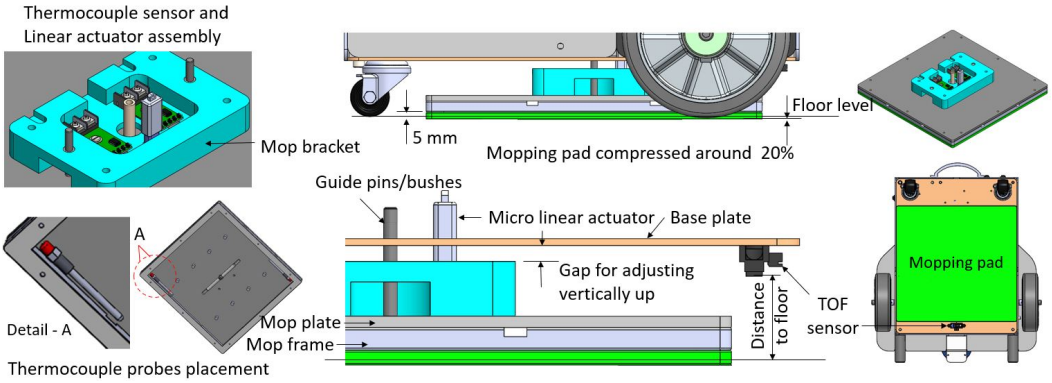


Figure 5. Mop temperature-monitoring and uniform floor contact system.

2.2. System Architecture Design

The system architecture, mainly the hardware and software architecture of the steam mopping robot, is illustrated in Figure 6 and explained as follows.

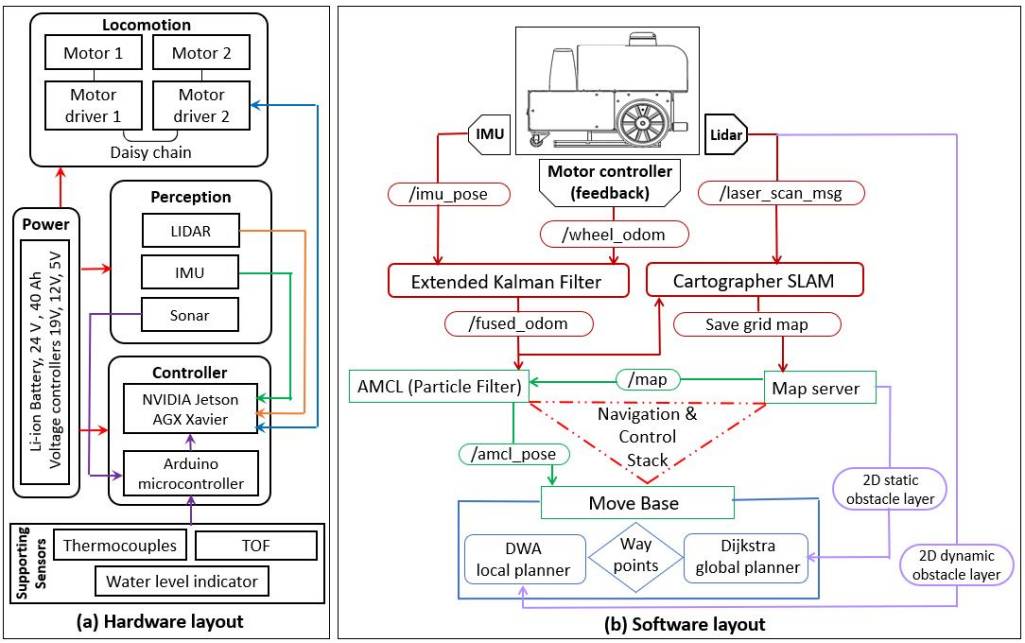


Figure 6. System architecture.

2.2.1. Hardware Architecture

The hardware architecture of the steam mopping robot has two processing units: an Nvidia Jetson AGX Xavier single-board computer (SBC) and Arduino UNO embedded microcontroller. The Nvidia Jetson AGX Xavier SBC is used as the central processing unit. It handles navigation- and perception-related tasks. RPLIDAR A2, Vectornav IMU VN-100, and SONAR sensors are used for mapping and perception. The Arduino UNO microcontroller acts as an embedded microcontroller that handles lower-level control tasks. It monitors and controls the steam mopping unit and manages the steam temperature's fixed threshold level. A water level indicator, thermocouples, and a TOF sensor are connected to the Arduino microcontroller, which are used to monitor the mopping unit's operation and control.

2.2.2. Software Architecture

The Robot Operating System (ROS) melodic framework was used to build the software modules and also act as the middleware to support and establish communication between other hardware units and software packages. The navigation stack and control system stack of the robot were build on top of ROS. The higher-level systems such as the 2D LIDAR, IMU, and motor controllers are directly connected through SBC using ROS nodes, which publish the sensor data in their corresponding ROS topics to support the navigation stack and control stack. The ROS Navigation Stack was adopted to implement the autonomous navigation, which subscribes the motor controller feedback, IMU feedback, and fusion with an extended Kalman fiter (EKF) to estimate the odometry data and uses an adaptive Monte Carlo localization (AMCL) package for the robot's localization. A move base package plans the path (global planner) and avoids the obstacles present on the map, and a local planner controls the robot to avoid dynamic obstacles by aligning the robot's motion with the generated global path. The Cartographer SLAM algorithm was used to generate a 2D occupancy grid map of the environment and for autonomous navigation. The control system stack manages the navigation goal and executes the trajectory algorithm to continuously monitor the state of the robot from the high-level sensor and lower-level sensors connected to the robot. Based on the feedback from the sensors, the control system stack fine-tunes the robot trajectory.

2.3. Control System Design

One challenge in the locomotion of this steam mopping robot is the extensive friction from the pad–floor interface due to the 30 × 30 cm large mopping pad size designed for maximum coverage, especially while mopping rough-textured carpets. Opting for a higher-power torque-rated motor or friction modeling with necessary sensors will add to the cost. Therefore, a trajectory-tracking algorithm is proposed, which consists of a high-gain friction estimator in tandem with a PD controller.

The differential drive mechanism adopted for the Snail robot is illustrated in Figure 7, with the inverse and forward kinematic equations used for the robot's maneuverability followed by an illustration of the motion controller's design.

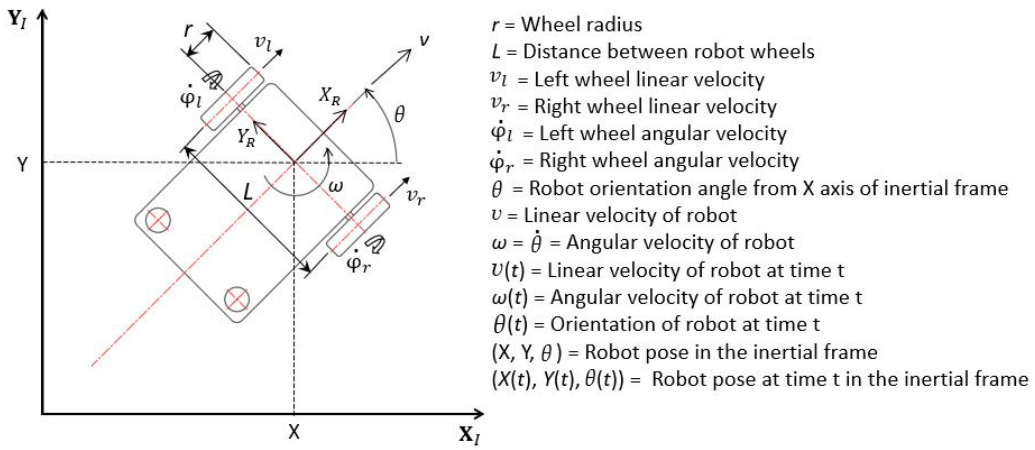


Figure 7. Differential drive mechanism.

With inverse kinematics, the left and right wheel tangential velocities can be found (Equation (1)) by applying the target linear and rotational velocities of the robot:

$$v_l = \frac{(2v - \omega L)}{2}$$
$$v_r = \frac{(2v + \omega L)}{2} \quad (1)$$

The control inputs the linear and angular velocities (v and ω, respectively), which are found as shown in Equation (2):

$$\begin{aligned} v &= \frac{r}{2}(\dot{\varphi}_r + \dot{\varphi}_l) \\ \omega &= \frac{r}{L}(\dot{\varphi}_r - \dot{\varphi}_l) \end{aligned} \tag{2}$$

As the pose (X, Y, and θ) functions depend on the functions of both the linear and angular velocities at time t, the new pose of the robot in the inertial frame is represented by

$$\begin{aligned} X(t) &= \int_0^t v(t)\cos[\theta(t)]dt \\ Y(t) &= \int_0^t v(t)\sin[\theta(t)]dt \\ \theta(t) &= \int_0^t \omega(t)dt \end{aligned} \tag{3}$$

The forward kinematics of the robot can be derived by expressing $X(t), Y(t)$, and $\theta(t)$ from Equation (3) in differential form (position vector):

$$\dot{p} = \begin{bmatrix} \dot{X} \\ \dot{Y} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} cos(\theta) & 0 \\ sin(\theta) & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v \\ \omega \end{bmatrix} \tag{4}$$

The mopping robot has to travel some predetermined path for satisfactory coverage. Hence, the motion control objective is to choose a suitable pair (v, ω) such that a desired path $(x_d(t), y_d(t))$ can be followed, starting from an arbitrary position $(X(0), Y(0))$. Given a desired path, the orientation $\theta_d(t)$ can be derived from $x_d(t)$ and $y_d(t)$. Hence, the trajectory tracking task can be fulfilled by enforcing $x \to x_d$ and $y \to y_d$. The kinematics involving only X and Y can be written as

$$\begin{bmatrix} \dot{X} \\ \dot{Y} \end{bmatrix} = \begin{bmatrix} cos(\theta) & 0 \\ sin(\theta) & 0 \end{bmatrix} \begin{bmatrix} v \\ \omega \end{bmatrix} \tag{5}$$

The problem with the model in Equation (5) is that we cannot uniquely determine (v, ω) from a particular $(\dot{X}, \dot{Y})$ pair due to the singularity in the system matrix (uncontrollable dynamics). To remove the singularity in Equation (5), we can differentiate the L.H.S. of Equation (5) as shown in Equation (6):

$$\begin{aligned} \ddot{X} &= \dot{v}cos(\theta) - vsin(\theta)\omega \\ \ddot{Y} &= \dot{v}sin(\theta) + vcos(\theta)\omega \end{aligned} \tag{6}$$

After rearranging, the new motion model can be expressed as

$$\begin{bmatrix} \ddot{X} \\ \ddot{Y} \end{bmatrix} = \begin{bmatrix} cos(\theta) & -\zeta sin(\theta) \\ sin(\theta) & \zeta cos(\theta) \end{bmatrix} \begin{bmatrix} f \\ \omega \end{bmatrix} = \Theta U \tag{7}$$

where $U = [f, \omega] \in R^2$ is the transformed control input pair and Θ is the new input matrix. Note that the linear velocity v and angular velocity ω are functions of the wheel velocities v_l and v_r, respectively (check Equation (1)). Thus, Equation (7) can also be expressed in terms of the wheel velocities. The control input f is the linear acceleration, which is realized by

$$\dot{\zeta} = f, \;\; v = \zeta \tag{8}$$

The compensator in Equation (8) is nothing but an integrator defined in the time domain. Note that the issue of singularity is resolved in Equation (7) with the new control pair. For an inversion-based control law, we have

$$U = \Theta^{-1}(\xi) \tag{9}$$

The system in Equation (7) is converted into a decoupled linear form:

$$\begin{bmatrix} \ddot{X} \\ \ddot{Y} \end{bmatrix} = \xi \tag{10}$$

where $\xi = [\xi_1, \xi_2]^T$ can be chosen as the simple PD feedback to solve the tracking problem for Equation (7).

However, the above control law does not consider any uncertainties in the acceleration of the robot described through Equation (7). The presence of a large mop may add significant friction (especially with carpet) to any acceleration in the X and Y directions. Therefore, a more practical model of the robot can be written as

$$\begin{bmatrix} \ddot{X} \\ \ddot{Y} \end{bmatrix} = \begin{bmatrix} \xi_1 + d_x \\ \xi_2 + d_y \end{bmatrix} \tag{11}$$

where d_x and d_y are the lumped uncertainties which may arise due to the variation in the friction forces from the mopping. In this regard, it is not conservative to assume that the variation in the friction forces with respect to time is slow. This leads to the following assumption:

Assumption 1.

$$||\dot{d}_x|| \approx 0, \;\; ||\dot{d}_y|| \approx 0. \tag{12}$$

To perfectly track a desired path, the impact of the uncertainties must be mitigated. Therefore, a fast disturbance estimator is proposed in the form of

$$\begin{aligned} \mu \dot{d}_1 &= -(d_1 + \frac{1}{\mu}\dot{X} + \xi_1) \\ \mu \dot{d}_2 &= -(d_2 + \frac{1}{\mu}\dot{Y} + \xi_2) \\ \hat{d}_x &= d_1 + \frac{1}{\mu}\dot{X} \\ \hat{d}_y &= d_2 + \frac{1}{\mu}\dot{Y} \end{aligned} \tag{13}$$

where $\mu \in (0, 1)$ is called the perturbation parameter, (d_1, d_2) are intermediate variables, and $(\hat{d}_x, \hat{d}_y)$ are estimates of the uncertainties. Note that Equation (13) is the dynamics for the proposed disturbance observer. The outputs of the observer are the estimates of disturbances d_x and d_y. The chosen parameter μ should be sufficiently small, depending on the sampling constraints. It can be derived that

$$\begin{aligned} \mu \dot{\hat{d}}_x &= d_x - \hat{d}_x \\ \mu \dot{\hat{d}}_y &= d_y - \hat{d}_y. \end{aligned} \tag{14}$$

The output of the estimator is fed to the controller, which is proposed in the form of

$$\begin{aligned} \xi_1 &= \ddot{x}_d - k_1\dot{e}_x - k_3 e_x \\ \xi_2 &= \ddot{y}_d - k_2\dot{e}_x - k_4 e_y \\ \begin{bmatrix} f \\ \omega \end{bmatrix} &= \hat{\Theta}^{-1}\begin{bmatrix} \xi_1 - \hat{d}_x \\ \xi_2 - \hat{d}_y \end{bmatrix} \\ v &= \zeta \end{aligned} \tag{15}$$

where $k_i,\ i = 1, 2$ are positive scalars acting as controller gains and the error variables are $e_x = X - x_d$ and $e_y = Y - y_d$.

With the control law (Equation (15)), the closed loop dynamics become

$$\begin{aligned} \ddot{e}_x &= -k_1\dot{e}_x - k_2 e_x + d_x - \hat{d}_x \\ \ddot{e}_y &= -k_3\dot{e}_y - k_2 e_y + d_y - \hat{d}_y \\ \mu\dot{\hat{d}}_x &= d_x - \hat{d}_x \\ \mu\dot{\hat{d}}_y &= d_y - \hat{d}_y \end{aligned} \tag{16}$$

which is in the form of a standard singularly perturbed model, with the error dynamics as a slow subsystem and the estimator dynamics as a fast subsystem.

Using the singularly perturbed technique and the smallness of μ, it can be derived that the fast estimator dynamics have a unique isolated root in the form of

$$\begin{bmatrix} d_x \\ d_y \end{bmatrix} = \begin{bmatrix} \hat{d}_x \\ \hat{d}_y. \end{bmatrix} \tag{17}$$

In other words, for a small enough μ, the estimated uncertainties will be equal to the actual uncertainties. If this root is substituted into the slow tracking dynamics, then the terms $(d_x - \hat{d}_x)$ and $(d_y - \hat{d}_y)$ will be zero. Hence, the tracking error dynamics become

$$\begin{aligned} \ddot{e}_x &= -k_1\dot{e}_x - k_2 e_x \\ \ddot{e}_y &= -k_3\dot{e}_y - k_2 e_y. \end{aligned} \tag{18}$$

The dynamics in Equation (18) are a stable Hurwitz system, and thus the tracking errors converge to zero such that

$$e_x \ \rightarrow \ 0, \qquad e_y \ \rightarrow \ 0. \tag{19}$$

3. Experiments, Simulations, and Discussion

The experimental section mainly validates the uniform steam dissipation, attaining the disinfection temperature, motion controller design mitigating the friction to follow the given trajectory, and cleaning efficiency as follows.

3.1. Evaluation of the Steam Temperature and Dissipation

The evaluation of the uniformity of steam dissipation at the disinfection temperature was conducted in two phases, including a steam flow simulation test and real-time steam flow dissipation test with a mopping pad.

3.1.1. Steam Flow Simulation

A steam flow simulation was conducted to evaluate the uniformity of steam dissipation in the cavity above the mop. The simulation was based on the transient Euler–Euler multiphase mixture model on ANSYS Fluent 2020 R2 to treat the transport of two interpenetrating continuum fluids, namely air and steam. Turbulence was accounted for using the k–ϵ Reynolds-averaged Navier–Stokes (RANS) model.

Superheated steam or saturated dry steam with a dryness fraction value X = 1 (no water droplets) is the ideal condition for the best steam performance for effectively removing dirt and disinfection. The dryness fraction (X) denotes the steam quality, and as per the standard definition, it is the ratio of the mass of dry steam to the combined mass of dry steam and mass of liquid in the mixture. However, due to the turbulence and splashing, environmental conditions, nozzle design and material etc., water molecules condense quickly, and hence steam often has water droplets. Thus, the steam's dryness fraction may vary from 0.90 to 0.95. Therefore, we tested these two dryness fraction conditions to validate our steam generation and dissipation design, assuring that the disinfection criterion (minimum 100 °C) was attained across the whole mopping pad area.

A quarter of the cavity was taken as the computational domain due to the symmetry in both lateral directions. A velocity magnitude equivalent with a steam flow rate of 2.2 kg/h, a steam volume fraction of 1.0, pressure of 2 bar, and temperature of 120 °C was imposed at the nozzle, as per the boiler selected for the Snail robot. Furthermore, a pressure outlet was imposed at the two small lateral outlets. The side boundaries were treated as walls with a convective heat transfer coefficient of 10 W/m^2-K and ambient temperature of 25 °C at the outer side. The top and bottom boundaries were assumed to be thermally insulated. The simulations were performed for four study cases (i.e., combination of steam qualities ($X = 0.9$ and $X = 0.95$) and nozzle designs (single and twin)). The thermal properties (density, specific heat capacity, thermal conductivity, enthalpy, and dynamic viscosity) were calculated based on the steam quality. The physical simulation time was set to 120 s.

The steam properties were calculated based on the saturated water and saturated vapor properties at a pressure of 2 bar and the steam quality via the following expression:

$$\mathrm{SP} = \mathrm{SP}_g X + (1 - X)\mathrm{SP}_l \tag{20}$$

where SP denotes the steam properties, SP_l represents the properties of saturated water, SP_g represents the properties of saturated vapor, and X is the steam quality.

Figure 8 shows the steam dissipation at different time instants for the two steam qualities and the two nozzle designs. For each plot, the top left area corresponds with the corner. The bottom right area corresponds with the nozzle's location in the single nozzle case. In the twin nozzle case, the nozzle inlet was positioned at almost half the distance from the bottom right to the top right corner. With the twin nozzle set-up, a temperature of around 100 °C was reached at the corner after 120 s for both steam qualities, which was not the case for the single nozzle.

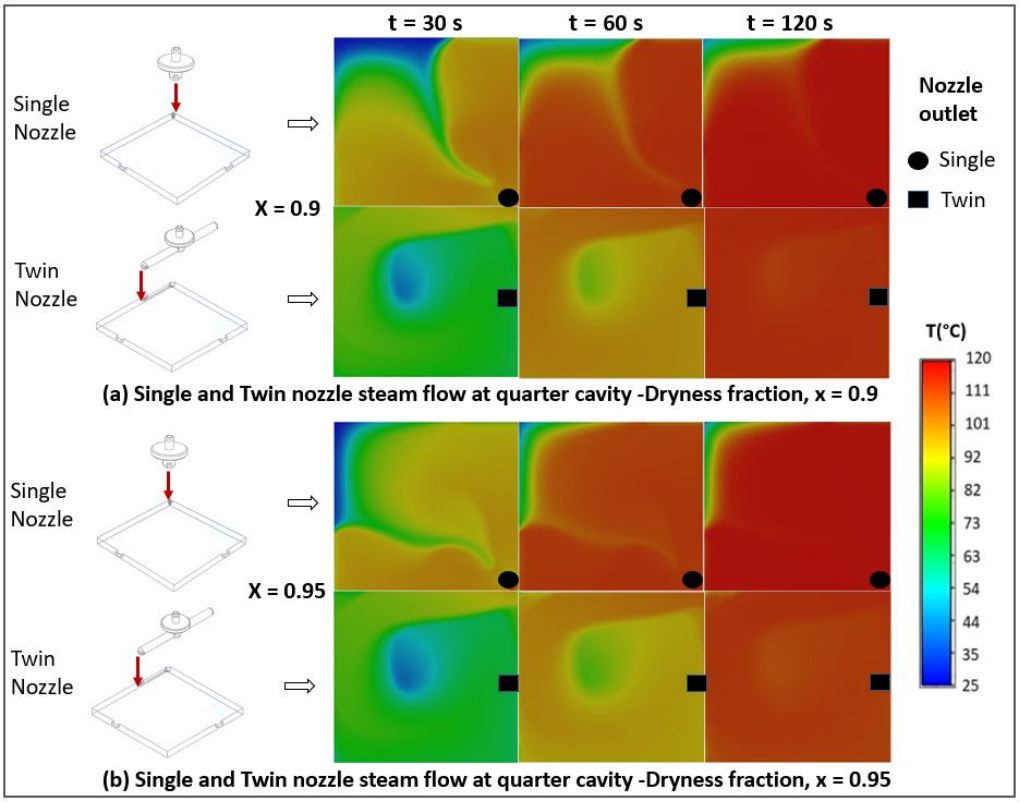


Figure 8. Top view of simulated temperature distribution at the mop–cavity interface.

The high corner temperature in the twin nozzle case was allowed by the steam jet impingement on the wall, owing to the nozzle position. The impinged heat-carrying jet was then deflected toward the corner, leading to a rotating convective flow. With the single nozzle, the steam tended to flow more radially without rotation around the corner. The rotating flow in the twin nozzle case enhanced the heat transfer throughout the cavity, as shown by the higher rate of temperature increase. This was also evident through the more uniform high temperature distribution at 120 s. The twin nozzle set-up also worked equally effectively for the two steam qualities, hence being more favorable than the single nozzle set-up.

3.1.2. Real-Time Steam Flow Test

A steam dissipation real-time test was conducted to assure uniform steam dissipation across the mopping pad–floor interface meeting the disinfection temperature. Here, the onboard thermocouple sensors monitor the steam temperature at the inner side of the moping pad. Hence, two additional thermocouple sensors were mounted diagonally on the floor (matching the pad's corners) to measure the steam temperature at the pad–floor interface in the extreme corners. The test was conducted on tile, carpet, concrete, vinyl, wood, and rug floors at room temperature (around 25 °C). Figure 9 shows the test set-up for the carpet floor. Accordingly, on the carpet, rug, and wooden floors, the steam temperature reached 100 °C within 115–125 s, whereas for the vinyl, tile, and concrete floors, this took 125–135 s. The onboard sensors read the maximum temperature of 105 °C at the this longest time period of 135 s. Hence, the target steam temperature for the onboard sensors was fixed at 105 °C for both the initial set-up and while mopping.

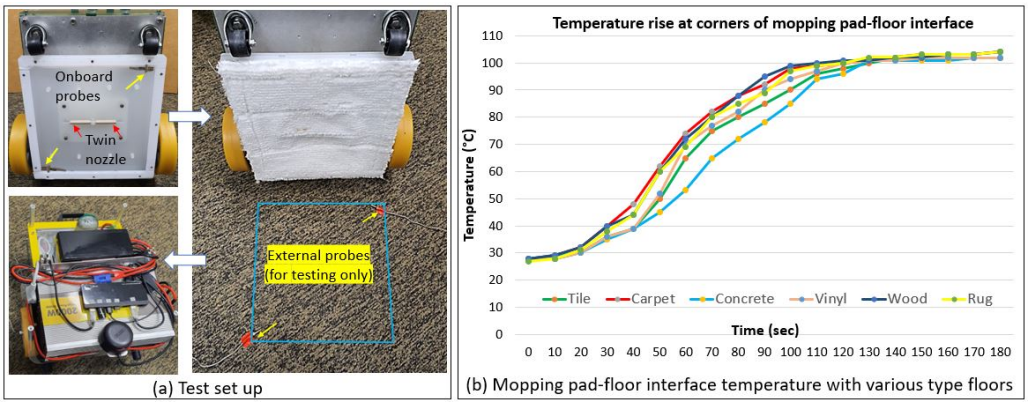


Figure 9. Rise in temperature at corners of pad–floor interface for various floors.

3.2. Motion Controller Performance Analysis and Comparison

To validate the motion controller performance, traverse speed feedback, and how closely the robot followed the given trajectory, tests were conducted before and after motion controller execution. The tests were conducted on tile, carpet, concrete, vinyl, and wooden floors at different linear and angular speeds (0.02–0.4 m/s and 0.2–1.0 rad/s, respectively). The tests were conducted at room temperature (25 °C) while assuring that the onboard thermocouple sensors and linear actuator mechanism were working correctly to maintain the steam disinfection temperature and constant mop contact with the floor before starting each test. In the robot's traverse feedback test, the actual velocities of the robot for each motion were recorded from the motor controller feedback and compared with the given velocity at every second. Accordingly, without a motion controller, there was obvious lagging, especially while driving through rough carpets, due to friction and wheel slippage. Figure 10 shows the traverse feedback test conducted at a linear speed of 0.35 m/s for 15 s before and after implementing the motion controller design. For the carpet, the actual velocity dropped to 0.335 m/s most of the time, which was improved to be very close (0.349 m/s) to the given velocity before and after the motion controller design, respectively.

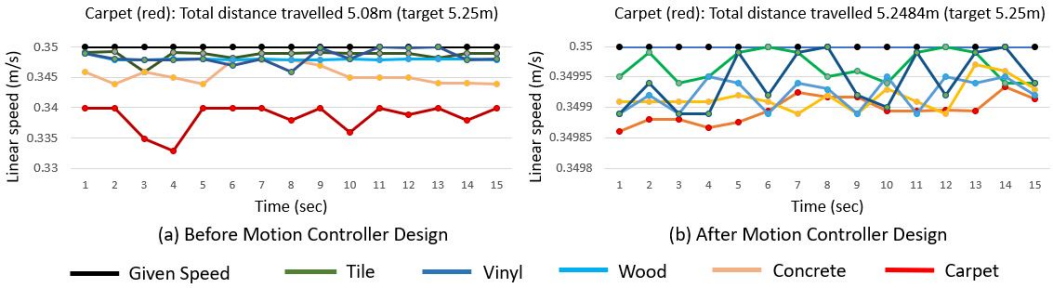


Figure 10. Snail robot's traverse feedback before and after motion controller design.

Furthermore, consistent lane tracking to avoid any deviation from the given path was tested on the same types of floors. A test bed size of a 1.2 × 4.2 m area was selected, and a 2D map of this test bed environment was generated using Cartographer SLAM to follow the planned path in a zig-zagging pattern using the waypoint coordinates of the map. Figure 11 shows this test bed's details together with the test results for a smooth concrete floor and rough carpet (the two extreme conditions) with and without executing the proposed motion controller (i.e., comparing the basic PID controller with this friction-estimated improved motion controller design). The motion controller comparison results are plotted in the same figure (B-1, B-2, C-1, and C-2), showing the performance improvement following the given trajectory with the motion controller design. The plotted trajectory was based on linear and angular speeds of 0.3 m/s and 0.2 rad/s, respectively, and we found similar changes at other speeds.

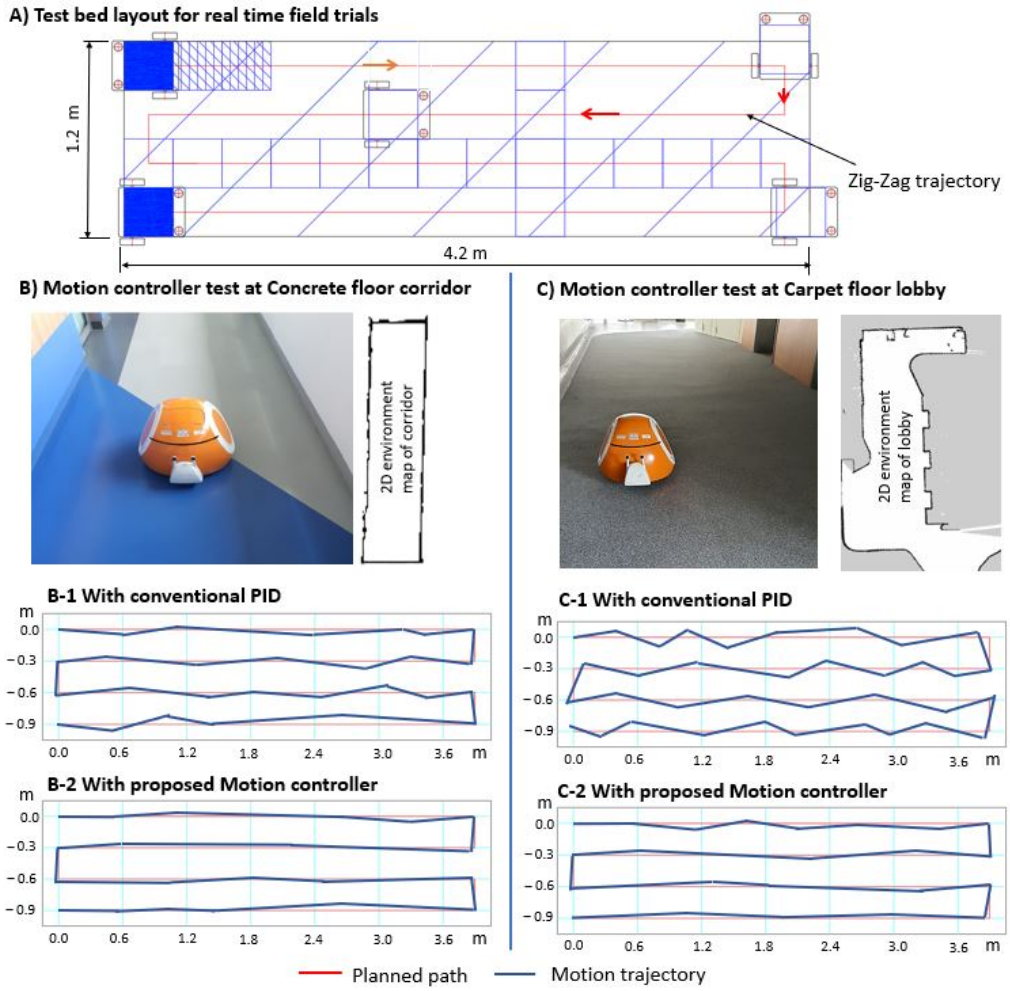


Figure 11. Autonomous traverse test following trajectory with motion controller design.

Furthermore, a quantitative comparison between this conventional PID and the proposed motion controller design was conducted based on the tracking error. As the proposed controller estimates the frictional disturbances and compensates for these uncertainties through control action, the robot can track the path better with this compared with the case where the PID is used. To quantify the improvements, the RMS values of the tracking errors were taken and compared in Table 1. The table also shows that the PID performance deteriorated, especially on carpet floors, where the friction effect was greater. However, the proposed controller performance was almost similar to the case with the concrete floor. This result proves the effectiveness of friction estimation and compensation mechanisms via the proposed control law.

Table 1. RMS Path tracking error (m) comparison.

Floor Surface	PID	Proposed
Concrete	0.021	0.015 (+34%)
Carpet	0.029	0.016 (+76%)

Note: RMS error = $\sqrt{(\frac{1}{n}(\sum_{ns} E_{pi}^2))}$, where E_{pi} is the magnitude of the path tracking error for the ith sample and ns is the number of samples obtained.

3.3. Cleaning Efficiency of Steam Mopping Robot

This section details the cleaning efficiency test of the steam mopping robot compared with a normal mopping robot by replacing the boiler-nozzle assembly of the Snail robot

with a normal water-spraying unit. We mainly tested removing a hard stain and the wet floor drying time on a concrete floor. We tested removing two similar hard, stubborn stains: one with steam and the other with the normal water mopping robot at the same speed. The steam mopping robot needed two iterations, while the water mopping robot needed four rounds to completely remove the stain, as shown in Figure 12. Similarly, it was observed that after the Snail robot's steam mopping, the wet surface dried within 1.5 min, whereas for the manual mopping with water, it took around six minutes for the surface to dry. Here, the superheated state of the condensed water helped to clean the hard stain and dry the floor quickly for the steam mopping robot. In addition, the steam mopping robot consumed 1 L of water to mop a 5 m^2 area, whereas normal water mopping took around 3.5 L for the tested concrete floor. This shows that the steam mopping robot assured a green and hygienic environment, avoiding chemicals and killing harmful germs and bacteria while being less expensive, avoiding big water tanks, and being more productive for the cleaning industry.

Figure 12. Cleaning efficiency of Snail when a removing hard stain.

4. Conclusions

An autonomous steam mopping robot design methodology was proposed and tested for eco-friendly cleaning and disinfection of typical indoor floors, such as tile, carpet, concrete, vinyl, sealed wood, and rugs. A low power-rated (2 bar, 1300 W) compact steam generation unit with twin nozzles and a 30 × 30 cm pad mopping unit was designed to balance lower energy consumption, attain a disinfection temperature, and increase the coverage area. The uniformity in steam dissipation across the mopping pad was validated using steam flow simulations and real-time tests, which assured the capability to maintain the disinfection temperature, set as 100 °C. A novel motion controller design was developed and validated to compensate for extensive pad–floor friction disturbances, especially for rough floors such as carpet. The experimental results show the cleaning efficiency and productivity of the proposed robot, mainly in removing the hard stain, fast floor drying after mopping, and the consumption of little water. The experimental results ensure that the proposed steam mopping robot design is ideal for the disinfection of floors, especially in hospitals and eldercare centers, where disinfection is crucial.

Author Contributions: Conceptualization, S.P. and M.R.E.; methodology, S.P., M.M.R. and B.R.; software, V.S. and S.P.; validation, S.P. and V.S.; formal analysis, S.P., M.R.E. and Z.S.S.; investigation, S.P. and M.R.E.; resources, M.R.E.; data, S.P., V.S. and Z.S.S.; writing—original draft preparation, S.P., M.M.R., Z.S.S. and B.R.; supervision, M.R.E.; project administration, M.R.E.; funding acquisition, M.R.E. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the National Robotics Programme under its Robotics Enabling Capabilities and Technologies (Funding Agency Project No. 192 25 00051) and the National Robotics Programme under its Robotics Domain Specific (Funding Agency Project No. W1922d0110), administered by the Agency for Science, Technology and Research.

Acknowledgments: The authors would like to thank the National Robotics Programme, the Agency for Science, Technology and Research, and the SUTD for their support.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Wordsworth, P.; Lee, R. *Lee's Building Maintenance Management*; Blackwell Science: Princeton, NJ, USA, 2001.
2. Ryu, B.H.; Cho, Y.; Cho, O.H.; Hong, S.I.; Kim, S.; Lee, S. Environmental contamination of SARS-CoV-2 during the COVID-19 outbreak in South Korea. *Am. J. Infect. Control* **2020**, *48*, 875–879. [CrossRef] [PubMed]
3. Wolkoff, P.; Schneider, T.; Kildesø, J.; Degerth, R.; Jaroszewski, M.; Schunk, H. Risk in cleaning: Chemical and physical exposure. *Sci. Total Environ.* **1998**, *215*, 135–156. [CrossRef] [PubMed]
4. Advanced Biotechnologies INC. Is UV Sterilisation Effective for Viruses and Bacteria? Available online: https://abionline.com/is-uv-sterilization-effective-for-viruses-and-bacteria (accessed on 14 May 2021).
5. Houser, K.W. *Ten Facts about UV Radiation and COVID-19*; Taylor & Francis: Abingdon, UK, 2020.
6. Wang, X.V.; Wang, L. A literature survey of the robotic technologies during the COVID-19 pandemic. *J. Manuf. Syst.* **2021**, *60*, 823–836. [CrossRef] [PubMed]
7. Choi, H.K.; Cui, C.; Seok, H.; Bae, J.Y.; Jeon, J.H.; Lee, G.E.; Choi, W.S.; Park, M.S.; Park, D.W. Feasibility of ultraviolet light-emitting diode irradiation robot for terminal decontamination of coronavirus disease 2019 (COVID-19) patient rooms. *Infect. Control. Hosp. Epidemiol.* **2021**, *43*, 232–237. [CrossRef] [PubMed]
8. Murphy, R.R.; Gandudi, V.B.M.; Adams, J. Applications of robots for COVID-19 response. *arXiv* **2020**, arXiv:2008.06976.
9. Song, L.; Wu, J.; Xi, C. Biofilms on environmental surfaces: Evaluation of the disinfection efficacy of a novel steam vapor system. *Am. J. Infect. Control* **2012**, *40*, 926–930. [CrossRef]
10. Morrone, G.; De Lorenzi, S.; Barrai, I.; Mosca, B.; Finzi, G.; Cugini, P. Comparing different methods of sanitization. in operating theatres: Use of disinfectant detergents or of steam? In Proceedings of the 34th National Congress Anmdo, Catania, Italy, 18 September 2008.
11. Bradbeer, S.J.; Coughlan, N.E.; Cuthbert, R.N.; Crane, K.; Dick, J.T.; Caffrey, J.M.; Lucy, F.E.; Renals, T.; Davis, E.; Warren, D.A.; et al. The effectiveness of disinfectant and steam exposure treatments to prevent the spread of the highly invasive killer shrimp, Dikerogammarus villosus. *Sci. Rep.* **2020**, *10*, 1–7. [CrossRef]
12. Bayzi Corporation. Steam Cleaning Studies. Available online: https://bayzi.com/steam-cleaning-studies (accessed on 15 February 2021).
13. Gillespie, E.; Williams, N.; Sloane, T.; Wright, L.; Kotsanas, D.; Stuart, R.L. Using microfiber and steam technology to improve cleaning outcomes in an intensive care unit. *Am. J. Infect. Control* **2015**, *43*, 177–179. [CrossRef]
14. Abernethy, M.; Gillespie, E.; Snook, K.; Stuart, R.L. Microfiber and steam for environmental cleaning during an outbreak. *Am. J. Infect. Control* **2013**, *41*, 1134–1135. [CrossRef]
15. Gillespie, E.; Wilson, J.; Lovegrove, A.; Scott, C.; Abernethy, M.; Kotsanas, D.; Stuart, R.L. Environment cleaning without chemicals in clinical settings. *Am. J. Infect. Control* **2013**, *41*, 461–463. [CrossRef]
16. Colloff, M.; Taylor, C.; Merrett, T. The use ofdomestic steam cleaning for the control of house dust mites. *Clin. Exp. Allergy* **1995**, *25*, 1061–1066. [CrossRef] [PubMed]
17. De Lorenzi, S.; Salvatorelli, G.; Finzi, G.; Cugini, P. Use of a steam generator for disinfection of hospital ward room surfaces. *Br. Microbiol. Res. J.* **2012**, *2*, 228. [CrossRef]
18. Nilfisk, Ltd. Why Steam Is the Best Solution for Cleaning When Hygiene Is Critical. Available online: https://new.nilfisk.com/global/articles/steam-solution-when-hygiene-is-critical (accessed on 10 January 2021).
19. Rosenzweig, M.; Ognjen, V. Steam Mop. US Patent 8,205,293, 26 June 2012.
20. Shaw, R.R. Steam Mop. US Patent 6,584,990, 1 July 2003.
21. kagoo.co.uk. Compare the Best Steam Cleaners. Available online: https://kagoo.co.uk/steam-cleaners (accessed on 10 January 2021).
22. digitec.ch. Steam Cleaners. Available online: https://www.digitec.ch/en/s1/producttype/steam-cleaners-126 (accessed on 10 January 2021).
23. Chan, J. The Best Affordable Robot Vacuums of 2021. Available online: https://www.reviewed.com/vacuums/best-right-now/best-affordable-robot-vacuums (accessed on 5 April 2021).
24. Alex Colon, A.M. The Best Robot Mops for 2021. Available online: https://www.tomsguide.com/best-picks/best-robot-mops (accessed on 5 April 2021).
25. LionsBot International Pte Ltd. LionsBot—Robots that Give Cleaners Superpowers. Available online: https://www.lionsbot.com/ (accessed on 15 August 2022).
26. Global Robotics Platform. Global Robotics—Cleaning Robots. Available online: https://www.allrobotsin.com/cleaning-robots/ (accessed on 15 August 2022).
27. Prabakaran, V.; Elara, M.R.; Pathmakumar, T.; Nansai, S. Floor cleaning robot with reconfigurable mechanism. *Autom. Constr.* **2018**, *91*, 155–165. [CrossRef]
28. Tun, T.T.; Huang, L.; Mohan, R.E.; Matthew, S.G.H. Four-wheel steering and driving mechanism for a reconfigurable floor cleaning robot. *Autom. Constr.* **2019**, *106*, 102796. [CrossRef]
29. Ilyas, M.; Yuyao, S.; Mohan, R.E.; Devarassu, M.; Kalimuthu, M. Design of sTetro: A modular, reconfigurable, and autonomous staircase cleaning robot. *J. Sens.* **2018**, *2018*, 8190802. [CrossRef]

30. Ramalingam, B.; Elara Mohan, R.; Balakrishnan, S.; Elangovan, K.; Félix Gómez, B.; Pathmakumar, T.; Devarassu, M.; Mohan Rayaguru, M.; Baskar, C. sTetro-Deep Learning Powered Staircase Cleaning and Maintenance Reconfigurable Robot. *Sensors* **2021**, *21*, 6279. [CrossRef]
31. Vega-Heredia, M.; Mohan, R.E.; Wen, T.Y.; Siti'Aisyah, J.; Vengadesh, A.; Ghanta, S.; Vinu, S. Design and modelling of a modular window cleaning robot. *Autom. Constr.* **2019**, *103*, 268–278. [CrossRef]
32. Muthugala, M.; Vega-Heredia, M.; Mohan, R.E.; Vishaal, S.R. Design and control of a wall cleaning robot with adhesion-awareness. *Symmetry* **2020**, *12*, 122. [CrossRef]
33. Sang, A.W.Y.; Moo, C.G.; Samarakoon, P.; Bhagya, S.; Muthugala, M.; Elara, M.R. Design of a Reconfigurable Wall Disinfection Robot. *Sensors* **2021**, *21*, 6096. [CrossRef]
34. Ramalingam, B.; Yin, J.; Rajesh Elara, M.; Tamilselvam, Y.K.; Mohan Rayguru, M.; Muthugala, M.; Félix Gómez, B. A human support robot for the cleaning and maintenance of door handles using a deep-learning framework. *Sensors* **2020**, *20*, 3543. [CrossRef]
35. Yin, J.; Apuroop, K.G.S.; Tamilselvam, Y.K.; Mohan, R.E.; Ramalingam, B.; Le, A.V. Table cleaning task by human support robot using deep learning technique. *Sensors* **2020**, *20*, 1698. [CrossRef] [PubMed]
36. Griffith, C.; Moore, G. An evaluation of the cleaning properties of a microfibre cloth. *Am. J. Infect. Control* **2005**, *33*, e35–e36. [CrossRef]
37. Wren, M.; Rollins, M.; Jeanes, A.; Hall, T.; Coen, P.; Gant, V. Removing bacteria from hospital surfaces: A laboratory comparison of ultramicrofibre and standard cloths. *J. Hosp. Infect.* **2008**, *70*, 265–271. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Strategic Energy Saving Technique for Multiple Visible Light Communication Access Points (VAP)

Vishal Kumar
Delhi Technological University, Delhi
vishalkumar_2k21phdec07@dtu.ac.in

Rohit Kumar
Delhi Technological University, Delhi
rohit.kumar@dtu.ac.in

Himani Mehta
Delhi Technological University, Delhi
himanimehta_2k21phdec16@dtu.ac.in

2023 15th International Conference on COMmunication Systems & NETworkS (COMSNETS) | 978-1-6654-7706-2/23/$31.00 ©2023 IEEE | DOI: 10.1109/COMSNETS56262.2023.10041406

***Abstract*—Visible light communication (VLC) system suffers from a major challenge of energy efficiency of the VLC access points (VAP). To encounter the challenge of energy wastage, a novel switching technique between active and idle modes for the multiple VAPs is proposed. The contrast of energy consumption between the proposed system and traditionally always active VAPs is analyzed. Further, the ideal probability of VAP is derived. In this work, the pilot signal is employed for changing the state of VAPs. Therefore, the consequence of false activation is well demonstrated.**

***Index Terms*—Visible light communication, Probability of true detection, Likelihood ratio test, Energy Efficiency, Receiver operating characteristic.**

I. Introduction

The rise in various activities such as the internet of things, gaming platforms, and online teaching platforms have led to expanding demands for boundless data in wireless communication. This in turn stipulates a broad range of spectrum and high data rates. Being a traditional way of communication, the radio frequency(RF) spectrum is overburdened day by day. Amidst the innumerable results, visible light communication (VLC) has been considered a reassuring wireless technique to acknowledge the insufficiency of radio frequency (RF) spectra, due to its superiority in transmission capacity, channel capacity, and link safety. It is one of the best solutions for indoor communication in the existing literature. In spite of its several advantages, VLC suffers from a major challenge of the energy efficiency of VAP. In VLC technology, the main aim is to employ visible light for communication as well as luminosity. Thus, there must be a proper trade-off between these two.

Approximately 66 $\%$ of the total power is apprehended to be consumed by the nodes at an idle state in wireless communication [1]. That shows a high amount of energy is underutilized or wasted. One of the recent technology to reduce this power consumption substantially is by utilizing a Wake-up system [2]. In [3], an algorithm is proposed in order to reduce the overall power expenditure by putting the small base stations (BS) in sleeping or energy-saving mode depending upon the demand of traffic for static traffic and dynamic traffic models. In [4], techniques for opportunistic sleep mode have been proposed. The suggested proposal make it possible for the minuscule cell BSs in RF communication to improve their downlink capability by an appropriate trade-off between quality of service (QOS) requirements of users and well-organized stability of load amid various base stations. In these two works, sleeping techniques are used for RF technology.

In [5], strategic sleeping for VLC access points is proposed for single VAP and single UE. In a single VAP, as the UE goes away from the VAP inside the room, the signal suffers from fading. To reduce the effect of fading in this article we have proposed multiple VAPs. In areas such as hospitals, RF is prohibited as it can cause a serious threat to patients, this proposed energy-efficient VLC system can be utilized. Other major applications lie in mines or airplanes where an energy-efficient system is of utmost importance.

In this article, we propose an energy-efficient VLC system based on the strategic sleeping of VAP. Herein, multiple access points (APs) are considered and based on the proposed methodology, theoretical and numerical analysis is done for energy consumption. Further, the ideal probability of the VAP is derived and analyzed from which it is inferred that it is needless for all the VAPs to be active all the time for communication with the User equipment (UE) at a specific position. A likelihood ratio test (LRT) is carried out to procure the probability of false activation and the probability of true detection of the pilot signal by VAPs.

This paper is outlined in different sections as: In section II proposed sleeping strategy and the system model is discussed. Section III is dedicated for derivation of idle probability of VAP and different parameters for the ROC curve. In section IV, numerical and simulation results were discussed and at last section V concludes this paper and future work is discussed.

II. System Model

In this work, 'k' number of VAPs are considered to be embedded in the ceiling of the room. The UE can be assumed to be located anywhere in the room. The overall proposed system model is demonstrated in Fig.(1). When UE wants to access a VAP then it broadcast a pilot signal (PS) to all the VAPs. The UE has the capability to track the signal-to-noise ratios (SNRs) of all the signals received from all the VAPs and hold them for a few seconds and choose the VAP whose signal wavelength is associated with the highest SNR. After receiving the PS, the VAP comes out of the idle state and starts transmitting the data. Once the connection is established rest all VAPs should be put in sleeping mode. One of the challenges

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 02,2023 at 06:43:30 UTC from IEEE Xplore. Restrictions apply.

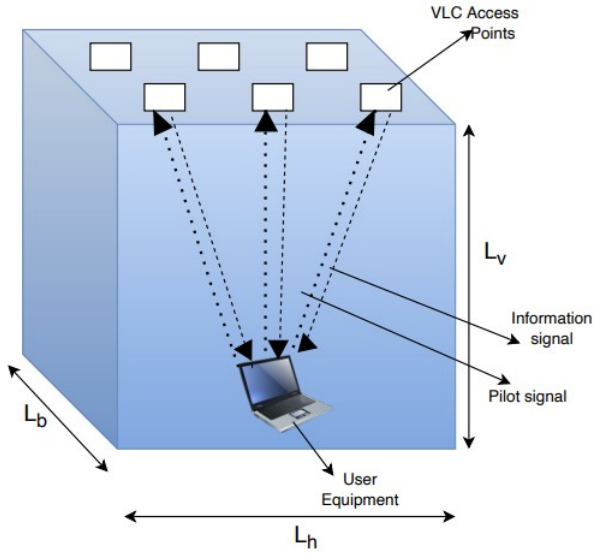


Fig. 1. System Model

in Multiple access points is signal interference due to the 'k' number of VAPs. To mitigate the effect of signal interference, the signal modulated at each VAP is dedicated to a different wavelength of visible light.

Multiple VAP's horizontal locations are modeled using a 2D homogeneous Poisson point process (PPP). Here, We assume that VAP is at a vertical distance of 'l_v' from the plane of UE. This PPP model is expressed as

$$\Psi_v = \{v_k, k \in \mathbb{N}\} \subset \mathbb{R}^2 \tag{1}$$

Here, 'v_k' is the horizontal distance between k^{th} VAP and the origin. Node density of VAP and UE is denoted as 'μ_v' and 'μ_u' respectively. DC channel gain of the VLC system is given by [6] as

$$h_{VLC} = \frac{A(m+1)R_p}{2\pi d^2}\cos^m(\alpha)U(\beta)G(\beta)\cos(\beta) \tag{2}$$

where '$U(\beta)$' is optical filter gain, '$G(\beta)$' is optical concentrator gain, 'A' is the physical area of the photodetector (PD), 'R_p' is responsivity of the PD, 'd' is the Euclidean distance between VAP and UE, 'β' is the angle of irradiance for VLC system.

$$m = \frac{-\ln 2}{\ln\left(\cos\left(\phi_{1/2}\right)\right)} \tag{3}$$

$$G(\beta) = \begin{cases} \frac{N^2}{\sin^2(\psi_c)} & ;0 \le \beta \le \psi_c \\ 0 & ;\beta > \psi_c \end{cases} \tag{4}$$

where 'N' is an internal refractive index, 'ψ_c' is a field of view (FOV) at the semi-angle of the concentrator, '$\phi_{1/2}$' is the LED's semi-angle at half power. By utilizing the geometrical feature of the VLC link [7],

$$\cos(\alpha) = \frac{l_v}{d} = \cos(\beta) \tag{5}$$

$$d = \sqrt{l_v^2 + v_k^2} \tag{6}$$

'ϑ_k' is the indicator of the transmission condition of VAP such that $\vartheta_k = 1$ indicates that VAP is active and transmission took place, $\vartheta_k = 0$ indicates VAP's sleep or idle state.

TABLE I
Notations and Descriptions

Notations	Descriptions
l_v	Vertical distance between the plane of VAP and UE
v_k	Horizontal distance between k^{th} VAP and the origin
μ_v	Node density of VAP
μ_u	Node density of UE
$U(\beta)$	Optical filter gain
$G(\beta)$	Optical concentrator gain
A	Physical area of the PD
R_p	Responsivity of the PD
β	Angle of irradiance
N	Internal refractive index
ψ_c	FOV at semi-angle of the concentrator
$\phi_{1/2}$	LED's semi-angle at half power
ϑ_k	Indicator of the transmission condition of VAP
P_a	Power consumed by that VAP which has the highest SNR
PL	Power consumed to send the pilot signals
$\hbar$	Coverage area of VAP
$\mu(\hbar)$	Standard Lebesgue measure
$1_{\hbar}(v_k)$	Metric of random counting
δ	Shape parameter
$\Lambda(Y)$	Likelihood ratio
Υ	Threshold of ratio test
χ	Transmitted power factor of pilot signal
P_F	Probability of false detection
P_T	Probability of true detection

Power consumed by VAP is denoted as P_{VAP} and expressed as

$$P_{VAP} = \begin{cases} P_a + PL & ;\vartheta_k = 1 \\ PL & ;\vartheta_k = 0 \end{cases} \tag{7}$$

where, P_a is the power consumed by that VAP which has the highest SNR and connected for the communication, PL is the power consumed to send the pilot signals.

III. Performance Analysis

In order to show the significance of the proposed sleeping strategy, we show that there is the possibility that some VAPs are not connected to any UE and are still in active mode. This can be elucidated from the probability of VAP to be in idle state. It is determined by the idle probability that the VAP will be in an idle state when no UEs are making data traffic requests so that VLC AP can be decisively shut off to save

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 02,2023 at 06:43:30 UTC from IEEE Xplore. Restrictions apply.

energy. Let's say $\hbar \subset \mathbb{R}^2$ is the coverage area of k^{th} VAP. Probability of n UE inside $\hbar$ is

$$P\left[\sum_{v_k} 1_{\hbar}(v_k) = n\right] = \mathbb{E}\left[\frac{(\mu_u \mu(\hbar))^n}{n!} \exp\{-\mu_u \mu(\hbar)\}\right] \quad (8)$$

where, $\mu(\hbar)$ is the standard Lebesgue measure of $\hbar$, $1_{\hbar}(v_k)$ is metric of random counting of $\hbar$ and is expressed

$$1_{\hbar}(v_k) = \begin{cases} 1 & ; v_k \in \hbar \\ 0 & ; \text{otherwise} \end{cases} \quad (9)$$

The probability density function (PDF) of '$\mu(\hbar)$' with 'δ' as shape parameter is given as [7]

$$f_{\mu(\hbar)}(t) = \frac{(\delta \mu_v)^{\delta}}{\Gamma(\delta)} t^{\delta - 1} \exp\{-\delta \mu_v t\} \quad (10)$$

The final expression for the probability of 'n' UEs present in the coverage area of VAP is derived

$$P\left[\sum_{v_k} 1_{\hbar}(v_k) = n\right] = \frac{\Gamma(n+\delta)}{n!\Gamma(\delta)} \left(\frac{\rho}{\delta+\rho}\right)^n \left(\frac{\delta}{\delta+\rho}\right)^{\delta} \quad (11)$$

Here, ρ is density ratio and given as $\rho = \frac{\mu_u}{\mu_v}$. Idle probability can be obtained by putting $n = 0$ and hence expressed as

$$P_{idle} = \left(\frac{\delta}{\delta+\rho}\right)^{\delta} \quad (12)$$

Initially, it is assumed that all VAPs are in a sleeping state. When it senses PS then it acknowledges that they have to come into the active stage and start sending data. In this process, it is probable that VAPs get activated by sensing some noise present in the environment. Therefore, it is vital to study the probability of false activation. LRT is carried out and the probability of false activation for one VAP is derived and similarly can be extended for k VAPs. H_1 and H_0 are the two hypotheses considered where H_1 denotes the case where VAP is activated by receiving the PS. On the other hand, H_0 represents that VAP is falsely activated by the noise. In this work, PS is modeled as an amplitude constraint signal and denoted with constant 'P'. η_j are statistically independent Gaussian random variable with S samples which represent the noise signal with σ_P^2 as variance. These two hypotheses are represented as

$$H_1 : \; y_j = P + \eta_j \quad ; j = 1, 2, 3, ..., S \quad (13)$$

$$H_0 : \qquad y_j = \eta_j \quad ; j = 1, 2, 3, ..., S \quad (14)$$

y_j is the received signal by access points of VLC. The probability density of the noise is expressed as

$$f_{\eta_j}(\eta) = \frac{1}{\sqrt{2\pi\sigma_P^2}} \exp\left(\frac{-\eta^2}{2\sigma_P^2}\right) \quad (15)$$

PDF of y_j under hypothesis H_1 is derived and found to be

$$f_{y_j}(y_j/H_1) = \frac{1}{\sqrt{2\pi\sigma_P^2}} \exp\left(\frac{-(y_j - P)^2}{2\sigma_P^2}\right) \quad (16)$$

Similarly, PDF of y_j under hypothesis H_0

$$f_{y_j}(y_j/H_0) = \frac{1}{\sqrt{2\pi\sigma_P^2}} \exp\left(\frac{-y_j^2}{2\sigma_P^2}\right) \quad (17)$$

A joint PDF of the received signal y_j for both hypotheses can be given as

$$f_{y_j}(y_j/H_1) = \prod_{j=1}^{S} \frac{1}{\sqrt{2\pi\sigma_P^2}} \exp\left(\frac{-(y_j - P)^2}{2\sigma_P^2}\right) \quad (18)$$

$$f_{y_j}(y_j/H_0) = \prod_{j=1}^{S} \frac{1}{\sqrt{2\pi\sigma_P^2}} \exp\left(\frac{-y_j^2}{2\sigma_P^2}\right) \quad (19)$$

To carry out LRT, the likelihood ratio is found. Likelihood ratio is denoted by '$\Lambda(Y)$' and is obtained as

$$\Lambda(Y) = \frac{\prod_{j=1}^{S} \frac{1}{\sqrt{2\pi\sigma_P^2}} \exp\left(\frac{-(y_j - P)^2}{2\sigma_P^2}\right)}{\prod_{j=1}^{S} \frac{1}{\sqrt{2\pi\sigma_P^2}} \exp\left(\frac{-y_j^2}{2\sigma_P^2}\right)} \quad (20)$$

By the application of logarithm and performing simplification we get

$$\ln \Lambda(Y) = \left(\frac{P}{\sigma_P^2} \sum_{j=1}^{S} y_j\right) - \frac{SP^2}{2\sigma_P^2} \quad (21)$$

LRT is expressed as

$$\ln \Upsilon \underset{H_0}{\overset{H_1}{\lessgtr}} \left[\left(\frac{P}{\sigma_P^2} \sum_{j=1}^{S} y_j\right) - \frac{SP^2}{2\sigma_P^2}\right] \quad (22)$$

'Υ' is the threshold of the ratio test and it is a function of cost and priori probabilities. Probability of PS false detection denoted by P_F and probability of true detection denoted as P_T can be given as

$$P_T = \int_{z_1} f_{y_j}(y_j/H_1) dy_j \quad (23)$$

$$P_F = \int_{z_0} f_{y_j}(y_j/H_0) dy_j \quad (24)$$

where, H_1 is true hypothesis in the neighbourhood of region z_1 for threshold Υ.

IV. Results and Analysis

Discussions of numerical and simulation results are presented in this section. Dimension of the room is considered to be of $l_v \times l_h \times l_b$ where specifically we have taken dimension as $(4 \times 10 \times 10)\ m^3$. δ is taken as 3.5 in this work. The transmitted power of one LED is $20mW$ and 40×40 LEDs array is used. Semi-angle at half power and FOV of the receiver of the proposed VAP is 70^o.

From Fig. (2), it can be clearly observed that as we increase the number of UEs in a coverage area of a particular VAP, that means the density ratio increases, and hence Probability of a VAP being idle reduces. The simulation results for energy consumption for the proposed Multiple VAPs with and without Strategic sleeping are shown in Fig.(3). From Fig.(3), it is

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 02,2023 at 06:43:30 UTC from IEEE Xplore. Restrictions apply.

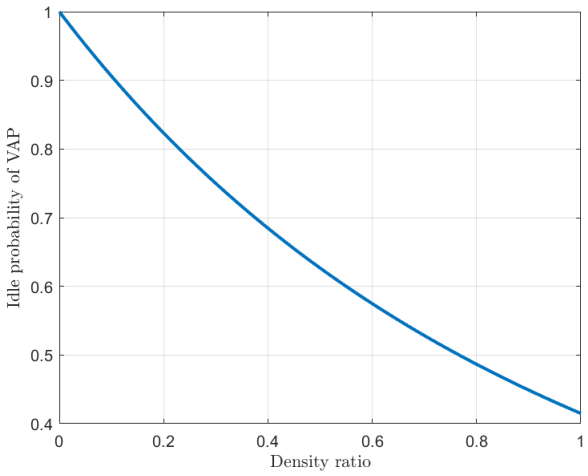


Fig. 2. Characteristic of Idle probability

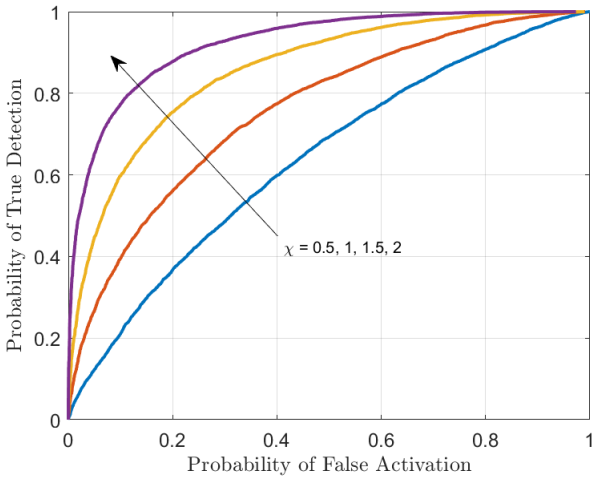


Fig. 4. Receiver operating characteristic

clearly evident that we can save a lot of energy by utilizing the proposed approach.

In Fig. (4), the simulation result of ROC is shown. ROC curve is plotted for different values of power transmitted factor $\chi = \{0.5, 1, 1.5, 2\}$ by UE in the form of the pilot signal. It is obvious that as we increase the power content of the PS the true detection increases and it is evident from Fig. (4). Moreover, it can be observed that at any point

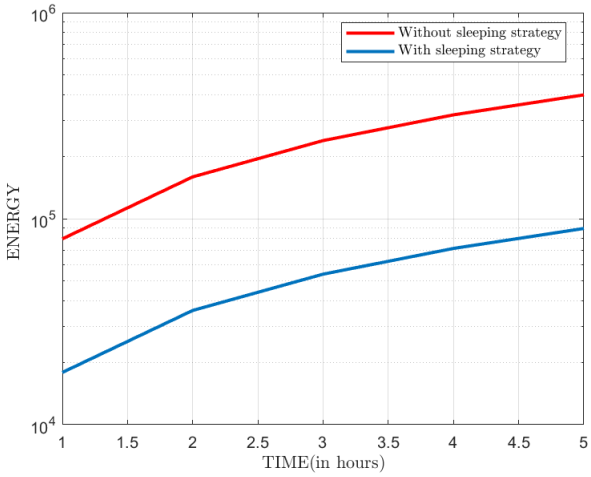


Fig. 3. Amount of Energy saving as a function of time

probability of true detection is greater than the probability of false activation which notably justifies our proposed strategical energy-saving method. Further, we try to analyze the impact of power received by UE at different positions for single VAP and multiple VAPs. From Fig. (5), we can observe the power received at different positions by UE for single VAP reduces while for multiple VAPs it remains uniform throughout the dimension of our room.

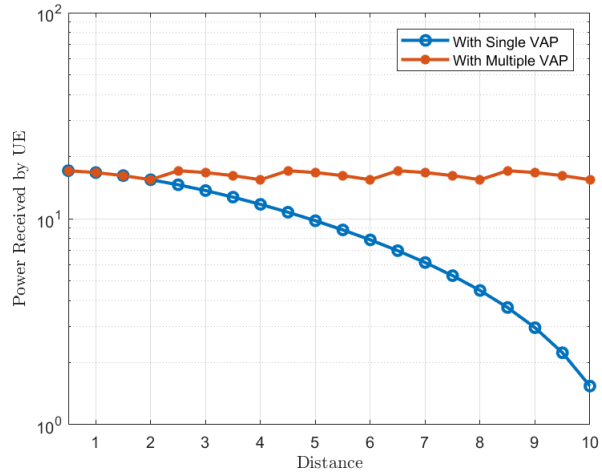


Fig. 5. Comparison between Multiple VAP and Single VAP

V. Conclusion

In this paper, an energy-efficient Multiple VAP system is proposed. Through the analysis of idle probability, it is shown that there is a probability of multiple VAPs being in an idle state. So, we can switch it OFF or put it in an energy-saving mode to save a lot of energy. Further, probability of false activation is also analyzed. One advantage of the proposed work also lies in the application of multiple VAPs to reduce signal fading.

This work can be extended by analyzing the different parameters like throughput and latency. As, we need to take care that while changing the state of a VAP, there should not be a loss of data. Re-configurable Intelligent Surface (RIS) can be incorporated to reduce the fading effect of signal when UE is mobile.

References

[1] Key World Energy Statistics 2014, International Energy Agency, 2014.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 02,2023 at 06:43:30 UTC from IEEE Xplore. Restrictions apply.

[2] Oller, J., Demirkol, I., Casademont, J., Paradells, J., Gamm, G.U. and Reindl, L., 2013, November. Wake-up radio as an energy-efficient alternative to conventional wireless sensor networks MAC protocols. In Proceedings of the 16th ACM international conference on Modeling, analysis simulation of wireless and mobile systems (pp. 173-180).

[3] C. Liu, B. Natarajan and H. Xia, "Small Cell Base Station Sleep Strategies for Energy Efficiency," in IEEE Transactions on Vehicular Technology, vol. 65, no. 3, pp. 1652-1661, March 2016, doi: 10.1109/TVT.2015.2413382.

[4] S. Samarakoon, M. Bennis, W. Saad and M. Latva-aho, "Opportunistic sleep mode strategies in wireless small cell networks," 2014 IEEE International Conference on Communications (ICC), 2014, pp. 2707-2712, doi: 10.1109/ICC.2014.6883733. S. Samarakoon, M. Bennis, W. Saad and M. Latva-aho, "Opportunistic sleep mode strategies in wireless small cell networks," , 2014, pp. 2707-2712, doi: 10.1109/ICC.2014.6883733..

[5] V. Kumar, A. Gupta and P. Garg, "Strategic Sleeping of Visible Light Communication Access Point," 2021 6th International Conference on Signal Processing, Computing and Control (ISPCC), 2021, pp. 387-392, doi: 10.1109/ISPCC53510.2021.9609445.

[6] L. Yin, W. O. Popoola, X. Wu and H. Haas, "Performance Evaluation of Non-Orthogonal Multiple Access in Visible Light Communication," in IEEE Transactions on Communications, vol. 64, no. 12, pp. 5162-5175, Dec. 2016, doi: 10.1109/TCOMM.2016.2612195.

[7] Ferenc, Járai-Szabó, and Zoltán Néda. "On the size distribution of Poisson Voronoi cells." Physica A: Statistical Mechanics and its Applications 385, no. 2 (2007): 518-526.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on March 02,2023 at 06:43:30 UTC from IEEE Xplore. Restrictions apply.

KSCE Journal of Civil Engineering
DOI 10.1007/s12205-023-0020-5

pISSN 1226-7988, eISSN 1976-3808
www.springer.com/12205

Hydraulic Engineering

Study of the Energy Dissipation over the Type-A Piano Key Weir

Deepak Singh[a] **and Munendra Kumar**[a]

[a]*Dept. of Civil Engineering, Delhi Technological University, New Delhi 110042, India*

ARTICLE HISTORY

Received 7 February 2022
Revised 18 July 2022
Accepted 5 January 2023
Published Online 22 February 2023

KEYWORDS

Energy dissipation overtype-A PKW
Cost-effective structure
Less footprint structure
Sensitivity analysis
Steps at outlet key

ABSTRACT

This study assesses the effects of geometrical variations, i.e., relative width ratio, magnification ratio, and cyclic variation, on the Piano Key Weir energy dissipation performance. In order to assess this, 12 type-A PKW models were tested by conducting an experimental investigation. The analysisincludes results from 280 tests and comprehensive findings of the fluid domain. The results demonstrated that the configuration of the PKWs has a greater influence on energy dissipation. The present study's findings show that the type-A PKW dissipates energy more efficiently than nonlinear or standard weirs. The energy loss over the PKW decreases as the magnification and relative width ratios increase, although the energy dissipation increases with key cycle numbers. Furthermore, the sensitivity analysis of the weir was examined by making steps at outlet key floors and found that the energy dissipation has been increased by 6.67% at low heads while reduced by 1.59% at higher heads. In addition, the author proposed two empirical projection equations for the optimal relative residual energy configuration downstream of the PKW. The proposed equations are the function of headwater ratio, relative width ratio, and magnification ratio. The results of this investigation are in perfect correlation with earlier studies.

1. Introduction

Piano Key Weir (PKW) is a modified labyrinth-type weir intended for applications with a small weir footprint and was first represented in 2003 by Lempérière and Ouamane (2003). PKWs can be alternatives to linear or other nonlinear overflow structures for providing a longer crest length for anequal spillway width (Leite Ribeiro et al., 2011). Lempérière and Ouamane (2003) desired to design a folded weir with a smaller base when sustaining a straightforward, cost-effective, and easily constructible structure. As a result of early collaborations between academia and industry, the concept evolved significantly. The discharge capacity of existing gravity dams in France has been considerably increased by PKW design (Laugier et al., 2017), and gated weirs have been replaced with an exciting alternative in Vietnam (Ho Ta Khanh, 2017). Since 2011, the regular evaluation of the PKW has been freely shared globally in an accessible and free environment through international workshops. As a result, three reference books (Erpicum et al., 2011, 2013 and 2017) have been published to highlight the state of the art through that time. Since 2006, over 35 PKWs have been built around the world. In France, the first PKW was constructed above the Goulours dam (Crookston et al., 2019), and the research has continued worldwide, averaging 15 large contributions a year since 2010. Recently, a detailed review study of the hydraulic behaviour and geometrical configuration of PKW was presented by Singh and Kumar (2021).

Interestingly, researchers have also demonstrated interest in PKW's applications as a side weir for flood release (Karimi et al., 2018). They found that using PKWs as side weirs results in significantly higher hydraulic efficiencies than linear side weirs. Much research has been done on the upstream characteristics of PKWs. PKW's downstream hydraulic qualities have only been studied in a limited number of investigations so far. Therefore, more research is required to assess the hydraulic behaviour downstream of the PKW. Energy dissipation is a vital component in examining the downstream aspects of PKW. As a result, information on the PKWs' energy dissipation must be recorded and evaluated. PKW with stepped spillways (Ho Ta Khanh et al., 2011a) or stilling basins (Pfister et al., 2017; Truong and Ho Ta Khanh, 2017) is the subject of systematic experimental studies. The effect of PKW inlet-outlet key slopes (*B/P*) on energy loss was assessed by Al-Shukur and Al-Khafaji (2018) and reported that the relative energy loss diminishes as the key slopes diminish. There's also a correlation between how far the hydraulic jump

CORRESPONDENCE Munendra Kumar munendrakumar@dtu.ac.in Dept. of Civil Engineering, Delhi Technological University, New Delhi 110042, India

© 2023 Korean Society of Civil Engineers

from the PKW's toe is and how much of an energy drain it is as the hydraulic jump distance increases.

Additionally, the research mentioned earlier studies (Leite Ribeiro et al., 2007; Bieri et al., 2011; Leite Ribeiro et al., 2011; Ho Ta Khanh et al., 2011a; Erpicum et al., 2013) (Ghaderi et al., 2020) are the specific research based on the different downstream elementsconsideration of labyrinths and PKWs (Leite Ribeiro et al., 2007, 2011) demonstrated the idea of stepped outlet floors of the PKW and a "ski-jump" for the spillway at St-Marc Dam. Six models were developed and tested to dissipate flow energy, including slab baffle blocks, natural rock protection, shifting ski jumps, and deflector walls. PKWs with a short stilling basin could be sufficient for small dam projects to dissipate energy; for medium- to large dams, PKWs with stepped spillways and a short stilling basin would be appropriate (Ho Ta Khanh et al., 2011b). Under standardized conditions, overstepped spillways disperse energy faster than ogee weirs (Erpicum et al., 2013).

Moreover, Silverstri et al. (2013) illustrated the critical effort of (Erpicum et al., 2011) by testing three different weirs (one classical OCW ("Ogee-crested weir") or two different PKWs) upstream of the stepped spillway by considering four spillway lengths. They found that downstream of a PKW, the flow energy at the spillway toe is much higher than downstream of an OCW, regardless of spillway length. Recently, a detailedstudy of energy dissipation over type-A PKWs based on the different relative width ratios was investigated by Eslinger and Crookston (2020). They found that the type-A PKW has better dissipating efficiency than its alternatives. In order to estimate the downstream PKW's residual energy, they proposed two empirical prediction equations. Singh and Kumar (2022a) and Singh and Kumar (2022b) performedan experimental and computational analysis to assess the residual energy at the toe of type-B PKW, respectively. They stated that the type-B PKW is extremely sensitive to low head or discharge. Moreover, Singh and Kumar (2022c) also developed new equations based on GEP to assess the residual energy over the type-A PKW.

In addition, the energy loss over the weir or PKWs is also connected to the research of local scour at the base of PKWs placed in canals and rivers. Jüstrich et al. (2016) presented a detailed study of jet scour on a mobile bed, focusing on assessing the scour and the adjacent ridge formation near the PKW. A skijump reduces the flowing jet's impact, and consequences enhance energy dissipation Noui and Ouamane (2013). Toe scour reduces the weir stability if a rigid foundation is not possible (Pfister et al., 2017). If the flow exiting the PKW is affected by the downstream face of foundations, scouring potential is reduced. By examining the homogenous non-cohesive granular bed material under distinct tailwater depth as well asdischarge combinations, Palermo et al. (2020) demonstrated the first PKW's scouring morphology downstream. In addition, Kumar and Ahmad (2022) studied the scouring behaviour of PKWs with and without using a solid apron downstream under both free and submerged flow conditions. Further, Kumar et al. (2021) studied the movement of the sediment particles across type-A PKWs. They observed that sediment speeds up as it enters the key and slows as it moves upstream toward the intake key.

Yazdi et al. (2021) examined the impact of PKW's geometry on downstream scouring. They claim that the weir configuration and the releasing rate affect the scour characteristics. If the particle shape is not spherical, the saltation length and height increase (Zhao et al., 2020). Experimental research on hydraulic jumps formedonthe pebbled rough bedwas presented by Bahmanpouri et al. (2021). Overall, the study learned something new about the hydraulic jump on a gravel bed, highlighting differences and similarities with those on a rigid bed. Kheilapour et al. (2022) demonstrated that increasing the magnification ratio and flow discharge in a labyrinth weir caused more flow interference downstream and more flow transverse curvature upstream while decreasing the discharge coefficient. In addition, Singh and Kumar (2022d) investigated the aeration performances of the different types (type-A, B, and C) of the PKW and found that the type-A PKW has better aeration efficiency than others.

Even though the fact that several studies on the flow properties, design guidelines, and scouring of PKWs and labyrinth weirs have been published, extensive research on the downstream flow regime, energy dissipation, and protection measures of the various types of PKWs is still lacking. In order to better understand this, thepresent study examines the effectsof differentgeometrical parameters, i.e., relative width ratio, magnification ratio, and cyclic variation, on the type-A PKW's energydissipation capacity and downstream flow bahviour. Moreover, the weir's sensitivity istested by employing the steps at the outlet key. In addition, to aid design efforts, data of the present study were used to propose the prediction of residual energy downstream of type-A PKWs.

2. Materials and Methods

2.1 Experimental Setup

The experiments were performed in a flume of dimensions 10 m long, 0.516 m wide, and 0.6 m deep in the "FM & HE Laboratory at DTU, Delhi." Flow to the channel is provided by a 20-horsepower pump and a network of 4-inch supply pipes, a flow regulating valve to control the discharge, quantitative supplying discharges of 50 L/s, and calibration using an orifice meter fitted to the flume supply pipe (0.25% uncertainty). Uniformity is maintained with a baffle wall or metal gate. A tailgate controls the downstream water depth. An ultrasonic level sensor of 4 – 20 mA (0.2% uncertainty) and a pointer gauge with the least count of 0.1 mm for the head over the weir crest is attached to the flume. A 4 – 20 mA electromagnetic flowmeter (accuracy ± 0.2%) is equipped with a supply pipe for the discharge measurement (see Fig. 1). The average flow velocities were measured by ADV ("Acoustic Doppler Velocimeter") device.

The current study estimates downstream energy dissipation using 12 type-A PKW models in various geometric features (see Table 1). The models were made with a transparent acrylic sheet of 8 mm thick and affixed with chloroform.The configurations of the models are as follows: magnification ratio ($L/W = 5 - 6$), weir

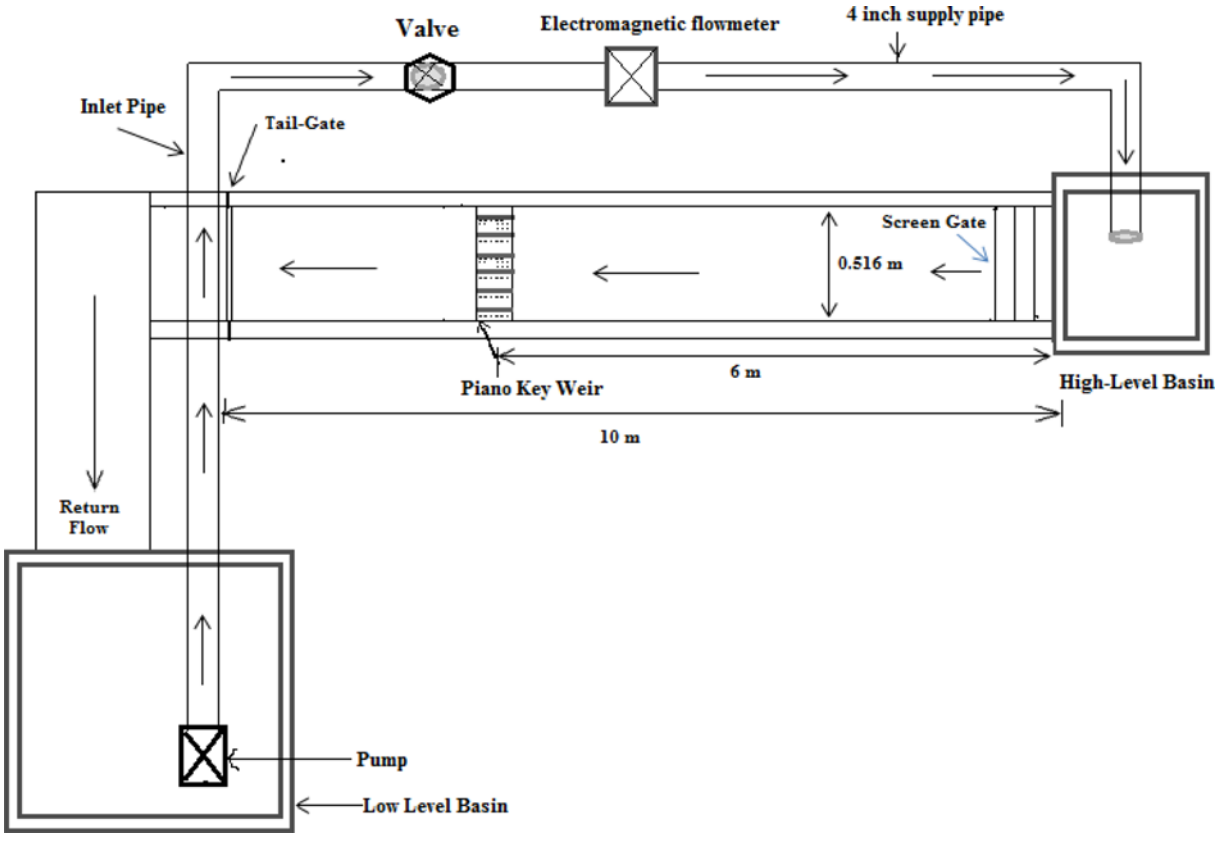


Fig. 1. Schematic Plan View of the Experimental Setup

Table 1. Range of Data Collected in the Present Study [type-A PK weirs]

Model No.	Range of Q (m^3/s)	Range of H_t (m)	$\frac{W_i}{W_o}$	W_i (m)	W_o (m)	P (m)	$\frac{L}{W}$	B (m)	B_i (m)	B_o (m)	Range of $\left(\frac{E_L}{E_1}\right)$	Range of $E_r = \frac{E_2}{E_1}$	N (No. of cycles)	No. of runs
PKW-1	0.005-0.050	0.0170-0.165	1.0	0.078	0.078	0.20	5	0.28	0.093	0.093	0.8413-0.153	0.158-0.8419	3	20
PKW-2	0.005-0.050	0.0168-0.167	1.1	0.082	0.075	0.20	5	0.28	0.093	0.093	0.8310-0.1460	0.1682-0.8504	3	20
PKW-3	0.005-0.050	0.0169-0.164	1.2	0.085	0.071	0.20	5	0.28	0.093	0.093	0.8180-0.1531	0.189-0.853	3	20
PKW-4	0.005-0.050	0.0168-0.154	1.3	0.089	0.068	0.20	5	0.28	0.093	0.093	0.8012-0.1601	0.198-0.843	3	20
PKW-5	0.005-0.050	0.0171-0.151	1.4	0.091	0.065	0.20	5	0.28	0.093	0.093	0.7931-0.1476	0.206-0.852	3	20
PKW-6	0.005-0.050	0.0168-0.160	1.5	0.094	0.063	0.20	5	0.28	0.093	0.093	0.7841-0.1301	0.215-0.861	3	20
PKW-7	0.005-0.050	0.0171-0.147	1.28	0.088	0.069	0.15	5	0.343	0.115	0.115	0.8110-0.1471	0.1889-0.8531	3	20
PKW-8	0.005-0.050	0.0169-0.145	1.28	0.064	0.05	0.15	5	0.259	0.086	0.086	0.833-0.136	0.167-0.8642	4	20
PKW-9	0.005-0.050	0.0172-0.146	1.28	0.05	0.039	0.15	5	0.208	0.069	0.069	0.853-0.1501	0.147-0.8498	5	20
PKW-10	0.005-0.050	0.0172-0.145	1.28	0.088	0.069	0.15	6	0.427	0.142	0.142	0.8001-0.1302	0.199-0.869	3	20
PKW-11	0.005-0.050	0.0168-0.141	1.28	0.064	0.05	0.15	6	0.322	0.107	0.107	0.8215-0.1492	0.179-0.8511	4	20
PKW-12	0.005-0.050	0.0181-0.148	1.28	0.05	0.039	0.15	6	0.259	0.086	0.086	0.8439-0.1538	0.1574-0.8461	5	20
PKW-7 (Stepped)	0.005-0.050	0.0168-0.147	1.28	0.088	0.069	0.15	5	0.343	0.115	0.115	0.867-0.1301	0.1331-0.8692	3	20
PKW-10 (Stepped)	0.005-0.050	0.0169-0.146	1.28	0.088	0.069	0.15	6	0.427	0.142	0.142	0.853-0.1471	0.1475-0.8527	3	20

height (P = 0.15 – 0.20 m), relative width ratio ($1.0 \le W_i/W_o \le 1.5$), and overhang portions ($B_i = B_o$) (see Table 1). The experiments were performed for desired discharges ($0.005\ m^3/s \le Q \le 0.05\ m^3/s$), head to weir ratio ($0.085 \le H_t/P \le 0.85$), for 0.20 m tall models, $0.11 \le H_t/P \le 1.0$ for 0.15 m tall models, and the total head over crest (0.0168 m $\le H_t \le$ 0.167 m), other dimensionless ratios ($0.43 \le (B_i/P = B_o/P) \le 0.95$).

The measurements were taken at 2P upstream from the PKWs' lateral centerline and 8P downstream after the downstream water level had stabilized (see Fig. 2). The downstream energy

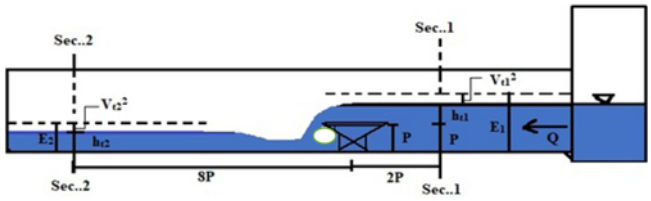


Fig. 2. Schematic Plot for Specific Energy Measurement

measurement locations were chosen based on the complex flow field exiting immediately downstream of the PKWs. With a nearly constant flow depth, the downstream needed for the flow field in a given cross-section gradually changes. The downstream flume's length is around 4 m, less than the upstream but enough to study the downstream morphology (i.e., hydraulic jump formation, study the downstream energy dissipator structures, local scour, design of downstream apron, etc.). The experiments in this study were conducted in a laboratory flume using the channel flow approach; thus, the study primarily focused on PKWs constructed at river barrages or serving as a control structure in canals.

2.2 Methodology

The present investigation was carried out under free-flow conditions, with no artificial control of the downstream or tailwater. Downstream and upstream PKW's energyhead was determined using the following formula:

$$E_i = P + h_{ti} + \frac{V_{ti}^2}{2g} \quad (1)$$

where E = energy at section i, V_t = velocity (m/s) at section i, P= weir height (m) (P = 0 m for downstream), h_t = head at section i, and i = section (i = 1, 2). The upstream depth ($P + h_{t1}$) and downstream depth flow h_{t2} are determinedusinga pointer gauge and an ultrasonic level sensor after the water surface becomes uniform for at least 3 to 5 minutes; as can be seen in Fig. 2, it is capable of increasing with an increase in flow rate. The mean approach flow velocity V_t was calculated using a Sontek ADV at 0.25, 0.5, and 0.75W flume cross-sections as the average velocity over the flume's width W. The velocity analysis revealed that these average cross-sectional velocities and mean approach velocities agreed (by ADV); consequently, the variation in H_t for the Q ranges is less than 5%. Table 1 presents the PKW's main geometric characteristics and the data set ranges provided in this study, and the geometrical sketch of type-A PKW is shown in Fig. 3.

The relative energy dissipation [$(E_1 - E_2)/E_1 = \Delta E/E_1$] and residual energy ($E_2/E_1$) were calculated using the values of E_1 and E_2. The geometrical as well as hydraulic parameters, are arranged in Eq. (2) determine energy dissipation across the PKW.

$$\frac{\Delta E}{E_1} = f(\rho, g, \mu, \sigma, V_t, H_t, L, L_{cy}, P, W, W_i, W_o, W_{cy}, S_i, S_o, B, B_i, B_o, R, h_s, l_s) \quad (2)$$

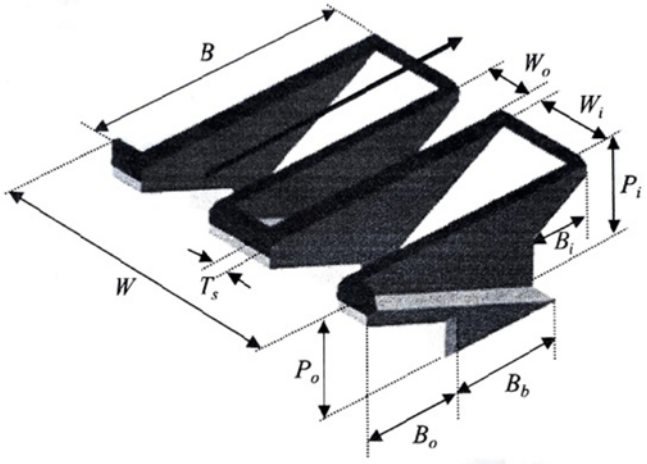


Fig. 3. Fundamental Parameters of a Type-A PKW – 3D View (Adapted from Pralong et al., 2011)

where ρ = density of flowing fluid, g = gravitational acceleration, and $H_t = (h_t + V_t^2/2\,g)$ total u/s head over the PKWs. In addition to the hydraulic parameters, the geometrical parameters of interest include L, and L is the total crest length. One cycle's crest length is measured in L_{cy}. $P$is the weir height. PKW's overall width is represented by W, and the widths of the outlet and inlet keys are given by W_o and W_i, correspondingly. W_{cy} represents the width of one cycle. The slopes of the outlet and inlet keys are S_o and S_i, correspondingly. Rrepresents the parapet height. The height and length of the steps are l_s and h_s, respectively, and the N is the number of units. Applying Buckingham theorem and determining the dimensionless parameters by taking P, g, and ρ as repeating variables (Eqs. (3), (4), (5)). F_r, R_e, and W_e represent the Froude number, Reynolds number, and Weber's, respectively.

$$\prod_1(H_t) = \frac{H_t}{P}, \prod_2(L) = \frac{L}{P}, \prod_3(L_{cy}) = \frac{L_{cy}}{P}, \prod_4(W) = \frac{W}{P},$$
$$\prod_5(W_i) = \frac{W_i}{P}, \prod_6(W_o) = \frac{W_o}{P}, \prod_7(W_{cy}) = \frac{W_{cy}}{P}, \prod_8(S_i) = S_i,$$
$$\prod_9(S_o) = S_o, \prod_{10}(B) = \frac{B}{P}, \prod_{11}(B_i) = \frac{B_i}{P}, \prod_{12}(B_o) = \frac{B_o}{P},$$
$$\prod_{13}(R) = \frac{R}{P}, \prod_{14}(V_t) = \frac{V_t}{\sqrt{g^P}} = F_r, \prod_{15}(\mu) = \left(\frac{\mu}{\rho g^{\frac{1}{2}} P^{\frac{3}{2}}}\right) = Re,$$
$$\text{and } \prod_{16}(\sigma) = \left(\frac{\sigma}{\rho g P^2}\right) = W_e \prod_{17}(l_s) = \frac{l_s}{P}, \text{ and } \prod_{18}(h_s) = \frac{h_s}{P} \quad (3)$$

$$\prod(L) \times \frac{1}{\prod(W)} = \frac{1}{P} \times \frac{P}{W} = \frac{L}{W}, \prod(W_i) = \frac{1}{\prod(W_o)} = \frac{W_i}{P} \times \frac{P}{W_o} = \frac{W_i}{W_o},$$
$$\prod(S_i) \times \frac{1}{\prod(S_o)} = \frac{S_i}{S_o}, \prod(B_i) = \frac{1}{\prod(B)} = \frac{B_i}{B}, \prod(B_o) \times \frac{1}{\prod(B)} = \frac{B_o}{B},$$
$$\prod(h_s) \times \frac{1}{\prod(l_s)} = \frac{h_s}{l_s} \quad (4)$$

$$\prod(L)\times\frac{1}{\prod(L_{cy})}=\frac{L}{L_{cy}}=N=\prod(W)\times\frac{1}{\prod(W_{cy})}=\frac{W}{W_{cy}}=N \quad (5)$$

The following equations are found by rearranging all parameters:

$$\frac{\Delta E}{E_1}=f\left(\frac{H_t}{P},\frac{L}{W},\frac{W_i}{W_o},\frac{S_i}{S_o},\frac{B_i}{B},\frac{B_o}{B},\frac{h_s}{l_s},N,F_r,R_e,W,\frac{R}{P}\right) \quad (6)$$

The PKW's upstream flow is usually subcritical because the flow hits laterally to the weir; thus, their hydraulic characteristics are not evaluated using the Froude number (Haghiabi et al., (2021). The primary causes of scale effects are model roughness and model approach conditions associated with turbulent boundary layer development, surface tension effects, associated aeration, vortex-formation problems, and cavitation phenomena. These scale effects can be avoided or reduced using model scales with sufficiently high Reynolds and Weber numbers. When comparing PKW energy dissipation to the maximum limit of energy dissipation, the relative PKW energy dissipation should be related to the relative minimum head-to-water depth ratio (H_t/P).

Using this point and ignoring the effects of viscosity and surface tension on the PKW's hydraulic characteristics (Mohammadzadeh-Habili et al., 2018). As a result, Eq. (6) can be modified as Eq. (7).

$$\frac{\Delta E}{E_1}=f\left(\frac{H_t}{P},\frac{L}{W},\frac{W_i}{W_o},\frac{S_i}{S_o},\frac{B_i}{B},\frac{B_o}{B},\frac{h_s}{l_s},N,\frac{R}{P}\right) \quad (7)$$

3. Experimental Results and Discussions

The PKW flow is highly aerated and three-dimensional, with splash and spray regions within the outlet keys andat the structure's base (see Fig. 4). The area of sprinkling and spraying increases slightly when H_t and the planar jet begin to descend on the crest. On the other hand, the aeration zone rose dramatically with H_t, partially due to increased local speed, which resulted in more significant turbulent mixing and advection levels (Singh and Kumar, 2022a). The nappes crossing across the PKW crest are figured to have three planar jets interrelating across an inclined bottom (see Fig. 4), with jet augmentation, jet contact between

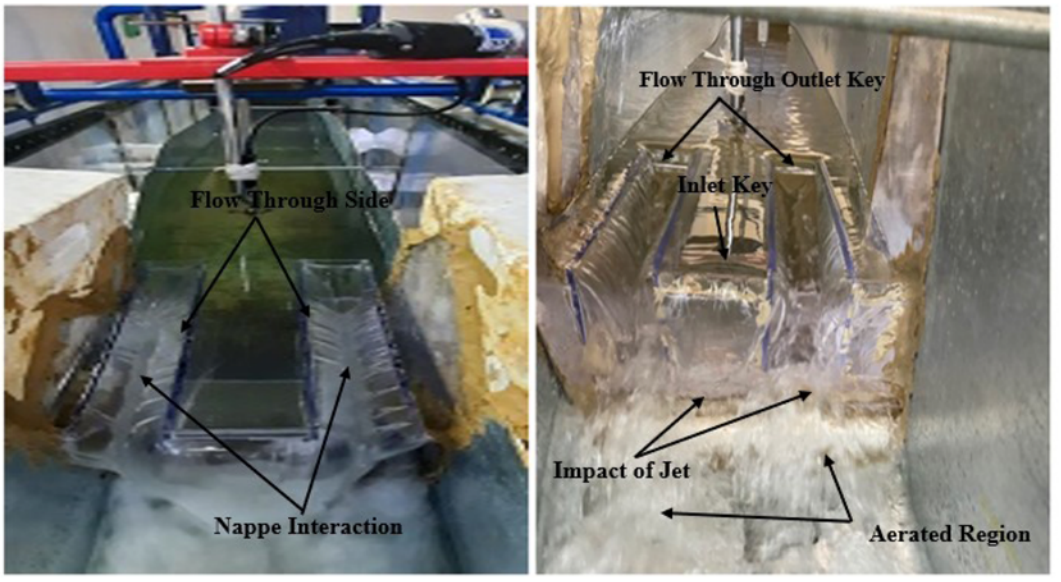


Fig. 4. Flow Pattern over PKW: (a) Flow Pattern over the PKW (Top view), (b) Flow Pattern over the PKW (Side view)

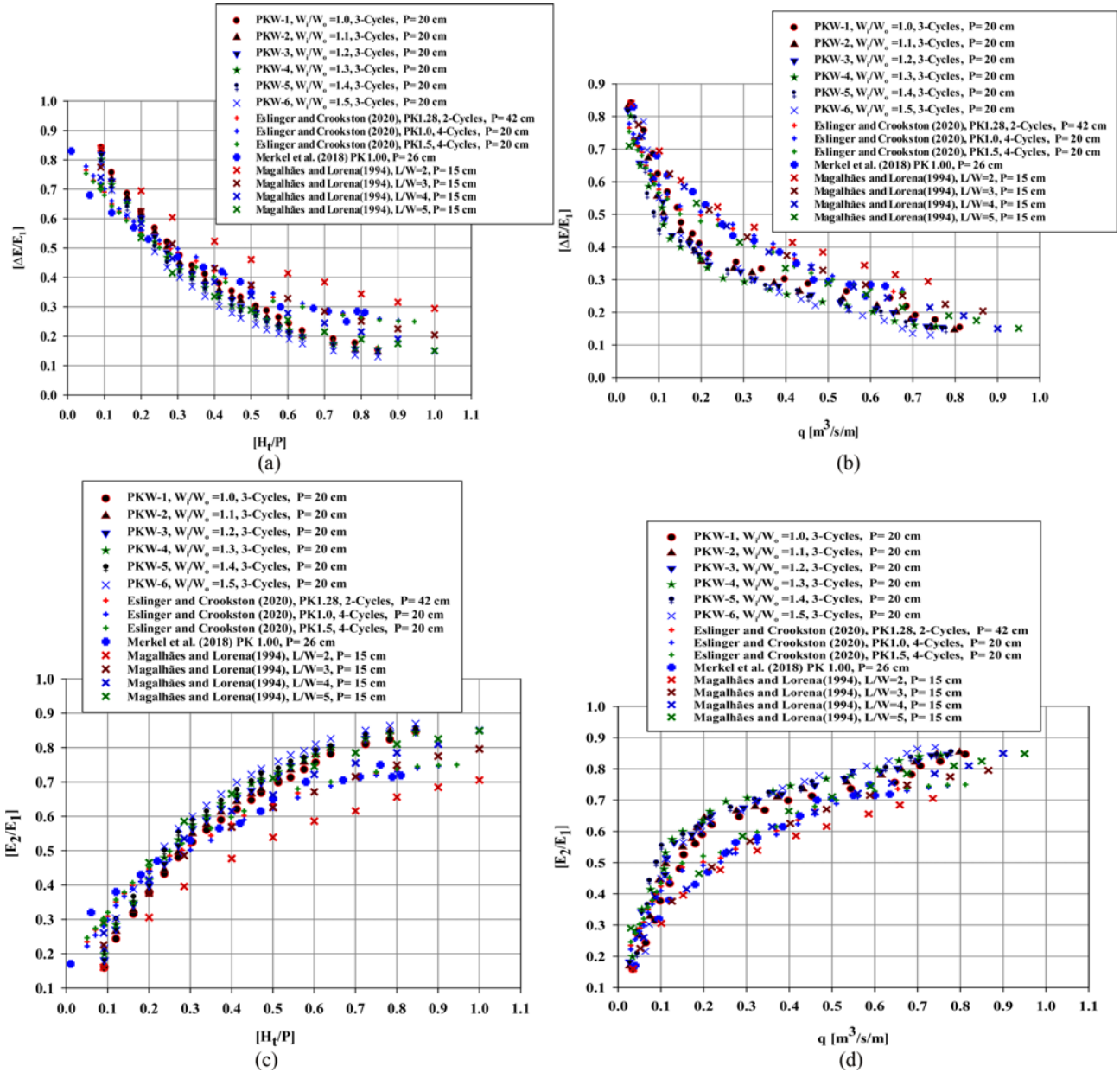


Fig. 5. Energy Dissipation of Flow over Type-A PKW for Different Width Ratios (W_i/W_o): (a) Relative Energy Dissipation ($\Delta E/E_1$) with respect to Head Water Ratio (H_t/P), (b) Relative Energy Dissipation ($\Delta E/E_1$) with respect to Unit Discharge (q), (c) Relative Residual Energy (E_2/E_1) with respect to Head Water Ratio (H_t/P), (d) Relative Residual Energy (E_2/E_1) with Respect to Unit Discharge (q)

opposite nappes with an opposite region of the water surface, a crossing area, and minor wakes at these upstream apexes. During the research, the local submergence rose and took up a more significant proportion of the $B_o + B_b$ streamwise distance in the area of $H_t/P \geq 0.25$ until most of the space was locally submerged. Submergence, in particular, reduces discharge efficiency and actively accounts for the loss in energy losses for greater H_t/P ratios, as described in the following section.

This research aimed to acquire information on how energy is dissipated across the type-A PKWs and give the designers and practitioners collective input. To this end, twelve laboratory-scale replicas were tested and examined experimentally, and Eq. (1) was used to calculate the upstream and downstream specific energy over the PKW. The experimental investigations were recorded and analyzed for each PKW with various hydraulic properties, like the relative energy dissipation ($\Delta E/E_1$) as well as relative residual energy (E_2/E_1) withthe H_t/P and unit discharges (q) are depicted in Figs. 5 to 12.

As detailedin earlier sections, the PKWs' energy dissipation characteristics depend on various non-dimensionless parameters. According to Pralong et al. (2011), various geometrical parameters (around 22 parameters) related to the PKWs (shown in Fig. 3) are frequently used in the design analysis and are sufficient to provide comprehensive information about the PKWs. The 3D

structure of the PKW is complicated, and it is not possible to vary all the parameters at the same time, so keeping all these constraints in mind, the authors consider only three dimensionless parameters in the present study; Eq. (7)'s remaining parameters (i.e., S_i/S_o, B_i/B, B_o/B, *hs/ls & R/P*) were considered constant for all the models. So, the data of the current work was assessed in three modes: The first was to evaluate the impact of the inlet-outlet width (W_i/W_o) ratio on energy lossacross the PKW when maintaining the *L/W* ratio as well as the number of cycle counts (*N*) unchanged. The second aspectis to understandthe effect of key cyclic variation or the number of key cycles (*N*) on energy dissipation across PKW when maintaining the constant W_i/W_o and *L/W* ratios. The third consideration was: the effect on energy dissipation of varying the *L/W* ratio while maintaining a constant width (W_i/W_o) ratio and the cyclic count.

The head over the PKW fluctuated between 0.0168 m ≤ H_t≤ 0.167 m in the current study, where few depths are smaller than 0.03 m, therefore, the scalinginfluencewould be vital for these key values (Erpicum et al., 2016). The projects that fall under the purview of physical modeling must be designed with large-scale parameters in mind (physical small dimensional models). Parameters that are not scaled, includingwater viscosity, surface tension, and atmospheric pressure, may impact the flow characteristics of these models. Low overflow depths may cause scale effects due to fluid parameters like surface tension and viscosity that could not be scaled. Although total head over PKW as well as scaled-to-prototype (for 1:3 scale ratio) head thresholds where insignificant scale effects were found, according to Tullis et al. (2020a), were proposed to be $H_t/P \geq 0.09$. The authors didn't account for the scale effect in the present work. There are additional opportunities to conduct experimental research by considering the scaling effects.

Moreover, the observed flow field data were compared to previously published literature. Table 2 illustrates the MAPE, RMSE, and R^2 between the current research and those of Anderson and Tulli (2013) and Young (2018) were calculated and comparison of*Q* for a given value of H_t. Afterward, the results from ($\Delta E/E_1$) and (E_2/E_1) were compared with datasets from PKW by Eslinger and Crookston (2020), rectangular labyrinth weirs of *P* = 26 cm by Merkel et al. (2018), and trapezoidal labyrinth weirs of *L/W* = 2, 3, 4, 5, *P* = 15 cm by Magalhães and Lorena (1994). Table 2 shows that the current experimental study herein is in good agreement with the published data.

3.1 Effects of Relative Width Ratio (W_i/W_o) on Energy Dissipation

The first goal was to investigate energy dissipation by varying W_i/W_o ratios in 6 laboratory-scaled type-A PKW replicas (as shown in Table 1, PKW-1 to 6). The rating curves for each PKW were created, examined, and compared to previous research. Data from laboratory-scale trapezoidal and rectangular labyrinth weirs (Magalhães and Lorena, 1994; Merkel et al., 2018) and PKW (Eslinger and Crookston, 2020) are presented only for comparison and reference purposes. The relative energy dissipation patterns of all the models were almost identical. In this study, the energy dissipation rates for $H_t/P < 0.28$ (about for all models) were found to be higher than in previously reported research. The dissipation ($\Delta E/E_1$) rate between $0.28 \leq H_t/P \leq 0.45$ wasnoticed intermixing in nature, with $H_t/P > 0.45$ having a lower rate of relative energy dissipation (see Figs. 5(a), 5(b)).

PKW type-A is less efficient at dissipating energy in higher heads, whereas it is the most effective at lower head levels. When H_t/P is less than 0.18, even a little modification (in H_t/P) shows a noticeable variation in the energy dissipation. As H_t/P increases, the rate of ($\Delta E/E_1$) relative energy dissipation decreases steadily. At $H_t/P > 0.81$ results in constant relative energy dissipation, but between $0.18 \leq H_t/P \leq 0.81$, the relative energy dissipation rate ($\Delta E/E_1$) slowly reduces as H_t/P rises (see Figs. 5(a), 5(b)). At low head levels or low discharges, the PKW dissipates relative energy more effectively; its relative dissipation decreases as it grows in head levels or discharges. Similar effects were observed by (Eslinger and Crookston, 2020). This could be because of the model roughness and model approach conditions caused by turbulent boundary layer development and surface tension effects. These variations can impact the stage-discharge correlation, nappe course, and entrainment of air, according to Tullis et al. (2020a). As the discharges or heads increase, and the nappe flow changes from aerated to partially aerated and drown, the surface and viscous effect decreases; the energy loss provided by jet

Table 2. *Q* comparison for Corresponding H_t between the Present Study and Published Data by (Anderson and Tullis (2013) and Young (2018))

	Models	Anderson and Tullis (2013)				Models	Young (2018)		
S. No.	Type-A	MAPE	RMSE	R^2	S. No.	Type-A	MAPE	RMSE	R^2
1	PKW-1	2.50%	0.018	0.990	7	PKW-7	2.84%	0.016	0.997
2	PKW-2	---	--	--	8	PKW-8	2.32%	0.021	0.975
3	PKW-3	---	---	--	9	PKW-9	3.13%	0.024	0.995
4	PKW-4	--	--	--	10	PKW-10	3.90%	0.033	0.988
5	PKW-5	--	--	--	11	PKW-11	3.45%	0.022	0.986
6	PKW-6	3.59%	0.023	0.991	12	PKW-12	4.81%	0.032	0.991
					13	PKW-7 (Steeped)	3.15%	0.029	0.991
					14	PKW-10 (Steeped)	3.11%	0.033	0.986

[1]MAPE= Mean Absolute Percentage Error; [2]RSME = Root Mean Square Error; [3]R^2 = Coefficient of determination

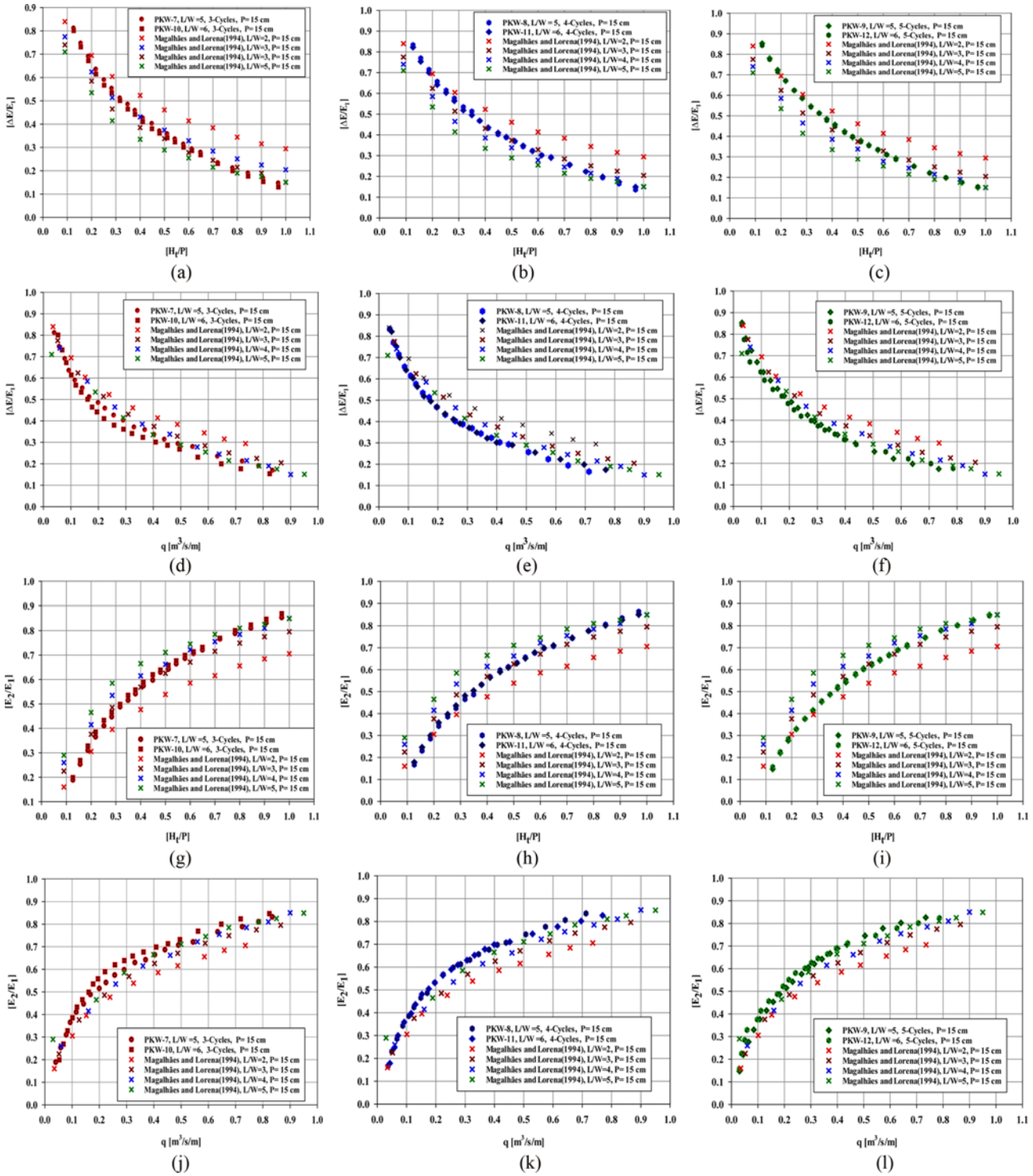


Fig. 6. Energy Dissipation of Flow Over Type-A PKW for Different Cycles (*N*): (a-f), and (g-l) Show the Variation of Relative Energy Dissipation ($\Delta E/E_1$) and Residual Energy (E_2/E_1) at PKWs Toe with Dimensionless Parameters H_f/P and the Specific Discharge *q*, respectively, for Different Cycles (*N*)

collision and downstream impact changes, resulting in less energy dissipation across the weir.

Figures 5(a), 5(b) and 5(c), 5(d) demonstrate the characteristics of the relative energy loss ($\Delta E/E_1$) and residual energy (E_2/E_1) at PKWs toe with dimensionless parameters H_f/P and the specific discharge q. Regardless of the W_i/W_o, the relative residual energy (E_2/E_1) at thebottom of the PKW senhances with H_f/P, especially for lower values. Figs. 5(a), 5(b) and 5(c), 5(d) show that the lowest value of W_i/W_o has the highest relative energy dissipation (i.e., $\Delta E/E_1$ = .841 or 84.10 percent of the related W_i/W_o = 1), which is 7.5% higher than the (Eslinger and Crookston, 2020) model for P = 0.20 m, W_i/W_o=1, and the highest value of W_i/W_o

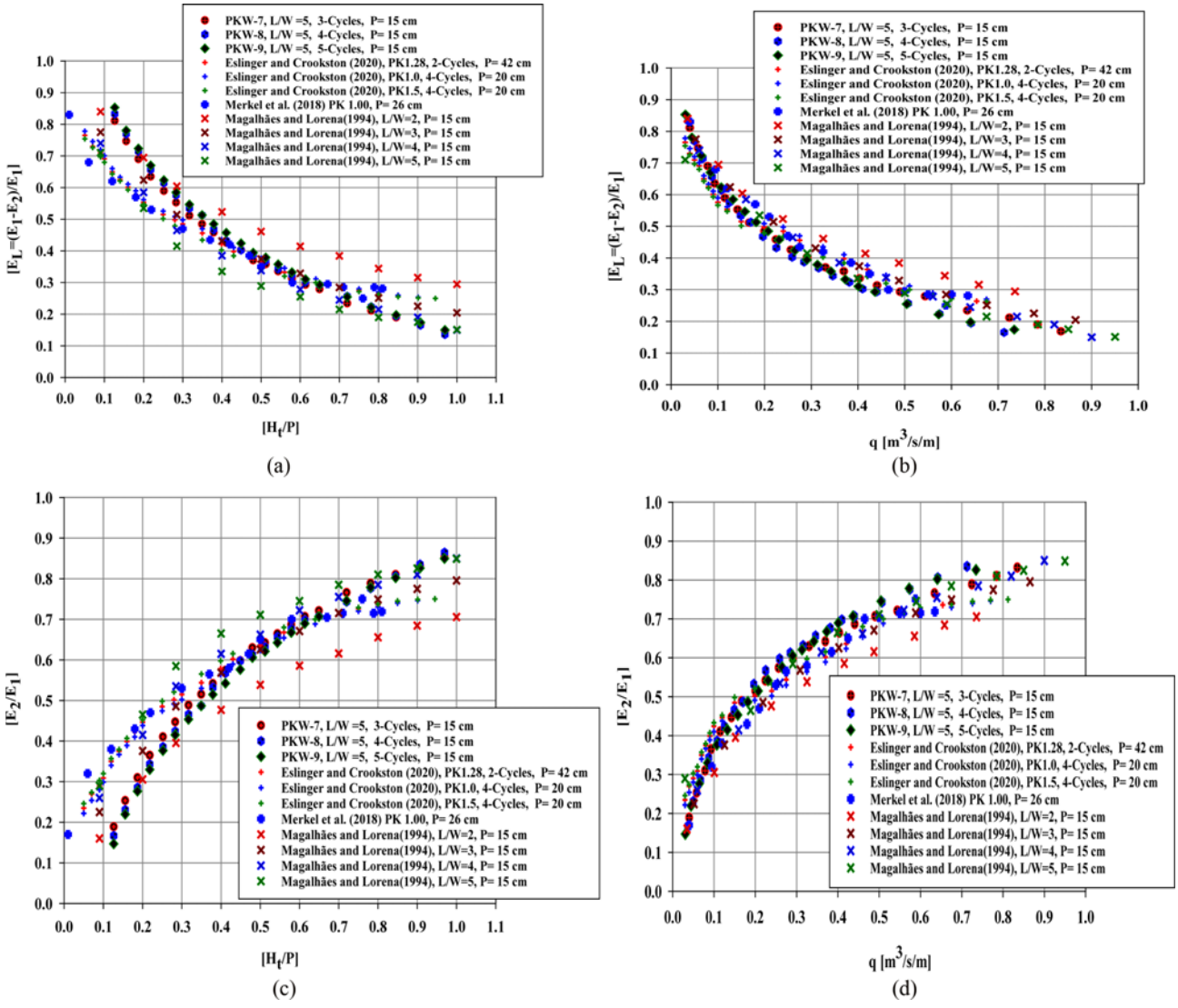


Fig. 7. Energy Dissipation of Flow over Type-A PKW for Different Cycles of Magnification Ratio $L/W = 5$: (a-b), and (c-d) Show the Variation of Relative Energy Dissipation ($\Delta E/E_1$) and Residual Energy (E_2/E_1) at PKWs Toe with Dimensionless Parameters H_t/P and the Specific Discharge q, respectively, for Different Cycles (for $L/W = 5$)

has the lowest energy loss ($\Delta E/E_1$ = .784 and 78.40 percent the related W_i/W_o = 1.5) which is 3.9% higher than the (Eslinger and Crookston, 2020) model for P = 0.20 m, W_i/W_o = 1.5. It means the lowest value of width ratio shows the highest energy dissipation and vice versa. This may be due to increasing the W_i value; head losses caused by flow passing the inlet key are reduced while the flow region is enhanced. A high W_i/W_o ratio in PKWs inlet keys optimizes flow approach and distribution, whereas a high W_i/W_o value in the outlet section promotes submergence influences. In the case of outlet sections, submergence effects may diminish the weir's discharge efficiency (areas where the weir crest height is higher than the flow depth in the outflow cycle).

3.2 Effects of the Number of Key Cycles (*N*) on Energy Dissipation

The second aspect of the investigationisdetermining how the key cycliccount influencesthe energy loss at the type-A PKW base. To this end, 6 PKW models (as shown in Table 1, PKW-7 to 12) were tested and assessed. The findings of the present section show that when the H_t/P ratio increases, the energy dissipation across type-A reduces. Furthermore, a detailed examination of Figs. 6 and 7 demonstrate that at low values of H_t/P (i.e., $H_t/P \leq$ 0.38 for both the L/W ratio), the relative energy dissipation rises with cycle count, while $H_t/P >$ 0.77 results in constant relative energy dissipation and between $0.38 \leq H_t/P \leq 0.77$, ($\Delta E/E_1$) ratesteadily drops as H_t/P rises. For a constant or limited width of the channel, the number of keycycles or units significantly affects the discharge capacity or flow characteristics of PKW. This portion of current research looked at the effect of cyclic influence on the dissipating energy at type-A PKWs. Increasing the cycle number reduces the hydraulic efficiency of the labyrinth family, which results in somewhat higher energy dissipation (Tulli et al., 2020b).

At PKW base, ($\Delta E/E_1$) rises with an increasing cyclic count with the same L/W ratio and the constant width ratio W_i/W_o. A downstream apex can be submerged by an interaction between nappes of adjacent weir walls, which reduces weir discharge efficiency (Crookston and Tulli, 2012). When the Labyrinth or PKW nappe settings change from aerated to moderately aerated, the head-discharge relationships for weirs of equal length but different cycle numbers can differ due to local submergence effects (Crookston and Tulli, 2013). Vertical drops and weirs have the same patterns in terms of energy dissipation and unit discharge (q) when comparing the two. Head-discharge relationships for labyrinth weirs of equivalent length but different cycle numbers can differ due to local submergence effects.

The findings suggest that the maximum cyclic count (i.e., $\Delta E/E_1$ = .8531 or 85.31 percent, corresponds to L/W = 5 and N = 5) has the highest relative energy dissipation. In comparison, the least keycyclic numbers ($\Delta E/E_1$ = .8010 and 80.10 percent correspond to L/W = 6, and N = 3) have the lowest relative energy dissipation (see Fig. 6). In the above discussion, PKWs depend on their key cycles (units) at $0.38 \leq H_t/P \leq 0.77$. Therefore, the cycle count is a convenient parameter.The current study's findings were compared to earlier reported data for trapezoidal labyrinth weirs (Magalhães and Lorena, 1994), and it was seen that all three models followed a similar pattern. However,

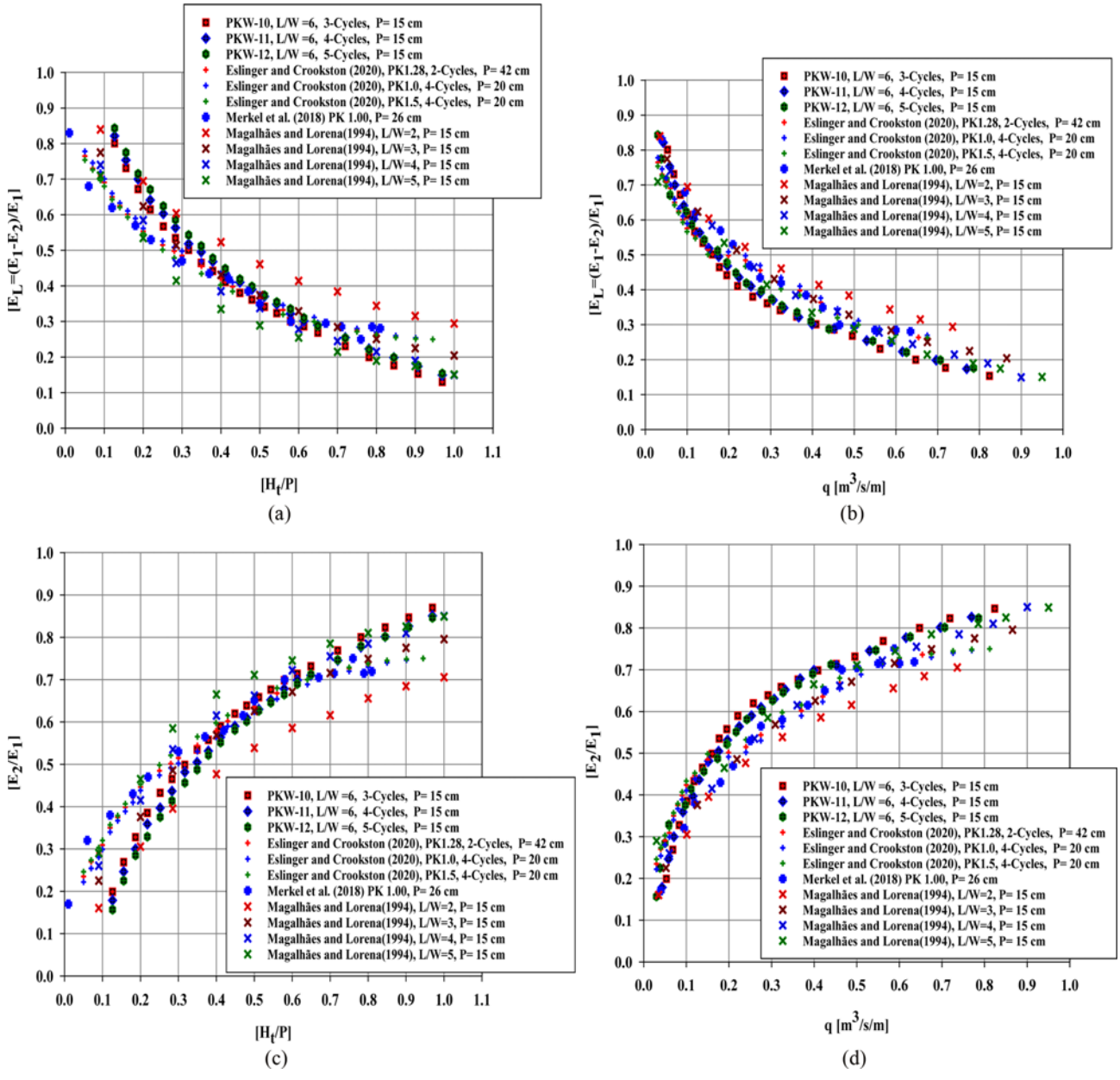


Fig. 8. Energy Dissipation of Flow over Type-A PKW for Different Cycles of Magnification Ratio L/W = 6: (a-b), and (c-d) Show the Variation of Relative Energy Dissipation ($\Delta E/E_1$) and Residual Energy (E_2/E_1) at PKWs Toe with Dimensionless Parameters H_t/P and the Specific Discharge q, respectively, for Different Cycles (for L/W = 6)

experimental uncertainties and different geometries can partly explain minor energy dissipation differences. Magalhães and Lorena (1994) also observed that the energy dissipation decreases as the H_t/P ratio increases. They suggested that the geometries greatly influence the hydraulic efficiency and energy dissipation across the labyrinth family.

3.3 The Influence of *L/W* Ratio on Energy Dissipation

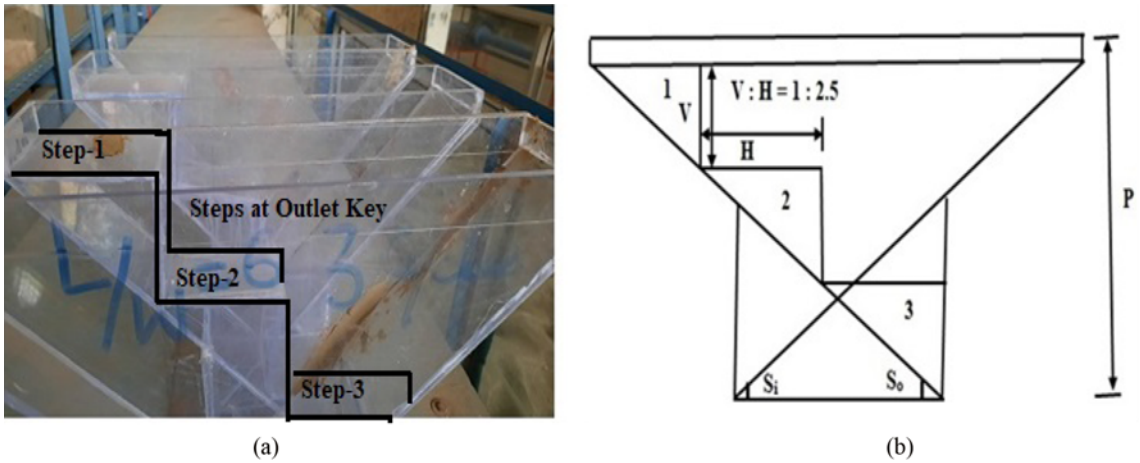
PKW discharge efficiency is significantly influenced by the *L/W* ratio, according to the literature. Leite Ribeiro et al. (2012) stated that PKW discharge carrying capacity increases as the *L/W* ratio increases. However, according to Magalhães and Lorena (1994), the magnification ratio (*L/W*) affects the folded shape weir's capability for displacing energy. The PKW's development crest length should be multiple times the weir's width for optimal discharge efficiency. By increasing the usable width *W*, the design effectively increases the crest length *L* (Laugier et al., 2017).

In additional insight regarding the impacts of varying *L/W* ratios on the energy assessment across type-A PKW, the test results of six PKW models (see Table 1, PKW-7 to 12) were examined. The present study results indicated that as the *L/W* ratio increases, the energy loss at PKWdrops. For labyrinth weirs with one-cycle trapezoidal crest profiles with different magnification ratios, (Magalhães and Lorena, 1994) observed similar effects. They concluded that the relative residual energy (E_2/E_1) at the base of labyrinth weirs rises as (*L/W*) ratio increases (i.e., with increasing (*L/W*) ratio, the relative energy loss at the bottom of the PKW decreases) (see Figs. 7 and 8). Further, Magalhães and Lorena (1994) reported the higher E_2/E_1 ratio for $L/W = 5$, whichcould be attributed to variations in *B* (the length of structures along the direction of flow), cycle geometry, and crest shape (overhanging profile). They also said that the geometry of the labyrinth weir has a considerable influence on the energy assessment.

The authors compared their results with already reported data to better understand the impacts of the relative width ratio, magnification ratio, and key cyclic influence on the energy across the PKW. The test results show that as the *L/W* ratio increases, the energy lossacross the type-A PKW decreases. Because the efficient crest length reduces with the rising head due to local submergence on the upstream apex, this approach's development length (*L*) changes with the head. With increasing head, the overcrossing occurs between the lateral jet and flow drowning, reducing PKW's dissipation efficiency.

The review has found that the energy loss over the type-A PKW is more significant than its alternatives, with the exception of $L/W = 2$ of the trapezoidal labyrinth weirs (Magalhães and Lorena, 1994), which follows the same trend. Figs. 7(a), 7(b) and 8(a), 8(b) reveal that the energy dissipation across type-A PKW was found to be greater than its alternatives for $H_t/P \leq 0.28$. However, the relative energy dissipation at $H_t/P > 0.6$ is lower than that of the trapezoidal labyrinth weirs tested by Magalhães and Lorena (1994), and the intermixing patterns were seen between $0.28 < H_t/P \leq 0.60$, butthe rate of energy dissipation follows the same trend as in previous studies. Figs. 7(a), 7(b) and 8(a), 8(b) depict that the relative energy dissipation ($\Delta E/E_1$) across the PKWs drops with the H_t/P, notably for lower H_t/P ratios, regardless of the magnification ratio (*L/W*). As illustrated in Figs. 7(c), 7(d) and 8(c), 8(d), the E_2/E_1 ratio at the PKWs bottomrises with (*L/W*) ratio, comparable to Magalhães and Lorena (1994). The energy losses of 0.853 and 0.14 correspond to relative residual energy at PKWs base ranging between 0.147 and 0.859. Consequently, particularly at low *Ht/P*, type-A PKWs dissipate energy very efficiently.

The *L/W* ratio of a PKW is animportant hydraulic parameter. Indeed, this ratio shows how well a design maximizes the crest length *L* by maximizing the available width *W*. In order to maximize discharge, the flow control section must be placed along the evolved crest length *L*. This indicates that the outlet should not be submerged due to flow conditions, and the inlet should not restrict the flow over the developed crest. The inlet section's sidewalls account for a large portion of the crest's length. Ideally, each inlet section should have a discharge capability equivalent to one of the developed crest length sections downstream. Outlet keys remove the stream's upstream flow and the inlet key's lateral overflow, so the performance of its outlet determines the PKW's efficiency.



(a) (b)

Fig. 9. Type-A PK Weir after Using the Steps at Outlet Key: (a) Laboratory Model, (b) Sectional View

4. Sensitivity Analysis

PKW's sensitivity was assessed using the steps instead of the sloping flooring (see Fig. 9). The stepped spillway dramatically improves energy dissipation, reduces the impact length of the flow on the downstream face of the dams, and reduces cavitation occurrences (Chanson, 1994; Parsaie et a., 2018). Hence, this investigation has focused on how the stepped outlet key affects PKW dissipation.

In order to assess the effects of using steps instead of a sloped floor at the outlet key area to improve the relative energy dissipation and, consequently, shorten the jet trajectories, two previously tested laboratory-scaled PKW models (see in Table 1, PKW-7 Stepped and PKW-10 Stepped) were updated with three uniform steps (all with a $1V$: $2.5H$ slope) at the outlet key, all other geometric configurations, on the other hand, remained unchanged. Thetesting heads varied between $0.12 \leq H_t/P \leq 0.98$, and all other dimensions are the same as discussed above.

A thorough understanding of flow behaviour across the PKW with and without steps is required for sensitivity analysis. According to Chanson (1994), the spillway face energy dissipation rate is significantly increased, and the demand for massive dissipators is eliminated or reduced. As per Parsaie et al. (2018), Froude number, depth to height ratio of the step, as well as drop number are essential factors influencing energy dissipation inflow across stepped spillways. Energy dissipation across the stepped spillway is influenced by the step's roughness (Torabi et al., 2018).

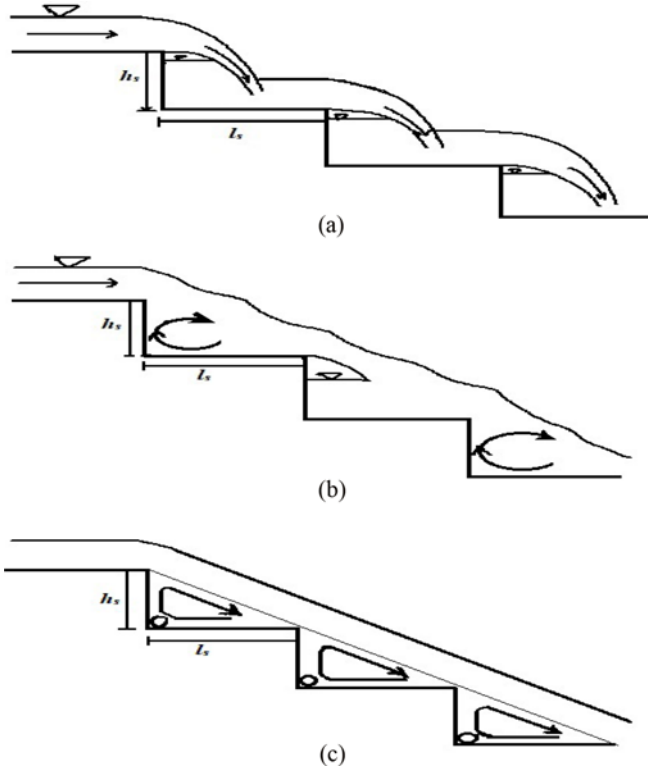


Fig. 10. Flow Regimes Above Stepped Weir: (a) Nappe Flow, (b) Transition Flow, (c) Skimming Flow

Following the observation, the three kinds of flow patterns via stepped spillways may be divided into napped, skimming, and transition flow (see Fig. 10). Low heads/discharges result in napped flow conditions. In this state, fluid moves from the top stage/step to the lower one (Fen et al., 2016). The clash of a flow jet with hydraulic leaps and steps, which can existentirely or partly, is the source of the energy loss. However, a skimming flow situation happens at high discharge levels, and in this case, pseudo-bottom forms between passing flow and steps. Keeping vortices under the quasi-bottom generated by the steps' exterior edges consumes higher energy in this situation. The vortices are reinforced by turbulent shear tensions between the recirculating fluid and the skimming stream (Chanson, 1996). It is critical to properly analyze the turbulent kinetic energy lossesover the steps, especially for high discharges related tothe skimming flow regime. Skimming flows are distinguished by free-surface aeration and high turbulence (Rajaratnam, 1990; Peyras et al., 1991; Chanson and Toombes, 2002; Gonzalez and Chanson, 2008). A transition regime is a state of flow that exists between skimming and napped flow.

In this investigation, the energy rating curves demonstrated asimilar pattern as without steps; however, steps exhibited more energy dissipation. Energy dissipation is greatly influenced by the slope and steps size of the stepped spillway (Parsaie and Haghiabi, 2019a, 2019b). Figs. 11(a), 11(b) and 11(c), 11(d) show a rating curves comparison with and without steps at the outlet key. The present study's data show that running water at the outlet key dissipates a significant amount of energy. The gain at a smaller H_t/P is about 5.60%, and 6.67%, whereas the decrease at a high H_t/P ratio is roughly 1.59 percentand 1.21 percent, respectively (see Fig. 11). At the low heads, water may flow down the steps as a freestream, skimming the constructed pseudo-bottom of the step edge. The free stream's shear stress transfer maintains the step cavities' turbulent recirculation. The free surface continually traps and releases the trapped air. At higher discharges or heads, the flow level in the outlet keys rises above the steps due to local submergence. Stepped weir capacity is reduced due to the complex air-water structures and energy dissipation processes associated with a two-phase mixture. As a result, using steps instead of sloping floors at the outlet key sections enhances the energy dissipation across PKW for low heads when reducing athigher heads. Thus, it may aid in designing energy-dissipative structures, reducing downstream scour, weir apron length, as well as hydraulic jump length.

5. Prediction of Estimation Equations

Lopes et al. (2006) made empirical predictions for the relative residual energy (E_2/E_1) downstream of the labyrinth weirs as a function of H_t/P and L/W based on an analysis of the Magalhães and Lorena (1994) experimental data. The results obtained from this study and past literature can be used for further model development. The authors elaborate on the current approach and

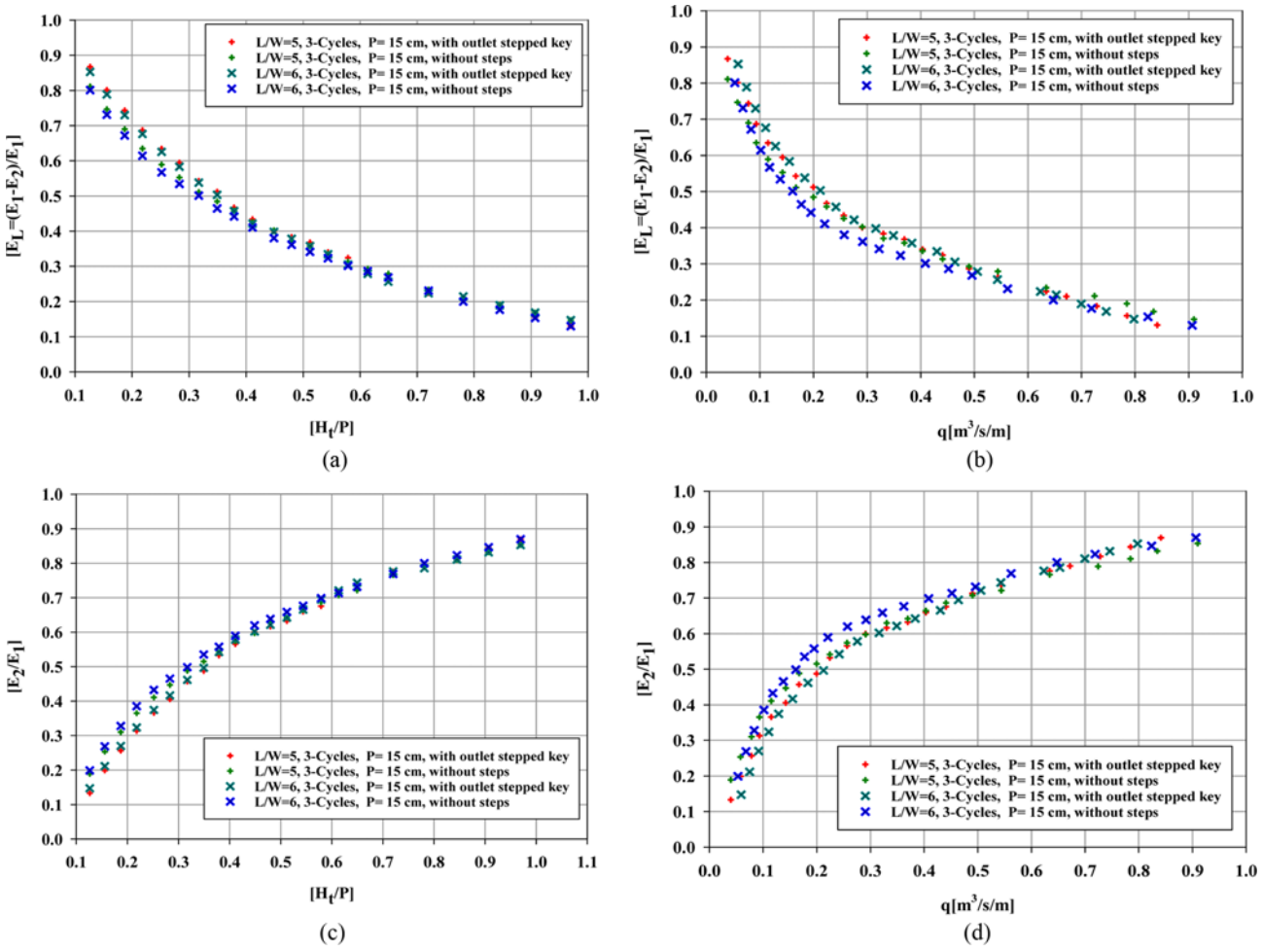


Fig. 11. Energy Dissipation of Flow over Type-A PKW with Using Steps at Out Let Key: (a-b), and (c-d) Show the Variation of Relative Energy Dissipation ($\Delta E/E_1$) and Residual Energy (E_2/E_1) at PKWs Toe with Dimensionless Parameters H_t/P and the Specific Discharge q, respectively, for Different L/W Ratios (for $N = 3$)

Table 3. Performance evaluation of Empirical Eqns. (8) and (9): (a) & (b) Coefficients Corresponding to Empirical Eqs. (8) and (9) Generated with the Help of Experimental Data [present study]

(a)

S. No.	Model geometry (PK weir Type-A)	a	b	c	MAPE	RMSE	R^2
1	PKW-1 (W_t/W_o = 1.0, 3 cycles)	0.9079	0.3205	0.001	2.0%	0.009	0.998
2	PKW-2 (W_t/W_o = 1.1, 3 cycles)	0.9308	0.3164	0.012	2.6%	0.017	0.996
3	PKW-3 (W_t/W_o = 1.2, 3 cycles)	0.9507	0.3263	0.010	2.5%	0.018	0.996
4	PKW-4 (W_t/W_o = 1.3, 3 cycles)	0.9282	0.3066	0.012	1.5%	0.013	0.996
5	PKW-5 (W_t/W_o = 1.4, 3 cycles)	0.9339	0.3016	0.007	1.5%	0.012	0.996
6	PKW-6 (W_t/W_o = 1.5, 3 cycles)	0.9575	0.3074	0.005	1.6%	0.012	0.996

(b)

S. No.	Model geometry (PK weir Type-A)	l	m	n	MAPE	RMSE	R^2
1	PKW-7 (L/W = 5, 3 cycles)	0.8660	0.3297	0.008	2.5%	0.013	0.999
2	PKW-8 (L/W = 5, 4 cycles)	0.8632	0.3424	0.025	1.8%	0.010	0.998
3	PKW-9 (L/W = 5, 5 cycles)	0.8573	0.3470	0.002	1.0%	0.005	0.999
4	PKW-10 (L/W = 6, 3 cycles)	0.8778	0.3268	0.005	1.7%	0.009	0.998
5	PKW-11 (L/W = 6, 4 cycles)	0.8557	0.3296	0.001	0.7%	0.004	0.999
6	PKW-12 (L/W = 6, 5 cycles)	0.8563	0.3433	0.005	1.8%	0.010	0.998

[1]MAPE = Mean Absolute Percentage Error; [2]RSME = Root Mean Square Error; [3]R^2 = Coefficient of determination

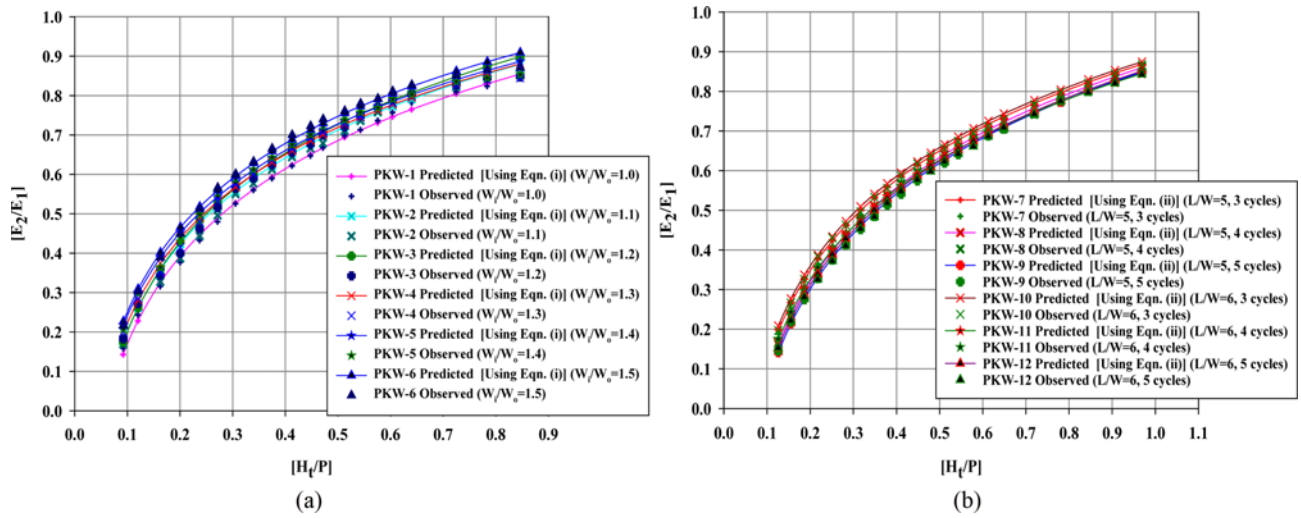


Fig. 12. Comparison Curve between Predicted and Observed Energy Values of Flow over Type-A PKW: (a) and (b) Observed and Predicted Values of Relative Residual Energy Using Eq. (8) and (9) and Tables 3(a) and 3(b), respectively

generate the prediction equations (Eqs. (8) and (9)) using the same concept. Thus, the first equation denotes the E_2/E_1 function of the (H_t/P) and (W_i/W_o) ratios.

$$\frac{E_2}{E_1} = a + b\,\ln\left[\frac{H_t}{P}\right] + c\,\ln\left[\frac{W_i}{W_o}\right] \tag{8}$$

whereas the second prediction equation (Eq. (9)) is the function of L/W and H_t/P.

$$\frac{E_2}{E_1} = l + m\,\ln\left[\frac{H_t}{P}\right] + n\,\ln\left[\frac{L}{W}\right] \tag{9}$$

The coefficients for Eqs. (8) and (9) are listed in Table 3, along with three statistical measures of MAPE, coefficient of determination (R^2), and RSME. Figs. 12(a) and 12(b) show the expected values of (E_2/E_1) using Eqs. (8) and (9), for each weir, the results were plotted against the original dataset along with the matching coefficients from Tables 3(a) and 3(b), respectively. The predicted relative residual energy (E_2/E_1) values consistently trend in the observed field data. The above equations subject to the present tests limitations 0.0168 m $\le H_t \le$ 0.167 m, 0.11 $\le H_t/P \le$ 1.0 for 0.15 m tall models, 0.085 $\le H_t/P \le$ 0.85 for 0.20 m tall models, 5 $\le L/W \le$ 6, 1.0 $\le W_i/W_o \le$ 1.5, 0.43 $\le (B_i/P = B_o/P) \le$ 0.95, P = 0.15 − 0.20 m, $B_i = B_o$ are alike for all models (see Table 1). The maximum MAPE between experimental and theoretical dataof the relative residual energy (E_2/E_1) is typically less than 2.6% for Eq. (8), and 2.5% for Eq. (9), implying that observed and predicted values are a good deal. It means the expected results are found to be within a small range of deviation from the actual ones. Hence, Eqs. (8) and (9) is a good tools to estimate the relative residual energy under the all free overflow condition.

6. Conclusions

A systematic experimental study was carried out to better understand energy dissipation at the base of the type-A PKW. To this end, twelve type-A PKW models were analyzed and evaluated. The present study's findings were compared to the energy dissipation downstream of the datasets for a rectangular, trapezoidal labyrinth and PKW of past literature, based on different hydraulic characteristics. To further understand how outlet key steps affect the type-A PKW's energy dissipation, a sensitivity analysis was performed by making the steps at the outlet key. A final set of first-order approximations to the author's empirical prediction equations for the best downstream configuration of the type-A PKW are provided. The following conclusions, observations, and recommendations are drawn from this study:

1. With a rising W_i/W_o ratio and magnification L/W ratio, energy dissipation across thetype-A PKWdiminishes. Thisenergy dissipation rate is nonlinear, and it is most recognizable at low heads. This may be due to the model roughness and model approach conditions caused by turbulent boundary layer development and surface tension effects.
2. The lowest width ratio (i.e., $\Delta E/E_1$ = .8621 or 86.21 percent the corresponding W_i/W_o = 1) had the greatest relative energy dissipation, while the largest width ratio ($\Delta E/E_1$ = .8077 or 80.77 percent the associated W_i/W_o = 1.5) had the lowest relative energy loss. As increase the W_i value; head losses caused by flow passing the inlet key are reduced while the flow area is enhanced, so overall dissipation efficiency decreases.
3. Low H_t/P (both L/W ratio) values increase relative energy dissipation, whereas higher H_t/P values result in intermixing patterns of energy dissipation.
4. The greatest cyclic count (i.e., $\Delta E/E_1$ = .8621 or 86.21 percent the corresponds to L/W = 5 & N = 5) resulted in the most relative energy dissipation, while the lowest cyclic number ($\Delta E/E_1$ = .8077 or 80.77 percent corresponds to L/W = 6 & N = 3) resulted in the least relative energy dissipation.
5. Lower H_t/P and higher L/W magnification ratios lead to higher

residual energy at the toe of PKW type-A (E_2/E_1), especially for lower values of H_t/P.

6. At low heads/discharges, the energy dissipation at type-A PKW base employing the steps at outlet key rather than the slanting floor improves (the gain is roughly 5.60 – 6.67 percent), whereas, at higher heads/discharges, it decreases (around -1.59 to -1.21 percent).
7. The proposed empirical equations imply that the maximum mean absolute percentage error between experimental and theoretical data is less than 2.6% for relative width ratio and 2.5% for the magnification ratio, showing that the observed and predicted values are good in the deal.

Overall, PKWs dissipate energy very efficiently, especially when the relative total upstream head is low. In addition, the PKW's shape substantially influences relative residual energy or energy dissipation at the weir's base. Consequently, it's vital to design and build energy-efficient hydraulic structures that limit scouring at the structures' downstream ends while increasing energy dissipation at the weir's base. These actual results are limited to the geometries that were tested, so more investigation into the energy loss downstream of the weir might be done with other PKW types. The study's contents, which include flow field measurements, a dataset, and the prediction residual energy equations, are intended to assist design engineers and researchers in building PKWs and determining downstream features such as energy dissipators and chutes. Additional research on energy dissipation and energy consumption over the PKW is recommended. This could include different geometries like PKW types B, C, and D, or it could use turbid water to calculate the energy surplus.

Acknowledgments

The authors would like to express their compassion and heartfelt appreciation to the Civil Engineering Department's faculty, staff, and the FM & HE laboratory technical team for their invaluable assistance. The authors would like to express their heartfelt gratitude and appreciation to Mr. Kewal Singh, a retired senior mechanic, for his practical assistance during the model fabrication process. This study was funded by the Delhi Technological University, India.

Nomenclature

B = Weir's length
B_b = Base length
B_i = Overhang length inlet portion
B_o = Overhang length outlet portion
E = Specific Energy
F_r = Froude Number
g = Gravitational acceleration
H_t = Total energy head
h_t = Piezometric head
h_s = Step's height
i = section
L = Total developed crest length
L_{cy} = Development ength of one cycle
l_s = Length of steps
$MAPE$ = Mean absolute percentage error
N = Key cycle's number
P = Height of the PKW
Q = Flow rate over PKW
R^2 = Coefficient of determination
R_e = Reynolds Number
$RSME$ = Root Mean Square Error
S_o = Outlet key slope
S_i = Inlet key slope
T_s = Sidewall thickness of PKW
V_t = Mean flow velocity
$V_t^2/2g$ = Approach velocity head
W = Width of PKW/Channel width
W_e = Weber Number
W_{cy} = Width of cycle
W_i = Inlet key width
W_o = Outlet key width
μ = Dynamic viscosity of flowing water
ρ = Liquid density
σ = Surface tension

ORCID

Deepak Singh https://orcid.org/0000-0001-9397-3220
Munendra Kumar https://orcid.org/0000-0003-3922-424X

References

Al-Shukur, Al-Khafaji GH (2018) Experimental study of the hydraulic performance of piano key weir. *International Journal of Energy and Environment* 9:63-70

Anderson RM, Tullis BP (2013) Piano Key Weir hydraulics and labyrinth weir comparison. *Journal of Irrigation and Drainage Engineering, ASCE* 139(3):246-253(2013), DOI: 10.1061/(ASCE)IR.1943-4774.0000530

Bahmanpouri F, Gualtieri C, Chanson H (2021) Flow patterns and free-surface dynamics in hydraulic jump on pebbled rough bed. Proceedings of the Institutionof Civil Engineers - Water Management, 1-47, DOI: 10.1680/jwama.20.00040

Bieri M, Federspiel M, Boillat JL (2011) Houdant, B, faramond, L, delorme, F energy dissipation downstream of piano key weirs — a case study of gloriettes dam (France). In Labyrinth and Piano Key Weirs — PKW 2011, CRC Press: Leiden, The Netherlands 120-130

Chanson H (1994) Hydraulics of nappe flow regime above stepped chutes and spillways. *Transactions of the Institution of Engineers, Australia. Civil Engineering* 36(1):69-76

Chanson H (1996) Prediction of the transition nappe/skimming flow on a stepped channel. *Journal of Hydraulic Research* 34(3):421-429, DOI: 10.1080/00221689609498490

Chanson H, Toombes L (2002) Experimental investigations of air entrainment in transition and skimming flows down a stepped chute. *Canadian Journal of Civil Engineering* 29(1):145-156, DOI: 10.1139/l01-084

Crookston B, Erpicum S, Tullis B, Laugier F (2019) Hydraulics of labyrinth and piano key weirs: 100 Years of prototype structures, advancements, and future research need. *Journal of Hydraulic Engineering ASCE* 145(12), DOI: 10.1061/(ASCE)HY.1943-7900.0001646

Crookston BM, Tullis BP (2012) Labyrinth weirs: Nappe interference and local submergence. *Journal of Irrigation and Drainage Engineering* 138:757-765, DOI: 10.1061/(ASCE)IR.1943-4774.0000466

Crookston BM, Tullis BP (2013) Hydraulic design and analysis of labyrinth weirs. II: Nappe aeration, instability, and vibration. *Journal of Irrigation and Drainage Engineering* 139:371-377, DOI: 10.1061/(ASCE)IR.1943-4774. 0000553

Erpicum S, Laugier F, Boillat JL, Pirotton M, Reverchon B, Schleiss AJ (2011) Labyrinth and Piano Key Weirs – PKW-2011. CRC Press, London, UK

Erpicum S, Laugier F, Ho Ta Khanh M, Pfister M, Labyrinth, Piano Key Weirs III – PKW-2017 (2017) CRC Press, London, U.K

Erpicum S, Laugier F, Pfister M, Pirotton M, Cicero G, Schleiss AJ (2013) Labyrinth and Piano Key Weirs II – PKW-2013. CRC Press, London, UK

Erpicum S, Tullis BP, Lodomez M, Archambeau P, Dewals BJ, Pirotton M (2016) Scale effects in physical piano key weir models. *Journal of Hydraulic Research* 54(6):692-698, DOI: 10.1080/00221686.2016.1211562

Eslinger KR, Crookston BM (2020) Energy dissipation of type a piano key weirs. *Water* 12(5):1253, MDPI, Basel, Switzerland, DOI: 10.3390/w12051253

Fen N, Kozlov DB, Rumyantsev IS (2016) Hydraulic studies of stepped spillways of various design. *Power Technology and Engineering* 49:337-344, DOI: 10.1007/s10749-016-0625-7

Ghaderi A, Daneshfaraz R, Dasineh M, Di Francesco S (2020) Energy dissipation and hydraulics of flow over trapezoidal-triangular labyrinth weirs. *Water (Switzerland)* 12(7), DOI: 10.3390/w12071992

Gonzalez CA, Chanson H (2008) Turbulence and cavity recirculation in air-water skimming flows. *Journal of Hydraulic Research, IAHR* 46(1):65-72, DOI: 10.1080/00221686.2008.9521843

Haghiabi AH, Ghaleh Nou MR, Parsaie A (2021) The energy dissipation of flow over the labyrinth weirs. *Alexandria Engineering Journal* THE AUTHORS, 0-4, DOI: 10.1016/j.aej.2021.08.075

Ho Ta Khanh M (2017) History and development of Piano Key Weirs in Vietnam from 2004 to 2016. In Erpicum, S, Laugier, F, Ho Ta Khanh, M, and Pfister, M. (2017). Labyrinth and Piano Key Weirs III - PKW 2017. CRC Press, London, U.K. 3-16

Ho Ta Khanh M, Sy Quat. D, Xuan Thuy D (2011a) PK weirs under design and construction in Vietnam. Labyrinth and piano key weirs-PKW 2011:225-232 (2011), CRC Press: London

Ho Ta Khanh M, Hien TC, Hai NT (2011b) Main results of the PK weir models test in Vietnam. (2004 to 2010). Labyrinth and piano key weirs-PKW 2011: 191-198, CRC Press: London (2011b)

Jüstrich S, Pfister M, Schleiss AJ (2016) Mobile riverbed scour downstream of a piano key weir. *Journal of Hydraulic Engineering* 142(11): 04016043:1-12, ASCEPress, DOI: 10.1061/(ASCE)HY.1943-7900.0001189

Karimi M, Attari J, Saneie M, Ghazizadeh MRJ (2018) Side weir flow characteristics: Comparison of piano key, labyrinth, and linear types. *Journal of Hydraulic Engineering* 144(12), DOI: 10.1061/(ASCE)HY.1943-7900.0001539

Kheilapour M, Mozaffari J, Asadolah S, Movahed M (2022) Experimental study of discharge coefficient of combination labyrinth weirs. *Journal of Hydraulic Structures* 7(4):1-9, DOI: 10.22055/jhs.2022.38073.1184

Kumar B, Ahmad Z (2022) Scour downstream of a piano key weir with and without a solid apron. *Journal of Irrigation and Drainage Engineering* 148(1):1-12, DOI: 10.1061/(ASCE)IR.1943-4774.0001647

Kumar B, Kadia S, Ahmad Z (2021) Sediment movement over Type-A piano key weirs. *Journal of Irrigation and Drainage Engineering* 147(6):04021018, DOI: 10.1061/(ASCE)IR.1943-4774.0001561

Laugier F, Vermeulen J, Blancher B (2017) Overview of design and construction of 11 piano key weir spillways developed in France by EDF from 2003 to 2006. In Erpicum, S., Laugier, F., Ho Ta Khanh, M., and Pfister, M. (2017). Labyrinth and Piano Key Weirs III - PKW 2017. CRC Press, London, U.K., 37-51 (2017)

Leite Ribeiro M, Boillat JL, Schleiss A, Laugier F, Albalat C (2007) Rehabilitation of St-Marc dam – Experimental Optimization of a Piano Key Weir. Proc. of 32nd Congress of IAHR., Vince, Italy

Leite Ribeiro M, Pfister M, Schleiss AJ, Boillat JL (2012) Hydraulic design of A-type Piano Key Weirs. *Journal of Hydraulic Research, Taylor & Francis* 50(4):400-408, DOI: 10.1080/00221686.2012.695041

Leite Ribeiro M, Boillat JL, Schleiss AJ (2011) Experimental parametric study for hydraulic design of PKWs. Proc. Int. Conf. Labyrinth and Piano Key Weirs, Liege B, 183-190, CRC Press, Boca Raton, FL

Lempérière F, Ouamane A (2003) The piano keys weir: A new cost-effective solution for spillways. *International Journal on Hydropower and Dams* 10(5):144-149

Lopes R, Matos J, Melo J (2006) Discharge capacity and residual energy of labyrinth weirs. In Hydraulic Model Report No. CH61/06, Proceedings of the International Junior Researcher and Engineer Workshop on Hydraulic Structures (IJREWHS '06), Montemor-o-Novo, Portugal, 2-4 September 2006, The University of Queensland: Brisbane, Australia 47-55

Magalhães AP, Lorena M, Perdas de Energiado do Escoamento Sobre Soleiras em Labirinto (1994) (Energy Losses in Flow Over Labyrinth Weirs), SILUSBA: Lisboa, Portugal, 203-211 (In Portuguese)

Merkel J, Belzner F, Gebhardt M, Thorenz C (2018) Energy dissipation downstream of labyrinth weirs. In Proceedings of the 7th IAHR International Symposium on Hydraulic Structures, Aachen, Germany 15-18

Mohammadzadeh-Habili J, Heidarpour M, Samiee S (2018) Study of energy dissipation and downstream flow regime of labyrinth weirs. *Iranian Journal of Science and Technology - Transactions of Civil Engineering, Springer International Publishing* 42(2):111-119, DOI: 10.1007/s40996-017-0088-6

Noui A, Ouamane A (2013) Study of optimization of the piano key weir. Labyrinth and piano key weirs II, CRC/Balkema, Leiden, Netherlands 175-182

Palermo M, Crookston BM, Pagliara SL (2020) Analysis of equilibrium morphologies downstream of a PKW structure. In Proceedings of the EWRI World Environmental and Water Resource Congress, Las Vegas, NV, USA 17-21

Parsaie A, Haghiabi AH (2019a) Inception point of flow aeration on quarter-circular crested stepped spillway. *Flow Measurement and Instrumentation* 69(August):101618, DOI: 10.1016/j.flowmeasinst.2019.101618

Parsaie A, Haghiabi AH (2019b) The hydraulic investigation of circular crested stepped spillway. *Flow Measurement and Instrumentation* 70(August):101624, DOI: 10.1016/j.flowmeasinst.2019.101624

Parsaie A, Haghiabi AH, Saneie M, Torabi H (2018) Prediction of energy dissipation of flow over stepped spillways using data-driven models. *Iranian Journal of Science and Technology - Transactions of Civil Engineering*, Springer International Publishing 42(1):39-53,

DOI: 10.1007/s40996-017-0060-5

Peyras L, Royet P, Degoutte G (1991) Ecoulement et dissipation sur les déversoirsengradins de gabions. *Journal La Houille Blanche* (1):37-47 (in French)

Pfister M, Justrich S, Schleiss AJ (2017) Toe scour formation at Piano Key weirs. In Erpicum, S, Laugier, F, Ho Ta Khanh, M, and Pfister, M Labyrinth and Piano Key Weirs III - PKW 2017 (2017). CRC Press: Leiden

Pralong J, Vermeulen J, Blancher B, Laugier F, Erpicum S, MacHiels O, Pirotton M, Boillat JL, Leite Ribeiro M, Schleiss AJ (2011) A naming convention for the Piano Key Weirs geometrical parameters. Labyrinth and Piano Key Weirs - Proceedings of the International Conference on Labyrinth and Piano Key Weirs, PKW 2011 271-278

Rajaratnam N (1990) Skimming flow in stepped spillways. *Journal of Hydraulic Engineering, ASCE* 116(4):587-591

Silverstri A, Archambeau P, Pirotton M, Dewals B, Erpicum S (2013) Comparative analysis of the energy dissipation on a stepped spillway downstream of a Piano Key weir. Labyrinth and piano key weirs II-PKW 2013:111-120. CRC Press, London

Singh D, Kumar M (2021) Hydraulic design and analysis of piano key weirs?: A Review. Arabian Journal for Science and Engineering, Springer Berlin Heidelberg, DOI: 10.1007/s13369-021-06370-4

Singh D, Kumar M (2022a) Energy dissipation of flow over the type-B piano key weir. *Flow Measurement and Instrumentation*, Elsevier Ltd, 83(November 2021):102109, DOI: 10.1016/j.flowmeasinst.2021.102109

Singh D, Kumar M (2022b) Gene expression programming for computing energy dissipation over type-B piano key weir. *Renewable Energy Focus, Elsevier Ltd* 41:230-235, DOI: 10.1016/j.ref.2022.03.005

Singh D, Kumar M (2022c) Study on aeration performance of different types of piano key weir. *Water Supply* 22(5):4810-4821, DOI: 10.2166/ws.2022.131

Singh D, Kumar M (2022d) Computation of energy dissipation across the type-A piano key weir by using gene expression programming technique Uncorrected Proof. *Watter Supply* 22(8):6715-6727, DOI: 10.2166/ws.2022.255

Torabi H, Parsaie A, Yonesi H, Mozafari E (2018) Energy dissipation on rough stepped spillways. *Iranian Journal of Science and Technology - Transactions of Civil Engineering, Springer International Publishing* 42(3):325-330, DOI: 10.1007/s40996-018-0092-5

Truong Chi H, Ho Ta Khanh M (2017) Experimental study for energy dissipation using stilling basin downstream of a P. K. weirs type A. Labyrinth and piano key weirs III-PKW 2017. CRC Press: Leiden

Tullis BP, Asce M, Crookston BM Asce M, Brislin J, Asce SM, Seamons T (2020b) Geometric effects on discharge relationships for labyrinth weirs. *Journal of Hydraulic Engineering* 146(10):1-10, DOI: 10.1061/(ASCE)HY.1943-7900.0001806

Tullis BP, Crookston BM, Young N, Asce M (2020a) Scale effects in free-flow nonlinear weir head-discharge relationships. *Journal of Hydraulic Engineering* 146(2):1-9, DOI: 10.1061/(ASCE)HY.1943-7900.0001661

Yazdi AM, Abbas Hoseini S, Nazari S, Amanian N (2021) Effects of weir geometry on scour development in the downstream of Piano Key Weirs. *Water Science and Technology: Water Supply* 21(1):289-298, DOI: 10.2166/ws.2020.272

Young NL (2018) Digital Commons @ USU Size-Scale Effects of Nonlinear Weir Hydraulics

Zhao C, Fang H, Liu Y, Dey S, He G (2020) Impact of particle shape on saltating mode of bedload transport sheared by turbulent flow. *Journal of Hydraulic Engineering* 146(5):04020034, DOI: 10.1061/(ASCE)HY.1943-7900.0001735

Journal of Physics: Conference Series

PAPER • OPEN ACCESS

Terahertz Wave Propagation Characteristics in Graded Teflon Based Solid-Core Photonic Crystal Fibre

To cite this article: Shikhar Dash *et al* 2023 *J. Phys.: Conf. Ser.* **2426** 012021

View the article online for updates and enhancements.

You may also like

- Design and simulation of liquid infiltrated photonic crystal fibre in terahertz frequencies
Izaddeen Yakasai, Pg Emeroylariffion Abas, Abdul Mu'iz Maidi et al.
- Photonic terahertz technology
Alvydas Lisauskas, Torsten Löffler and Hartmut G Roskos
- Design and analysis of low loss solid-core hexagonal photonic crystal fiber for applications in terahertz regime
K Renuka Rani and K Chitra

ECS 244th Electrochemical Society Meeting

October 8 – 12, 2023 • Gothenburg, Sweden

50 symposia in electrochemistry & solid state science

Abstract submission deadline:
April 7, 2023

Read the call for papers & **submit your abstract!**

This content was downloaded from IP address 14.139.251.98 on 02/03/2023 at 06:37

Terahertz Wave Propagation Characteristics in Graded Teflon Based Solid-Core Photonic Crystal Fibre

Shikhar Dash[1],Varun Mathur[1], Nilesh Pandey[1], Ravindra K. Sinha[2]

[1] Department of Applied Physics, Delhi Technological University, Bawana Road, Delhi-110042, India

[2] TIFAC-Center of Relevance and Excellence in Fiber Optics and Optical Communications Department of Applied Physics, Delhi Technological University, Bawana Road, Delhi –110 042, India

E-mail: shikhardash_2k18ep080@dtu.ac.in

Abstract. In the past few years, photonic crystal fibre (PCF) in the Terahertz spectrum is drawing significant attention because of its multifaceted applications in high-speed data communication, spectroscopy, sensing, etc. The proposed PCF structure is based on a Teflon solid core surrounded by porous cladding. Teflon has been used as a background material because of its high flexibility and the possibility of making longer PCFs, which is capable of guiding waves in the THz region. The porous cladding consists of circular air-holes positioned hexagonally whose diameters increase with each subsequent layer keeping the pitch constant thus providing a graded-index profile. The diameters of air holes in each hexagonal layer are 1 mm, 1.2 mm and 1.5 mm respectively, maintaining a constant pitch of 2mm and a perfectly matched layer (PML) thickness of 3mm. We have examined the transmission characteristics of the proposed profile in the frequency range of 0.8 THz to 2.5 THz. The modal solution of the profile is solved using the Finite Element Method based on COMSOL Multiphysics version 5.5. Using the graded-index profile, we have obtained a large effective modal area of around 5.6 mm^2 at 1 THz and a low confinement loss of 2.25 $\times$ 10^{-4} dB/km at the same frequency. Due to this large effective modal area, we get higher damage threshold and weaker non-linear effects, making it suitable for high power applications.

1. Introduction

Recent applications in the field of sensing, security, medical surgery and imaging have resulted in significant attention to Terahertz (THz) waves in these recent years. Terahertz frequency is a part of the electromagnetic spectrum that lies in between the microwave and the infrared region and is characterized by a frequency ranging from 0.1 THz to 10 THz. Wide range applications of THz wave can be attributed to its various properties, for example the property of penetration into fabrics and polymers are very useful for security and cleaning purposes, the non-ionizing nature finds application in medical scanning techniques and the higher bandwidth is suitable for high-speed communication over long distances. Along with the multitude of advantages, it still faces a significant challenge in its practical implementation as most of the systems use free space propagation. This results in high absorption loss in the atmosphere which leads to transmitter-receiver misalignment and consecutively information loss.

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

Now to overcome these difficulties, various solutions have been suggested in order to obtain low loss flexible waveguide such as metallic wires, metallic waveguides with dielectric coating for THz wave propagation . But nearly all of these suffer from high loss due to material absorption, bending losses and coupling during the wave propagation. So, a novel method of light guidance is required for further progress in this field.

In recent years, Photonic crystal fibre (PCF) [1, 2] in the THz regime have gained quite a significant attention. Photonic Crystal Fibre (PCF) provides several advantages over Single-Mode Fibre (SMF) and other previously suggested waveguides see ref [3], bearing low chromatic dispersion, low confinement loss and a high fraction of power confined to the core [4]. Moreover, they are much more easier to fabricate thus offering control over the optical properties by changing the core radius, number and size of air holes, pitch, geometry, etc.

Several different methods such as the multipole method, simple effective index method, radial effective index method [5] can be used to find the PCF modal solutions for various transverse profiles. However, the Finite Element Method (FEM) has a significant advantage over others. The FEM breaks down a large system into small parts known as finite elements using space discretization. The equations that model these finite elements are then collectively assembled into a large system of equations that formulates the entire problem.

There has been a significant progress in the study of the Terahertz wave propagation in PCF, see [6, 7, 8, 10], where extensive analysis is done on the transmission characteristics of solid-core PCF fibre using Teflon as the background material for getting low confinement losses below 10^{-7} cm^{-1} dB/km . Farhana Akter Mou et al. [7] have proposed a PCF structure that contain circular sectored cladding and a square-shaped core. This proposed geometry has confinement loss of about 10^{-15} cm^{-1} and an effective area close to 10^{-8} m^2. Rahman et al. [7] have employed the finite element technique to investigate the modal properties in a Topas based hollow core PCF. Alternatively, Md. Saiful Islam et al. [6] have proposed a zeonex based PCF structure having an array of elliptical-shaped air holes surrounded by porous cladding comprising a circular hole in a hexagonal arrangement. This structure bears a confinement loss of around 10^{-9} cm^{-1} dB/km and an effective mode area of around 105 μm^2.

In this paper, we will investigate a PCF structure comprising of Teflon as the background material with a solid core surrounded by porous cladding. Teflon fibre inherently provides us with low nonlinearity in the THz region and offers a feasible fabrication of solid core PCF with a large mode area. Morever, it provides us with a high damage threshold, making it suitable for high power application and as a medium to transport intense pulse in THz region as passive elements. The cladding consists of circular air holes arranged in a hexagonal structure with diameters increasing with each subsequent layer keeping the pitch constant, thus resulting in a graded profile. With these modifications in the profile, we were able to generate a large effective modal area along with a low confinement loss, thus making it useful for high power applications.

2. Physical structure and theoretical modelling

A 2-D cross-section of the employed PCF structure is shown below in Fig.1. The fibre design has been implemented using the Finite Element Method (FEM) [9]. In our work, we use the frequency domain FEM for modal analysis of the given profile. Under FEM, instead of solving the region as a whole, the area is divided into a finite number of elements to which the principal variations are applied. Here, we have subdivided the region into triangular elements with a Maximum element size = 0.01391, Minimum element size = 1.32×10^{-4}, Curvature factor = 0.6 and Maximum element growth rate = 1.5.

In the proposed fibre design (see Fig 1), we have used polytetrafluoroethylene (Teflon) as the background material because of its high flexibility [11] and easy access. This makes it suitable for long length fibres capable of guiding THz wave. Teflon could be easily fabricated into a low confinement loss profile , its nonlinearity is 100 times that of silica. The nonlinear refractive

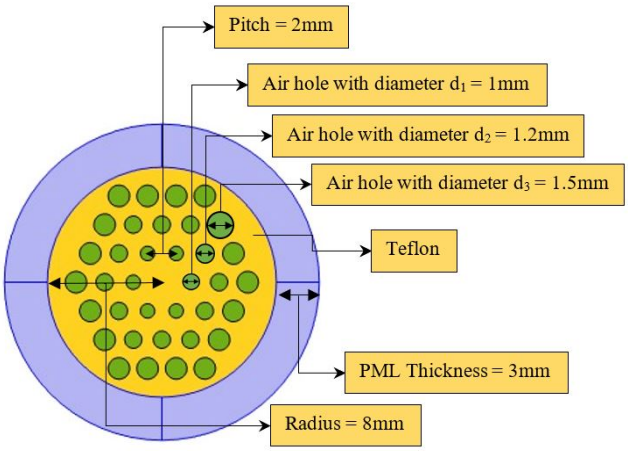


Figure 1. 2-D cross section of the proposed PCF

index of Teflon (n_2) is $8.3 \times 10^{-18}(m^2/W)$. The Refractive index of the bulk region is taken to be $1.445 + \text{j}0.00119$ for the frequency range between 0.8 THz to 2.5 THz. This ensures a low material dispersion in our fibre design.

The Teflon core is surrounded by a porous cladding which consists of circular air-holes positioned hexagonally [12], whose diameters increase with each subsequent layer keeping the pitch constant. This will lead to a graded-index profile. The diameters of air holes in each hexagonal layer are taken to be 1 mm, 1.2 mm and 1.5 mm respectively, thus maintaining a constant pitch of 2 mm and a perfectly matched layer (PML) of the thickness of 3 mm to prevent any surrounding effects like radiation or noise.

3. Numerical results and discussion

In this work, we present the modal solution of the proposed profile which has been solved using the Finite Element Method. Further, we analyse the mode profiles for different polarizations. We determine the effective index , dispersion and the Group Velocity Dispersion (GVD) indicating the important propagation characteristics. Also, we were able to figure out the confinement loss and the effective mode area for a solid core .Moreover, the behaviour of birefringence and non-linear coefficient for this graded PCF profile has been demonstrated for THz frequency region ranging from 0.8 THz to 2.5 THz.

In Fig 2, we see the electric field distribution of two fundamental mode profiles. One field distribution is polarized in the x-direction and the other in the y-direction.As we know that the proposed structure is entirely symmetrical in x and y direction, there is no significant contrast in their propagation characteristics along any direction. As mentioned above, the near indistinguishability of the x and y linearly polarized mode of our PCF allow us to can quantify this characteristic using the effective refractive index of these modes in the THz frequency range. The quantity that describes this polarization dependent propagation characteristic is termed as birefringence [13]. In Fig. 4, we have plotted the variation of birefringence with frequency in THz range. Birefringence (B) is defined as:

$$B = |n_x - n_y| \tag{1}$$

Where n_x and n_y represent the effective refractive index of the x and y polarization respectively. For this fibre model, we see in Fig.4 that the order of birefringence is very small and it decreases gradually with an increase in frequency. In our analysis, upon decreasing the frequency the mode expands outwards in the hexagonal porous cladding region. In comparison

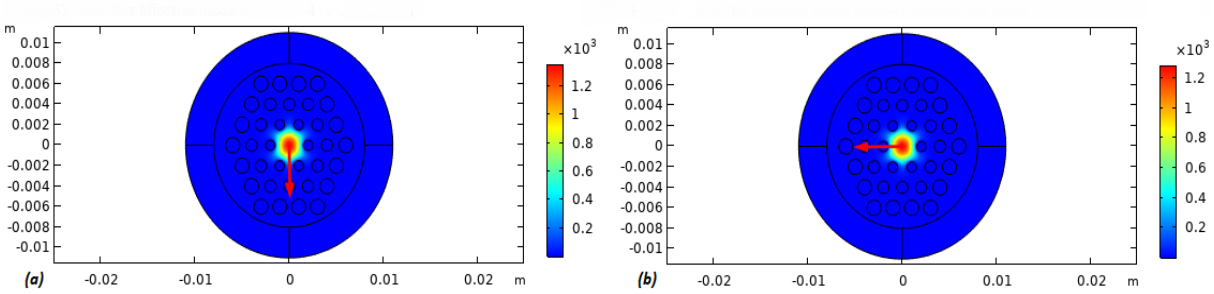


Figure 2. Cross sectional view of mode profile of y-polarized (a) and x- polarized (b) mode at 2.5 THz at n_{eff} of $1.44448 - j0.0011902$.

to core, the cladding is slighlty asymmetrical in the two orthogonal directions i.e., x and y directions. Hence, this results in a slight difference in the effective index in the cladding region as compared to the core. Thus, as we increase the frequency, the mode is more confined to the core region leading to a decrement in birefrigence. The case of low birefringent geometries and their importance are being explored to this date and this result could serve as a reference in the study of low birefringence polarization sustaining profiles [14].

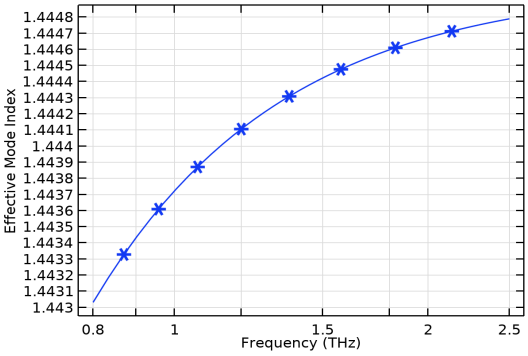


Figure 3. Variation of real part of the refractive index of the designed profile with frequency in the range 0.8 THz to 2.5 THz.

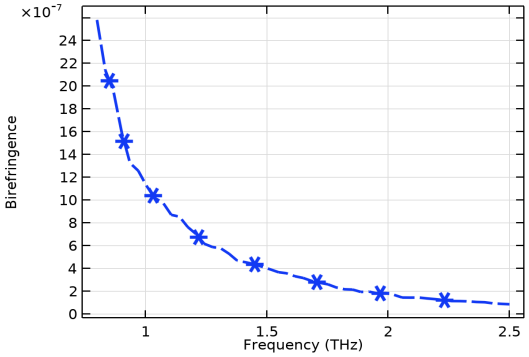


Figure 4. Variation of birefringence with frequency in the range 0.8 THz to 2.5 THz

Another important parameter for the transmission in the Terahertz frequency range is the confinement loss. The confinement loss [4] originates because of the imperfect fibre structure and the leaky modes that exists within the profile. A high confinement loss results in a short propagation length. This can be changed depending on parameters like the number of air holes , number of successive layers, air-hole diameter, and the pitch. Ideally, it should be small in magnitude which allows for a large propagation length in PCFs. Confinement loss is calculated using the following relation:

$$L_c = \left(\frac{4\pi f}{c}\right) Im\left(n_{eff}\right), \left[cm^{-1}\right] \tag{2}$$

Where, L_c represents the confinement loss, f is the operating frequency, c denotes the speed of light in a vacuum and $Im(n_{eff})$ is the imaginary part of the complex effective refractive index of the guided mode covered by the electric mode fields in the core region.For our case, we

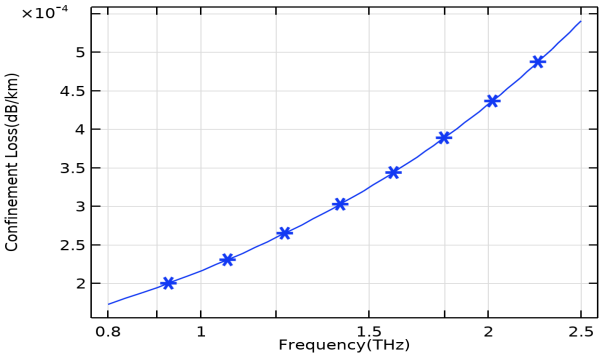


Figure 5. Variation of confinement loss (dB/km) is plotted with respect to increasing frequency in the THz region for the proposed profile.

introduce a hexagonal graded profile by gradually increasing air holes diameters in successive layers (Fig.1). Here, Fig 5, represents the variation of confinement loss in the THz frequency range.We find this loss to be around $10^{-3} dB/km$.

Optical pulses propagate through PCFs at a velocity often known as the group velocity. The relation elucidating the frequency dependence of the group velocity is called as group velocity dispersion [15]. We can see the variation of the group velocity dispersion in the THz frequency range in Fig 6. The graph indicates that our profile is in the normal region of operation since GVD decreases i.e., becomes more negative with an increase in frequency.

A useful characteristic of propagation is the chromatic dispersion of the profile. The sum of waveguide and material dispersion gives chromatic dispersion. The waveguide dispersion can be determined by the waveguide parameters and is easily variable in contrast with material dispersion. We have tried to reduce the dispersion by gradually increasing the air hole diameters in our hexagonal profile. Chromatic dispersion [16] is positive and is calculated using :

$$D(\lambda) = -\frac{\lambda}{c}\frac{d^2 Re\left[n_{yeff}\right]}{d\lambda^2} \tag{3}$$

Where, c is the velocity of light in a vacuum and λ is the operating wavelength. In Fig 7, the plot between dispersion and frequency has been shown. It can be seen that the dispersion decreases gradually from 0.028 dB/km at 0.8 THz to a minimum value of about 0.01 dB/km at 2.25 THz. After 2.25 THz, the dispersion begins to increase with frequency.

Another quantity of relevance is the effective modal area. PCFs with a large effective modal area find applications in high power transmission and also suppress the unwanted non-linear effects. The effective mode area is a measure of the transverse characteristics. A low value of the effective area imply large nonlinear effects and vice versa. It can be controlled by varying the structure size and air hole size. It is calculated using :

$$A_{eff} = \frac{\left[\int I(r) r dr\right]^2}{\left[I^2(r) dr\right]^2} \tag{4}$$

Where, $I(r) = [E(t)]^2$ is the intensity of electric field in transverse direction inside the fibre core. In this profile, we improvised the effective modal area for propagation taking place at less than 1 THz, as can be seen in Fig 8. In our study, we have taken our bulk refractive index to be constant for the same frequency range as studied by Chen Z Y et al. [11] i.e. 1.445 + j0.00119. From Eq (2), we can notice that the confinement loss varies directly with frequency and the imaginary part of effective index. For our case upon calculating the different mode profiles, we

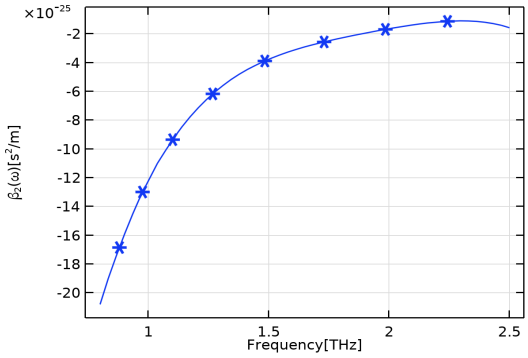


Figure 6. Variation of group velocity dispersion with frequency in the range 0.8 THz to 2.5 THz

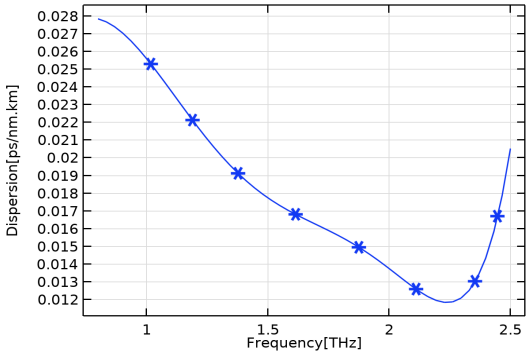


Figure 7. Variation of total dispersion of the proposed profile with frequency in the range 0.8 THz to 2.5 THz

realise that the imaginary part of effective index decreases very gradually with frequency. With the confluence of these variations, the net effect leads us to a confinement loss which gradually increases with an in frequency. Hence even though appropriate variation might not be possible with the above assumption it does provide us with a range in which our confinement loss lies. With these considerations, the mode is more or less confined to the core itself.

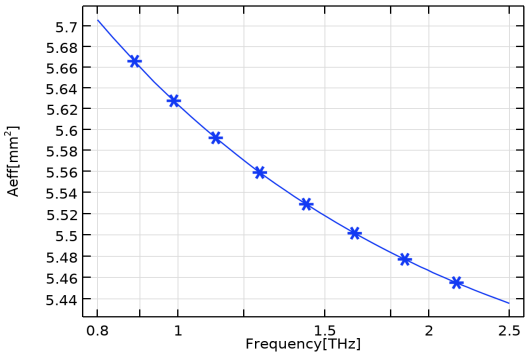


Figure 8. Variation of effective mode area of the proposed PCF with frequency ranging from 0.8 THz to 2.5 THz.

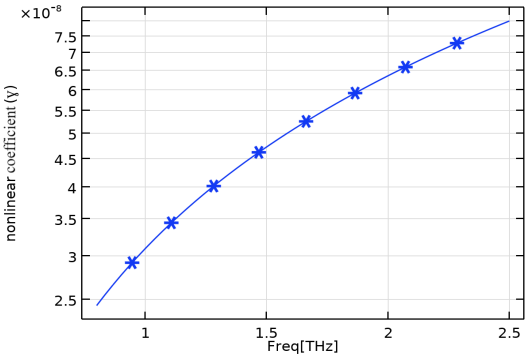


Figure 9. Variation of nonlinear coefficient with frequency in the range 0.8 THz to 2.5 THz.

Next, we discuss about the nonlinear effects of the proposed Teflon based PCF. Theoretically, Teflon has a bulk nonlinear refractive index n_2 of about 100 times higher as compared to silica [17]. This index is a measure of Kerr nonlinearity present in the profile. Also, nonlinear interaction also depends on the effective mode area of the profile. Higher the A_{eff} , the greater it helps in suppressing the nonlinear effects as in a structure. The above statement can be illustrated using the nonlinear coefficient γ. The nonlinear coefficient of a PCF is given by:

$$\gamma = \frac{2\pi n_2}{\lambda A_{eff}} \tag{5}$$

As can be seen above, the nonlinear coefficient is directly proportional to the nonlinear refractive index (n_2) and inversely proportional to the effective area (A_{eff}) [18]. Fig. 9 represents

the variation in nonlinear coefficient [17] with frequency for our solid core graded index Teflon based hexagonal PCF.The graph shows that the nonlinear coefficient reduces with an increase in frequency which is desirable for certain applications.

4. Conclusions

In this paper, we have implemented a Teflon based solid core PCF in a graded hexagonal manner. We also highlighted the propagation characteristics in the THz region. Results obtained have indicated an enhancement on effective mode index area and confinement loss for this profile over others [7] and [11].

Numerical and graphical results in this paper showed a high effective area of $5.62 \times 10^{-6}\ m^2$ at 1 THz, and a low non-linearity with a nonlinear coefficient of about 3×10^{-8} at 1 THz. Due to the proposed graded index profile, we were able to get a substantially low confinement loss of about $10^{-9}\ dB/cm$. Further, analysis on GVD and dispersion with varying frequencies in the THz region indicated a nearly dispersion flattened fibre with a dispersion of about around 0.025 [ps/nm km] at 1 THz. Owing to the symmetric nature of the profile, we got a low birefringence whose physical significance is yet to be explored. The fibre proposed and simulated (COMSOL V 5.5) in this paper could be used for making long enough PCFs as it comprises of Teflon as the bulk material. Additional advantages of obtaining a large effective area and low confinement losses make it suitable as a channel for high power transmission over long distances.

References

[1] Tajima K , Kurokawa K , Nakajima K , and Sankawa I 2005 *NTT Tech. Rev.* **3** 57-–61
[2] Meade R D V, Johnson S G and Winn J N 2008 *Photonic Crystals (Princeton University Press)* chapter 9 pp 156–189
[3] Wang K and Mittleman D M 2004 *Metal wires for terahertz wave guiding Nature* **432** 376–9
[4] Mia M B, Ani A B, Chowdhury K R, and Faisal M 2018 *2nd IEEE Int. Conf. Telecommun. Photonics* (ICTP) vol 2017-Decem, pp 10—13
[5] Chiang K S 1987 Radial effective-index method for the analysis of optical fibers *Appl. Opt.* **26** 2969
[6] Islam M S et al. 2018 *Opt. Mater. (Amst)* **79** 336—39
[7] Mou F A, Rahman M M, Islam M R, and Bhuiyan M I H 2020 *Sens. Bio-Sensing Res.* **29** 100346
[8] Han H, Park H, Cho M, and Kim J 2002 *Appl. Phys. Lett.* **80** 2634-–36
[9] Smitha T V , Madhura S, shreya N , and Udupa S 2021 *IRO J. Sustain. Wirel. Syst.* **3** 68-–86
[10] Bosu A, Basak C , Sharmin S, Hossain K M T, Urmi S Z and Mahfuz M Al 2019 *Int. Conf. Adv. Sci. Eng. Robot. Technol.* vol 2019 (ICASERT) pp 1—6
[11] Chen Z Y, Su H , Wang C.-D., Chen Q.-Z.,and Ruan S.-C. 2011 *Int. Symp.Photoelectron. Detect. Imaging 2011 Terahertz Wave Technol Appl.* vol 8195 p 81950Y
[12] Sinha R K and Varshney S K 2003 *Microw. Opt.Technol. Lett.* **37** 129-–132
[13] Agbemabiese P A and Akowuah E K 2020 *Sci. Rep.* **10** 1–12
[14] Schneider H, Harms H, Papp A, and Aulich H 1978 *Appl. Opt.* **17** 3035
[15] Mogilevtsev D, Birks T. A., and Russell P. S. J. 1998 *Opt. Lett.* **23** 1662
[16] Saitoh K , Koshiba M,Hasegawa T, and Sasaoka E 2003 *Opt. Express* **11** 843
[17] Vishwakarma R, Kumar P , Saha S, Rout G, and Roy J S 2017 *8th Int. Conf. Comput. Commun. Netw. Technol.* (ICCCNT) pp 3-–6
[18] Dabas B and Sinha R K 2010 *Opt. Commun.* **283** 1331–37

Physica Scripta

ACCEPTED MANUSCRIPT

Topological nature of large bulk band gap materials Sr_3Bi_2 and Ca_3Bi_2

To cite this article before publication: Sanjeev - *et al* 2023 *Phys. Scr.* in press https://doi.org/10.1088/1402-4896/acb7b3

Manuscript version: Accepted Manuscript

Accepted Manuscript is "the version of the article accepted for publication including all changes made as a result of the peer review process, and which may also include the addition to the article by IOP Publishing of a header, an article ID, a cover sheet and/or an 'Accepted Manuscript' watermark, but excluding any other editing, typesetting or other changes made by IOP Publishing and/or its licensors"

This Accepted Manuscript is © **2023 IOP Publishing Ltd**.

During the embargo period (the 12 month period from the publication of the Version of Record of this article), the Accepted Manuscript is fully protected by copyright and cannot be reused or reposted elsewhere.
As the Version of Record of this article is going to be / has been published on a subscription basis, this Accepted Manuscript is available for reuse under a CC BY-NC-ND 3.0 licence after the 12 month embargo period.

After the embargo period, everyone is permitted to use copy and redistribute this article for non-commercial purposes only, provided that they adhere to all the terms of the licence https://creativecommons.org/licences/by-nc-nd/3.0

Although reasonable endeavours have been taken to obtain all necessary permissions from third parties to include their copyrighted content within this article, their full citation and copyright line may not be present in this Accepted Manuscript version. Before using any content from this article, please refer to the Version of Record on IOPscience once published for full citation and copyright details, as permissions will likely be required. All third party content is fully copyright protected, unless specifically stated otherwise in the figure caption in the Version of Record.

View the article online for updates and enhancements.

This content was downloaded from IP address 14.139.251.98 on 06/02/2023 at 05:24

Topological Nature of Large Bulk Band Gap materials Sr_3Bi_2 and Ca_3Bi_2

Sanjeev[a], Mukhtiyar Singh[b] , Ramesh Kumar[a,*], Sunita Srivastava[c] and Tankeshwar Kumar[c]

[a] *Department of Physics, Guru Jambheshwar University of Science & Technology, Hisar-125001, India*

[b] *Department of Applied Physics, Delhi Technological University, Delhi-110042, India*

[c] *Department of Physics, Central University of Haryana, Jant-Pali, Mahendergarh-123031, India*

**Corresponding author: ramesh85@gjust.org*

Abstract

Topological materials are an emerging class of materials attracting the attention of the scientific community due to their potential applications in the fields of spintronics and quantum computing. Using first-principles calculations, the structural, electronic, and topological properties of Sr_3Bi_2 and Ca_3Bi_2 compounds without and with spin-orbit coupling are investigated. In the absence of spin-orbit coupling, the projected bulk band structure revealed that the Sr_3Bi_2 compound host a type-I Dirac point along the F-Γ direction. Since the compound possesses time-reversal and space-inversion symmetries, this Dirac point is associated with the nodal line. The existence of a type-I nodal ring around the Γ-point in the $k_z = 0$ planes, as well as a drumhead-like surface state within the nodal ring, suggested that Sr_3Bi_2 is a type-I nodal-line semimetal with no spin-orbit coupling. The inclusion of spin-orbit coupling introduced an energy gap of 0.36 eV between the valence band and conduction band at Dirac point. The topological surface states forming a Dirac cone between the bulk bandgap for (001) surface of Sr_3Bi_2 compound is calculated with spin-orbit coupling. The Z_2 topological invariants (1;000), as calculated by using parity product criteria, suggested that Sr_3Bi_2 is a strong topological insulator. Ca_3Bi_2, another compound with a similar crystal structure, is also predicted to behave similarly to Sr_3Bi_2 compound without and with spin-orbit coupling. This research broadens the application of topological insulators and existing platforms for developing novel spintronic devices.

Keywords: Topological Insulator, Spin-orbit coupling, Dirac point, Nodal line semimetal, Surface States, Density Functional Theory, Spintronics

1. Introduction

Topological quantum materials are novel materials and widely investigated fields in material science because of their rich scientific relevance and broad application possibilities [1-5]. Due to the effort of the scientific community, various classes of materials have been predicted theoretically and realized experimentally which include topological insulators (TIs) [1-4] [6, 7], Dirac semimetals (DSMs) [8-10], Weyl semimetals (WSMs) [11-13], nodal line semimetals (NLSMs) [14] [15] and topological crystalline insulators (TCIs) [16-18]. The TIs consist of bulk energy gap and exotic conducting surface states (SS) which are robust against external perturbations like structural disorders and nonmagnetic dopants [1-7]. These topological SS along with superconducting state or ferromagnetic state, make TIs useful in various potential applications such as quantum computing, spintronics, enhanced thermoelectric effects, optical recording, high-performance field-effect transistors, laser photonics and optoelectronics [19-21].

Besides TIs, topological semimetals (TSMs) are investigated in condensed matter physics and material science [5, 8-15]. Theoretically predicted exotic phenomena in TSMs include the anomalous Hall effect, negative longitudinal magnetoresistance, linear non-saturable magnetoresistance, very high carrier mobility and other exotic phenomena, which were experimentally observed in TSMs. TSMs are systems in which crossover of the conduction and valence bands in the Brillouin zone (BZ) is non-accidental, which means that it cannot be eliminated by perturbations on the Hamiltonian without destroying any of its symmetries. These TSMs are classified as WSMs and DSMs if the bands cross at discrete points, and NLSMs if the band crossing is along the closed curve [22, 23].

The two types of NLSMs are type-I and type-II, as determined by the slopes of crossing bands. In a type-I NLSM, the two bands cross along transverse axes with opposing slopes, but in a type-II NLSM, the two bands are slanted so that their slopes are identical along a single transverse axis. SOC usually

Accepted Manuscript

has an impact on these materials topological properties. Typically, SOC transforms nodal-line materials into other topological non-trivial phases like TIs [24], WSMs [25], DSMs [26], and even other varieties of nodal-line materials [27, 28].

The topological materials are characterised by the band-inversion, Z_2 topological invariants and conducting surface states. Band inversion is the basic feature of TIs which indicates the change in the orbital character of the conduction band (CB) and valence band (VB) of a material. For example, in the first generation of 3-dimensional TIs, Bi_2Se_3 or Bi_2Te_3, a topologically non-trivial state is because the p_z orbitals of Bi and Se (Te) have opposing parity at the Γ-point [29]. The first proven Dirac semimetal, Na_3Bi, exhibits an inversion of the Na 3s and Bi 6px, y bands [30]. Some ternary compounds, in addition to the binary ones, were also identified as topological materials, such as $TlBiSe_2$ family [31-33], $GeBi_2Te_4$ [34], $SnBi_2Te_4$ [35], $TaNiTe_5$ [36], $TaPdTe_5$ [37, 38] and so on. Experimentally, TIs can be verified by searching for the Dirac cone shape and gapless metallic SS in the surface band structure, which can be measured by angle-resolved photoemission spectroscopy (ARPES) and scanning tunnelling microscopy [39-42].

Bismuth (Bi) metal plays an important role in the topological material family due to the large SOC effect. It can donate as well as accept electrons to make a wide range of materials including insulators, semiconductors, or metals. Bi act as an electron donor (Bi^{3+}) in many TIs such as the Bi_2Se_3 family [29, 43, 44], $TlBiSe_2$ family [31-33] and BiTeX (X = I, Br, Cl) [45-50] due to the presence of high electronegativity elements like I, Br, Cl, F, Te, Se, or S which easily accept the electrons from Bi element to empty the p-orbitals of Bi atom. In some compound, Bi act as an electron acceptor (Bi^{3-}) such as the NaCaBi family [51-53], Mg_3Sb_2 [54], Mg_3Bi_2 [55, 56] and XX'Bi (X = Na, K, Rb; X' = Ca, Sr) [57] due to the presence of very low electronegativity element like alkali metals and alkaline-earth elements which easily donate the electrons to Bi atom to form fully filled p-orbitals. Using the compositional and structural searches of M-Bi (M = Ca, Sr and Ba) binary systems, it was found that the Sr_3Bi_2 compound crystallizes in the hR45 structure at its ground state [58].

In this work, the calculated bulk band structure using the first-principles calculation predicted that Sr_3Bi_2 is a semimetal in the absence of SOC. The presence of band-inversion between opposite parity eigenvalue bands near the Fermi-level at the Γ-point, a type-I nodal ring around the Γ-point in the $k_z = 0$ planes, and a drumhead-like surface state inside the nodal ring all suggested that Sr_3Bi_2 is a type-I NLSM without SOC. The SOC opens the large energy gap (~0.36 eV) at the line nodes, but retains the band inversion, resulting in a topological insulator-like band structure. In addition, the surface states are calculated for the (001) surface of Sr_3Bi_2 compound. Lastly, the parity product criteria [59] were used to calculate the value of Z_2 topological invariants (1,000), which further confirms that Sr_3Bi_2 is a topological insulator. Ca_3Bi_2, another compound with a similar crystal structure, are also predicted to behave similarly to Sr_3Bi_2 compound in the absence and presence of SOC as presented in the appendixes.

2. Computational Details

The structural and electronic properties of the binary compound Sr_3Bi_2 have been computed by using the density functional theory method as implemented in the Quantum-Espresso software [60, 61]. The GGA approximation in the form of Perdew-Burke-Ernzerhof (PBE) functional was utilized for exchange-correlation energy [62]. The electron-ion interactions were included by using a Norm-Conserving pseudopotential. The cutoff energy for wavefunctions was set at 60 Ry with a cutoff charge density of 240 Ry, to truncate the plane-wave basis set expansion of wavefunctions. The Γ-centered Brillion Zone (BZ) with an 8x8x8 Monkhorst-Pack grid of k-points for all calculations is employed. Further, the variable cell relaxation calculations were carried out to determine the relaxed cell parameter and to minimize the forces between the atoms by using the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm. Already optimized structure with scalar relativistic pseudopotentials as the initial structure was taken for including the influence of SOC. The force on atoms was further minimized by using the full relativistic pseudopotential with SOC. The maximally localized Wannier functions were constructed by employing the Wannier90 software [63] and the surface band structure was computed with the help of the wannier_tools code utility [64].

Accepted Manuscript

3. Results and discussion

3.1. Crystal structure of Sr_3Bi_2

The dynamical stability of Sr_3Bi_2 material with hR45 structure in equilibrium conditions are checked by calculating the cohesive energy (E_{coh}) per atom using the following formula,

$$E_{coh} = [E_{tot} - (n_{Bi}E_{Bi} + n_{Sr}E_{Sr})] / n_{tot} \quad \ldots (1)$$

where E_{tot} gives the total energy of the Sr_3Bi_2 system. E_{Bi} and E_{Sr} stand for the energies of single isolated Bi and Sr atoms, respectively. n_{Bi}, n_{Sr} and n_{tot} are the numbers of atoms of Bi, Sr and total number of atoms in Sr_3Bi_2 system. The calculated cohesive energy per atom is negative, i.e., -3.42 eV/atom, which means Sr_3Bi_2 is stable in the hR45 configuration as mentioned in Fig. 1 (a) and (b).

Further, Mg_3Sb_2 [54] and Mg_3Bi_2 [55, 56] were also reported with a hexagonal structure having space group P3-m1 (No. 164). The cohesive energy per atom in this structure for Sr_3Bi_2 system are also calculated which comes out to be positive i.e., + 9.65 eV per atom. Thus, Sr_3Bi_2 is unstable in the P3-m1 hexagonal phase.

The present work used the rhombohedral crystal structure of Sr_3Bi_2 with space group R3- (No. 148) for all types of calculations. The optimized primitive unit cell for Sr_3Bi_2 with the optimized value of lattice parameter a = 9.1270 Å and rhombohedral angle α = 66.146° is shown in Fig. 1(b). In the relaxed structure, the Bi atom is situated at 6f site (0.09164, 0.44523, 0.70242), Sr1 atom at 1b site (0.5, 0.5, 0.5), Sr2 atom at 2c site (0.31983, 0.31983, 0.31983) and Sr3 atom at 2f site (0.37394, 0.08409, 0.81048). These results are in good comparison with the previously reported work [58]. Structural details of Ca_3Bi_2 material are given in appendix A.

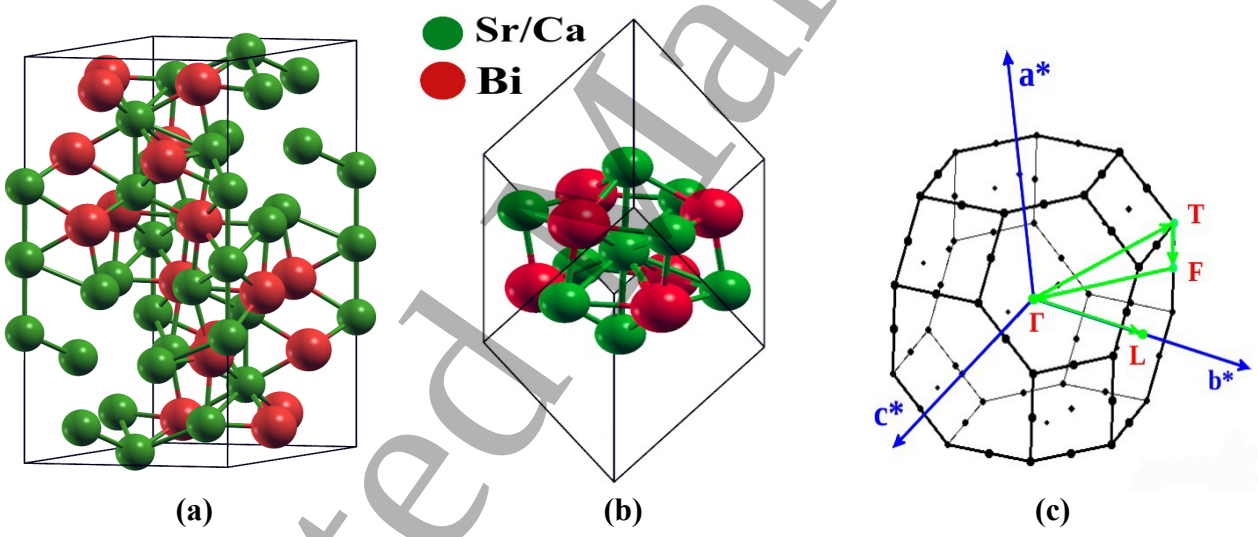


Figure 1: The hexagonal unit cell (a) and rhombohedral primitive cell of Sr_3Bi_2. (c) The first Brillion zone with high-symmetric points along which band structures are plotted.

3.2. Electronic properties of Sr_3Bi_2

The partial density of states (PDOS) and the projected bulk band structure along the momentum space path Γ–T–F–Γ–L are computed for studying the electronic properties without and with SOC of the Sr_3Bi_2. The projected bulk band structure of Sr_3Bi_2 compound in the absence of SOC is shown in figure 2a. By inspecting figure 2a, the VB and CB are crossed along the F-Γ direction, but well separated in the vicinity of Fermi-level at all other k-points in momentum space. Thus, the band structure of this material possessed semimetallic feature. Further, it is clearly observed that the Sr_3Bi_2 exhibits a band anticrossing feature at the Γ-point, with CB minima which is mainly due to the Bi-p orbitals (shown by red-coloured dots) and VB maxima, is mainly due to the Sr-s orbitals (shown by cyan-coloured dots). This band anticrossing feature indicates that there is a band-inversion between the s orbital of Sr and p orbitals of Bi atom due to the crystal-field splitting, which predicts the topological non-trivial nature of Sr_3Bi_2 compound. The band crossing in the F-Γ direction is gapless,

Accepted Manuscript

and the two bands around the crossing point have the opposite slope sign, which is characteristic of a type-I Dirac point. In contrast, a small gap appears between the two bands along the Γ-T direction. In general, in a spinless solid-state system having both space-inversion and time-reversal symmetries, the node points must form a 1D line as opposed to discrete points because the spinless Hamiltonian of two bands can always be assumed to have real-valued coefficients and the dimension of nodal solutions in 3D momentum space is 1. Therefore, in spinless Sr_3Bi_2 there must exist a nodal line going through the band crossing point. In figure 2b, the certain nodal points form a 1D ring around the Γ-point and lie precisely in the $k_z = 0$ planes. Usually, topological nodal line materials showed drumhead surface states on either side of the projected nodal line. To verify the existence of drumhead surface states, the band spectrum of the (001) surface of a semi-infinite Sr_3Bi_2 slab is computed. A Sr atomic layer or a Bi atomic layer could be used to terminate the (001) surface. In both ways, the drumhead-like surface states evolved inside the nodal line, as shown in figures 2 (c, d). Hence, Sr_3Bi_2 is a type-I nodal-line semimetal in the absence of SOC.

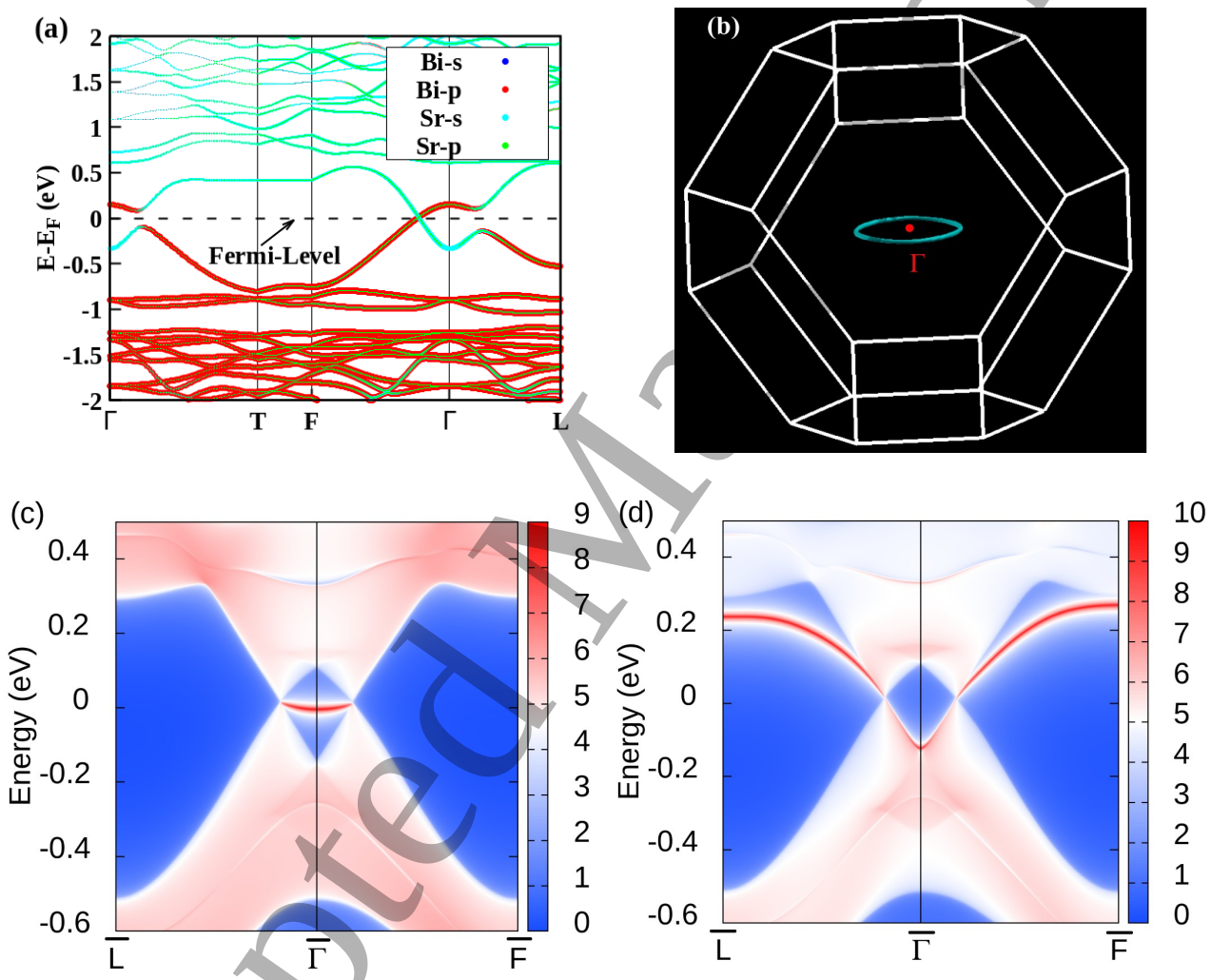


Figure 2: (a) Projected bulk band structure of Sr_3Bi_2 without SOC. (b) Illustration of type-I nodal ring in the bulk BZ. Computed (001) surface band structures of Sr_3Bi_2 for (c) Bi-terminated and (d) Sr-terminated surface without SOC.

The inverted bandgap value at Γ-point is approximately 0.48 eV. The inverted bands below and above the Fermi-level at Γ-point belong to the A_g and A_u irreducible representations (IRs) of the C_i (-1) point group, respectively. These IRs A_g and A_u are symmetric to the identity operator E. However, A_g is symmetric and A_u is anti-symmetric to the reflection through the centre of inversion. By inspecting the convention of IRs of C_i (-1) point group on Crystallographic Server notations at the Γ-point, these two bands belong to the Γ_1^+ and Γ_1^- notations are found, respectively. Thus, it is

Accepted Manuscript

concluded that the band-inversion occurs between the two opposite parity bands in the vicinity of Γ-point, which confirmed the topological nature of Sr_3Bi_2 compound without SOC.

The spinful case is studied by including the SOC for computing the electronic bulk band structure of Sr_3Bi_2 as given in figure 3a. Now the bulk nodal ring is separated, resulting in a continuous separation between the CB and VB at all k-points in the momentum space. However, it does not affect the band inversion in the vicinity of Γ-point. Thus, the topological nature of Sr_3Bi_2 compound is preserved due to the persistence of band-inversion even in the presence of SOC. But the energy gap between the inverted bands at Γ-point becomes 0.67 eV and the direct bandgap between the CB minima and VB maxima are now 0.36 eV which is large than the thermal energy at room temperature. After the band symmetry analysis of double point group C_i (-1), energy band just below and above the Fermi-level is represented by the Γ_2^+ and Γ_2^- notations, respectively. Based on band-inversion, Sr_3Bi_2 is a TI with a direct band gap of 0.36 eV and the inverted bandgap at Γ-point of approximately 0.67 eV. The corresponding projected band structure and (001) surface band structure of Ca_3Bi_2 are given in appendix B.

Exotic conducting topological SS are another significant property of the topologically non-trivial phase of a material. These SS serve as footprints to examine the topological non-trivial nature of a material. Thus, the topological SS of Sr_3Bi_2 compound for (001) surface is computed to enhance the understanding of topological non-trivial nature of the compound with SOC. Firstly, the maximally localized Wannier functions (MLWFs) are derived by choosing s and p orbitals of the Bi atom and s orbital of the Sr atom. Further, a tight-binding Hamiltonian is formed by employing these MLWFs with the help of Wannier90 software. The band structure obtained by this tight-binding model based on MLWFs exactly matches the one previously calculated using the DFT approach, confirming the correctness of the tight-binding Hamiltonian.

These findings show that the (001) surface have conducting topological SS that cross each other to form a single Dirac cone as shown in figure 3b. The surface carriers spins are locked to the direction of its momentum, similar to gapless edge states. These SS are found in the insulating bulk bandgap between the CB and VB and shows the Dirac-like dispersion. The presence of topological SS is a significant attribute of a TIs which makes them conducting at the surface. It not only provides solid proof of nontrivial topology, but it also acts as a direct link to the experiments. Thus, Sr_3Bi_2 behaves as a STI at ambient conditions. The calculated topological SS of Ca_3Bi_2 for (001) surface is shown in the figure 6b in appendix B. The presence of Dirac cone formed by the gapless SS confirmed the TIs nature of Ca_3Bi_2 compound.

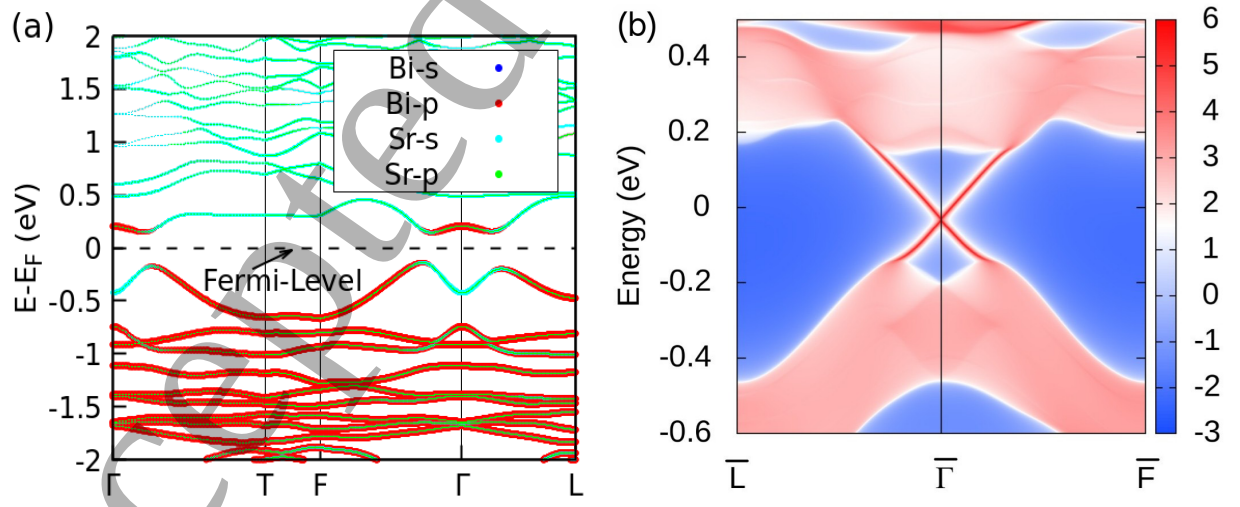


Figure 3: (a) Projected bulk band structure of Sr_3Bi_2 with SOC. (b) Gapless surface states of Sr_3Bi_2 terminated in the (001) surface. The (001) surface contains a 2D Dirac cone at Γ-point.

To understand the formation of bands and insulating nature of Sr_3Bi_2 compound, the orbital contributions of Bi-s, p, d and Sr-s, p states to the total density of states without SOC in the PDOS plot are projected as shown in Figure 4a. The Bi-p states dominate the states below the Fermi-level

Accepted Manuscript

(known as VB). However, Sr-s states play an important role in the lower portion of VB. The states just above the Fermi-level (known as CB) are dominated by the Bi-p and after that by the Sr-s states with a weak hybridization between them. After the inclusion of SOC, the gap between VB and CB is open near the Fermi-level as shown in figure 4b. The orbital contribution in the formation of VB remains almost equal as in the case of without SOC but in the CB, the hybridization between Bi-p and Sr-s states becomes strong.

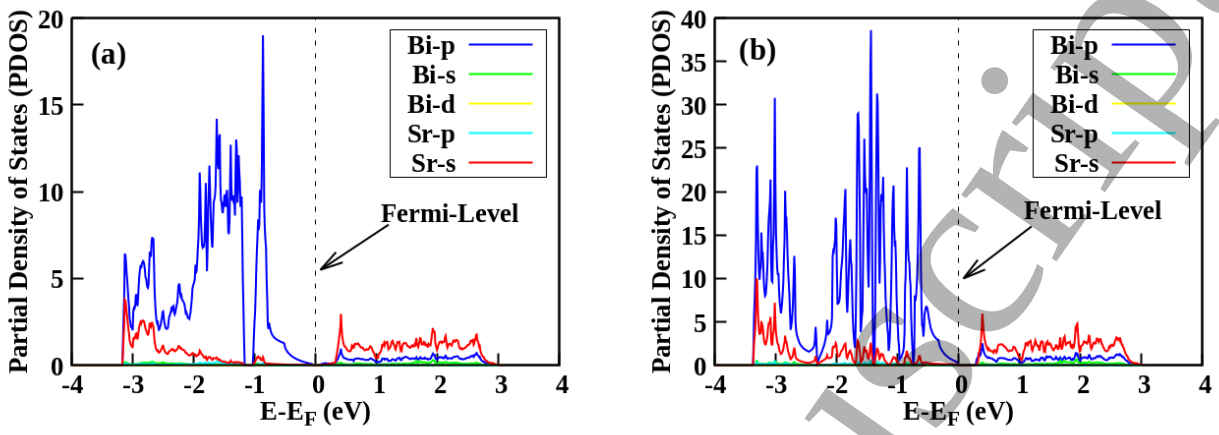


Figure 4: PDOS of Sr_3Bi_2 compounds (a) without SOC and (b) with SOC.

3.3. Parity Analysis of Sr_3Bi_2

Crystal structure of Sr_3Bi_2 possesses the inversion symmetry along with the identity symmetry operation. If any structure possesses inversion symmetry then the parity product criteria as proposed by Fu and Kane [59], can be used to calculate the Z_2 topological invariants, ν_0; (ν_1,ν_2,ν_3). The BZ of 3D system has eight time-reversal invariant momenta (TRIM) points which are obtained by the following formula,

$$\Gamma_{i=(n_1n_2n_3)} = \frac{1}{2}(n_1G_1 + n_2G_2 + n_3G_3) \quad (2),$$

where G_j are primitive reciprocal lattice vectors in k-space with $n_j = 0$, or 1. In case of Sr_3Bi_2 system, these eight TRIMs are 1Γ, 3L, 3F and 1T having coordinates in reciprocal space {(0, 0, 0)}, {(0, 0, 0.5), (0, 0.5, 0), (0.5, 0, 0)}, {(0, 0.5, 0.5), (0.5, 0, 0.5), (0.5, 0.5, 0)} and {(0.5, 0.5, 0.5)}, respectively.

The Z_2 topological invariants are determined by solving the following equation,

$$\delta_i = \prod_{m=1}^{N} \xi_{2m}(\Gamma_i) \quad (3),$$

where $\xi_{2m}(\Gamma_i) = \pm 1$ represent the parity eigenvalue of the $2m^{th}$ occupied energy band at TRIMs Γ_i. Due to the Kramers degeneracy at TRIMs, the eigenvalues of $2m^{th}$ and $(2m-1)^{th}$ occupied bands are equal, i.e., $\xi_{2m} = \xi_{2m-1}$. For 3D systems, four independent topological invariants ν_0; (ν_1,ν_2,ν_3), are commonly used to classify the topological class of system. These invariants can be calculated by using the following relations,

$$(-1)^{\nu_k} = \prod_{i=1}^{8} \delta_i \quad (4),$$

$$(-1)^{\nu_k} = \prod_{n_k=1, n_{j\neq k}=0,1} \delta_{i=(n_1n_2n_3)} \quad (5).$$

If the calculated value of ν_0 is nonzero then the system is a strong topological insulator (STI) otherwise we have to calculate the values of (ν_1,ν_2,ν_3). If the system has $\nu_{1,2,3} \neq 0$ then it is known as weak topological insulators (WTI), whereas topological invariant 0; (000) indicates the normal insulator (NI).

The Sr_3Bi_2 system has 48 occupied bands with SOC which predicts the topological classification of a system. Hence we used this complete set of bands to compute the values of Z_2 topological invariants, ν_0; (ν_1,ν_2,ν_3). Further, the parity eigenvalue of two consecutive bands, i.e. $2m^{th}$ and $(2m-1)^{th}$ occupied bands are the same and hence there are only 24 eigenvalues of the parity operator. At each TRIM, the parity eigenvalues of 24 occupied bands with even band index (ordered by energy) from a total of 24 occupied bands are computed. Multiplying over the parity eigenvalues of these 24 bands, the parity

Accepted Manuscript

of each TRIM, $\delta_{i=1,2,...8}$ in Eq. (3) are obtained. The parity of all 24 occupied bands and lowest unoccupied band at each TRIM for Sr_3Bi_2 system are given in Table 1. The Z_2 topological invariant ν_0 is determined by multiplying over the parities of all TRIMs using Eq. (4), whereas $\nu_{k=1,2,3}$ is derived by multiplying over the parities of TRIMs in the same plane using Eq. (5). The product of parity eigenvalues at eight TRIM for Sr_3Bi_2 system is listed in Table 2.

Table 1: Parity eigenvalues of Sr_3Bi_2 system at each TRIM for 24 occupied bands and lowest unoccupied band. The corresponding band energy increases from left to right. The product of parities at each TRIM is given in the last column.

TRIM	Parity $\{\xi_{2m}(\Gamma_i)\}$	Product (δ_i)
(0.0, 0.0, 0.0)	+ + + - - - + + - - + - + + - - + + - + + - - + ; -	-1
(0.5, 0.0, 0.0)	- + - + - + + - - - + + - - + + - - + + - + - + ; +	+1
(0.5, 0.5, 0.0)	+ - - + + - + - + + - - + - + - + - - - + + - + ; +	+1
(0.5, 0.5, 0.5)	+ - + + - - - - + + + - - + + - + + - - + + - - ; +	+1

Table 2: The parity product (δ_i) of 24 occupied bands at the eight TRIM and the corresponding parity product of eight TRIM to compute the Z_2 invariant (ν_0) for Sr_3Bi_2 system.

TRIM	1Γ	3L	3F	1T	Product	Z_2 invariant (ν_0)
Parity	-1	+1	+1	+1	-1	1

From the above calculations, the value of Z_2 topological invariant (ν_0) comes out to be nonzero, i.e., 1, thus the Sr_3Bi_2 is a strong topological insulator. The parity information and Z_2 topological invariant calculations for Ca_3Bi_2 are tabulated in Table 3 and 4 as given in appendix C.

4. Conclusions

In conclusion, by using the first-principles calculations, we investigated the structural, electronic and topological properties of the Sr_3Bi_2 and Ca_3Bi_2 compounds without and with SOC. These compounds are stable in hR45 crystalline phase with space group R3- (No. 148), but found to be unstable with a hexagonal structure having space group P3-m1 (No. 164), which is in contradiction to earlier reported works for Mg_3Sb_2 and Mg_3Bi_2 compounds. The Sr_3Bi_2 material is a type-I nodal-line semimetal without SOC. The nodal line is protected by the time reversal and spatial inversion symmetries and is created by continuous type-I nodal points. The projected bulk band structure and the corresponding band symmetry analysis predict the existence of band-inversion between the opposite parity eigenvalue bands near the Fermi-level at the Γ-point. The drumhead-like surface states hanging inside the nodal line were found for both surface terminations either Bi atomic layer or Sr atomic layer of the (001) surface projection of Sr_3Bi_2. When SOC is added, the nodal line becomes highly gapped, and a topological insulator state is formed. The value of the bandgap is found to be 0.36 eV, and the band-inversion persists, with an increase in the value of the inverted bandgap to 0.67 eV. As expected, conducting topological surface states with Dirac-like dispersion between the bulk bandgap are found for the (001) surface. The value of Z_2 topological invariants (1;000) are calculated by using the parity product criteria, which predicted that Sr_3Bi_2 is a strong topological insulator. It is important to comment that Ca_3Bi_2, also exhibits similar properties. The presence of large bandgap value in these materials at ambient conditions in the strong topological insulating phase, it is expected that scanning tunnelling microscopy and angle-resolved photoemission spectroscopy (ARPES) may provide unambiguous experimental proof of the topological surface states in the future. Based on these intriguing findings, Sr_3Bi_2 is a promising candidate material to realize the quantum spin Hall (QSH) effect at ambient conditions.

Accepted Manuscript

Appendix A: Structural Details of Ca_3Bi_2

It has been reported that Ca_3Bi_2 is found stable in hR45 structure with space group R3-, No. 148. We computed the cohesive energy per atom for Ca_3Bi_2 like Sr_3Bi_2. The value of cohesive energy per atom is -3.56 eV/atom, which indicates that Ca_3Bi_2 is stable in this structure. The optimized primitive unit cell (rhombohedral structure) for Ca_3Bi_2 with the optimized value of lattice parameter a = 8.7053 Å and rhombohedral angle α = 65.684° is given in Figure 1 (b). In the relaxed structure, the Bi atom is situated at 6f site (0.09088, 0.44531, 0.70477), Ca1 atom at 1b site (0.5, 0.5, 0.5), Ca2 atom at 2c site (0.32244, 0.32244, 0.32244) and Ca3 atom at 2f site (0.37580, 0.08336, 0.80894).

Appendix B: Electronic Properties of Ca_3Bi_2

We computed the projected bulk band structure along the momentum space path Γ–T–F–Γ–L to investigate the electronic properties of Ca_3Bi_2. It is easy to see that the electronic bulk band structure shown in Figure 5 (a) and the nodal line around the Γ-point in the $k_z = 0$ planes in the BZ shown in Figure 5 (b) of the Ca_3Bi_2 compound without SOC are like those of the Sr_3Bi_2 compound shown in Figures 2 (a) and (b). Thus, we can conclude based on discussion, as in the case of Sr_3Bi_2, that Ca_3Bi_2 is also a type-I NLSM without SOC.

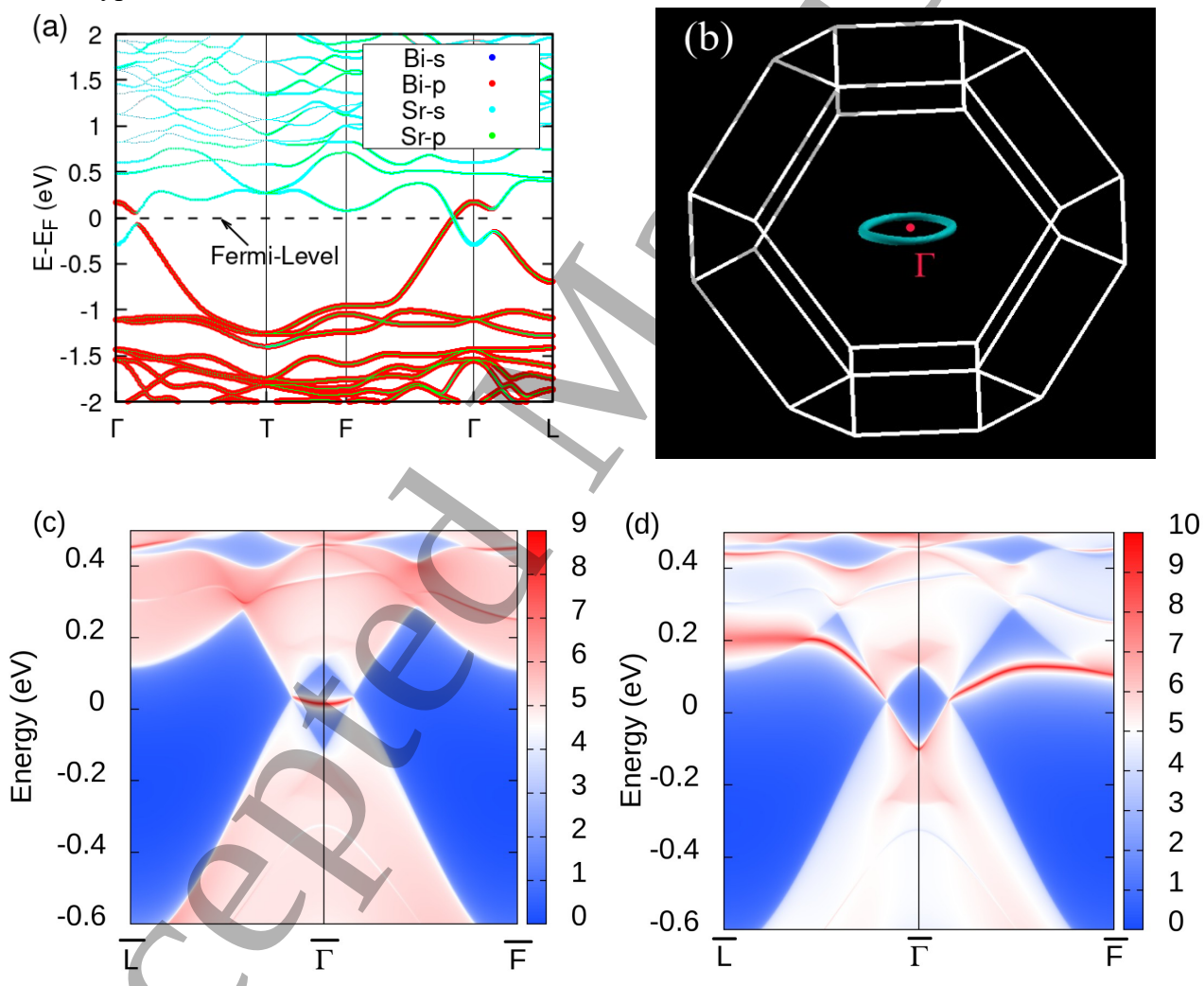


Figure 5: (a) Projected bulk band structure of Ca_3Bi_2 without SOC. (b) Illustration of type-I nodal ring in the bulk BZ. Computed (001) surface band structures of Ca_3Bi_2 for (c) Bi-terminated and (d) Ca-terminated surfaces without SOC.

Accepted Manuscript

As shown in Figure 6 (a), the inclusion of SOC opens a band gap of 0.26 eV near the Γ-point, converting Ca_3Bi_2 to TIs. The corresponding gapless SS for the (001) surface of the Ca_3Bi_2 compound is found to form the Dirac cone between the VB and CB and is shown in Figure 6 (b).

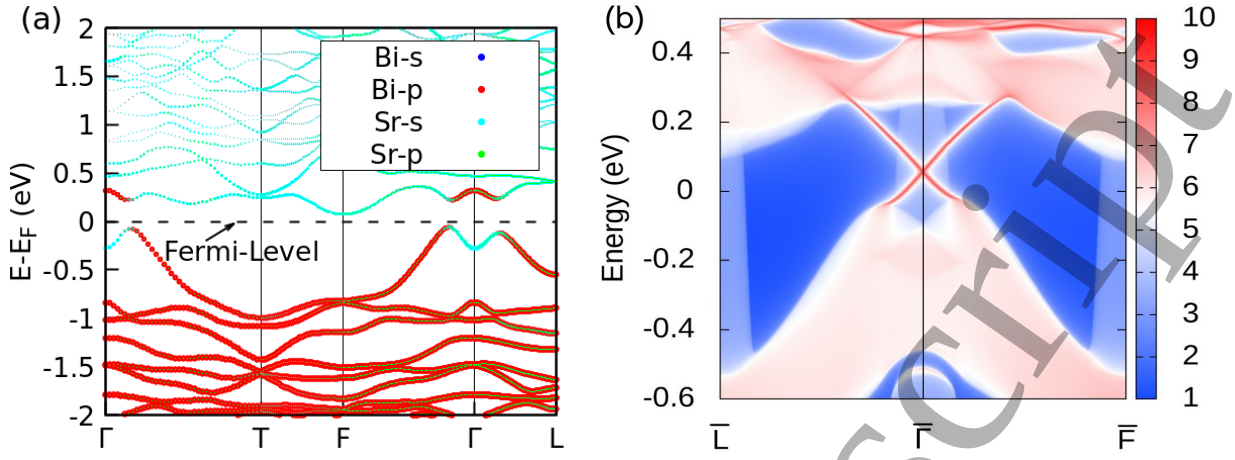


Figure 6: (a) Projected bulk band structure of Ca_3Bi_2 with SOC. (b) Gapless surface states of Ca_3Bi_2 terminated in the (001) surface. The (001) surface contains a 2D Dirac cone at Γ-point.

Appendix C: Topological Nature of Ca_3Bi_2

The parity of all 24 occupied bands and lowest unoccupied band at each TRIM for Ca_3Bi_2 system are given in Table 3 and the product of parity eigenvalues at eight TRIMs for Ca_3Bi_2 system is listed in Table 4.

Table 3: Parity eigenvalues of Ca_3Bi_2 system at each TRIM for 24 occupied bands and lowest unoccupied band. The corresponding band energy increases from left to right. The product of parities at each TRIM is given in the last column.

TRIM	Parity $\{\xi_{2m}(\Gamma_i)\}$	Product (δ_i)
(0.0, 0.0, 0.0)	+ + + - - - + + - - + - + + - - + + - + + - - + ; -	-1
(0.5, 0.0, 0.0)	- + - + - + + - - + - + - + - + - + - + - + - + ; -	+1
(0.5, 0.5, 0.0)	+ - - + - + + - + + - - + - - + - + - - + + + - ; +	+1
(0.5, 0.5, 0.5)	+ - + + - - - - + + + - - + + - + + - - + + - - ; -	+1

Table 4: The parity product (δ_i) of 24 occupied bands at the eight TRIM and the corresponding parity product of eight TRIM to compute the Z_2 invariant (ν_0) for Ca_3Bi_2 system.

TRIM	1Γ	3L	3F	1T	Product	Z_2 invariant (ν_0)
Parity	-1	+1	+1	+1	-1	1

From the above calculations, the value of Z_2 topological invariant (ν_0) is nonzero, i.e., 1 and hence the Ca_3Bi_2 is a strong topological insulator.

Acknowledgements

One of the authors (Sanjeev) is thankful to the University Grant Commission (UGC) India for giving financial support through the UGC-SRF fellowship. All the authors would like to thank **Pt. Deendayal Upadhyaya Innovation and Incubation Centre (PDUIIC),** Guru Jambheshwar University of Science and Technology, Hisar, India, for providing the funds under RUSA 2.0 Grant for setting the computational facility at Department of Physics, GJUS&T, Hisar.

Accepted Manuscript

Conflict of Interest

All the authors declared that they do not have any conflict of interest.

References

[1] Hasan M Z and Kane C L 2010 Colloquium: Topological insulators *Rev. Mod. Phys.* **82** 3045.

[2] Qi X-L and Zhang S-C 2011 Topological insulators and superconductors *Rev. Mod. Phys.* **83** 1057–1110.

[3] Bansil A, Lin H and Das T 2016 Colloquium: Topological band theory *Rev. Mod. Phys.* **88** 021004.

[4] Chiu C-K, Teo J C Y, Schnyder A P and Ryu S 2016 Classification of topological quantum matter with symmetries *Rev. Mod. Phys.* **88** 035005.

[5] Armitage N P, Mele E J, and Vishwanath A 2018 Weyl and Dirac semimetals in three-dimensional solids *Rev. Mod. Phys.* **90** 015001.

[6] Yan B and Zhang S-C 2012 Topological materials *Rep. Prog. Phys.* **75** 096501.

[7] Ando Y 2013 Topological insulator materials *J. Phys. Soc. Jpn.* **82** 102001.

[8] Liu Z K, Zhou B, Zhang Y et al. 2014 Discovery of a three-dimensional topological dirac semimetal, Na_3Bi *Science* **343** 864–867.

[9] Liu Z K, Jiang J, Zhou B et al. 2014 A stable three-dimensional topological dirac semimetal Cd_3As_2 *Nat. Mater.* **13** 677–681.

[10] Fujioka J, Kriener M, Hashizume D, Yamasaki Y, Taguchi Y and Tokura Y 2021 Alloying-induced enhancement of thermopower in the Dirac-semimetal system $Cd_{3-x}Zn_xAs_2$ *Phys. Rev. Materials* **5** 094201.

[11] Xu S-Y, Belopolski I, Alidoust N et al. 2015 Discovery of a Weyl fermion semimetal and topological Fermi arcs *Science* **349** 613–617.

[12] Lv B Q, Weng H M, Fu B B et al. 2015 Experimental discovery of Weyl semimetal TaAs *Phys. Rev. X* **5** 031013.

[13] Jin L, Zhang X, Liu Y, Dai X, Shen X, Wang L and Liu G 2020 Two-dimensional Weyl nodal-line semimetal in a d^0 ferromagnetic K_2N monolayer with a high Curie temperature *Phys. Rev. B* **102** 125118.

[14] Zhang X, Yu Z-M, Zhu Z, Wu W, Wang S S, Sheng X-L and Yang S A 2018 Nodal loop and nodal surface states in the Ti_3Al family of materials *Phys. Rev. B* **97** 235150.

[15] Wang X-B, Ma X-M, Emmanouilidou E et al. 2017 Topological surface electronic states in candidate nodal-line semimetal CaAgAs *Phys. Rev. B* **96** 161112(R).

[16] Shiozaki K and Sato M 2014 Topology of crystalline insulators and superconductors *Phys. Rev. B* **90** 165114.

[17] Wang Z, Alexandradinata A, Cava R J and Bernevig B A 2016 Hourglass fermions *Nature* **532** 189-194.

[18] Ma J, Yi C, Lv B et al. 2017 Experimental evidence of hourglass fermion in the candidate nonsymmorphic topological insulator KHgSb *Sci. Adv.* **3** 1602415.

[19] Sharma A, Bhattacharyya B, Srivastava A K, Senguttuvan T D and Husale S 2016 High performance broadband photodetector using fabricated nanowires of bismuth selenide *Sci. Rep.* **6** 19138.

Accepted Manuscript

[20] Wang Z, Mu H, Zhao C, Bao Q and Zhang H 2016 Harmonic mode-locking and wavelength-tunable Q-switching operation in the graphene–Bi_2Te_3 heterostructure saturable absorber-based fiber laser *Optical Engineering* **55** 081314.

[21] Crowley J M, Tahir-Kheli J and Goddard W A 2015 Accurate Ab Initio Quantum Mechanics Simulations of Bi_2Se_3 and Bi_2Te_3 Topological Insulator Surfaces *J. Phys. Chem. Lett.* **19** 3792–3796.

[22] Weng H, Dai X and Fang Z 2016 Topological semimetals predicted from first-principles calculations *J. Phys.: Condens. Matter* **28** 303001.

[23] Fang C, Weng H, Dai X and Fang Z 2016 Topological nodal line semimetals *Chinese Phys. B* **25** 117106.

[24] Yamakage A, Yamakawa Y, Tanaka Y and Okamoto Y (2016) Line-Node Dirac Semimetal and Topological Insulating Phase in Non-centrosymmetric Pnictides CaAgX (X = P, As) *J. Phys. Soc. Jpn*. **85** 013708.

[25] Weng H, Fang C, Fang Z and Dai X 2016 Coexistence of Weyl Fermion and Massless Triply Degenerate Nodal Points *Phys. Rev. B* **94** 165201.

[26] Yu R, Weng H, Fang Z, Dai X and Hu X 2015 Topological Node-Line Semimetal and Dirac Semimetal State in Antiperovskite Cu_3PdN *Phys. Rev. Lett.* **115** 036807.

[27] Bian G et al. 2016 Drumhead Surface States and Topological Nodal-Line Fermions in $TlTaSe_2$ *Phys. Rev. B* **93** 121113(R).

[28] Bian G et al. 2016 Topological Nodal-Line Fermions in Spin-Orbit Metal $PbTaSe_2$ *Nature Commun.* **7** 10556.

[29] Zhang H, Liu C-X, Qi X-L, Dai X, Fang Z and Zhang S-C 2009 Topological insulators in Bi_2Se_3, Bi_2Te_3 and Sb_2Te_3 with a single Dirac cone on the surface *Nature Physics* **5** 438-442.

[30] Wang Z et al. 2012 Dirac semimetal and topological phase transitions in A_3Bi (A= Na, K, Rb) *Phys. Rev. B* **85** 195320.

[31] Lin H, Markiewicz R S, Wray L A, Fu L, Hasan M Z and Bansil A 2010 Single-Dirac-cone topological surface states in the $TlBiSe_2$ class of topological semiconductors *Phys. Rev. Lett.* **105** 036404.

[32] Zhang Q, Cheng Y and Schwingenschlogl U 2015 Emergence of topological and topological crystalline phases in $TlBiS_2$ and $TlSbS_2$ *Sci. Rep.* **5** 8379.

[33] Rajaji V, Arora R, Sarma S C, Joseph B, Peter S C, Waghmare U V and Narayana C 2019 Phonon signature of multiple topological quantum phase transitions in compressed $TlBiS_2$: A combined experimental and theoretical study *Phys. Rev. B* **99** 184109.

[34] Neupane M et al. 2012 Topological surface states and Dirac point tuning in ternary topological insulators *Phys. Rev. B* **85** 235406.

[35] Li Y et al. 2021 Topological Dirac surface states in ternary compounds $GeBi_2Te_4$, $SnBi_2Te_4$ and $Sn_{0.571}Bi_{2.286}Se_4$ *Chin. Phys. B* **30** 127901.

[36] Li Y et al. 2022 Coexistence of ferroelectric-like polarization and Dirac-like surface state in $TaNiTe_5$ *Phys. Rev. Lett.* **128** 106802.

[37] Chen Z et al. 2021 Three-dimensional topological semimetal phase in layered $TaNiTe_5$ probed by quantum oscillations *Phys. Rev. B* **103** 035105.

[38] Jiao W-H et al. 2020 Topological Dirac states in a layered telluride $TaPdTe_5$ with quasi-one-dimensional $PdTe_2$ chains *Phys. Rev. B* **102** 075141.

Accepted Manuscript

[39] Zhang T et al. 2009 Experimental Demonstration of Topological Surface States Protected by Time-Reversal Symmetry *Phys. Rev. Lett.* **103** 266803.

[40] Alpichshev Z et al. 2010 STM Imaging of Electronic Waves on the Surface of Bi_2Te_3: Topologically Protected Surface States and Hexagonal Warping Effects *Phys. Rev. Lett.* **104** 016401.

[41] Lou R et al. 2017 Evidence of topological insulator state in the semimetal LaBi *Phys. Rev. B* **95** 115140.

[42] Nayak J et al. 2018 Electronic properties of topological insulator candidate CaAgAs *J. Phys.: Condens. Matt.* **30** 045501.

[43] Xia Y et al. 2009 Observation of a large-gap topological-insulator class with a single Dirac cone on the surface *Nature Physics* **5** 398-402.

[44] Chen Y L et al 2009 Experimental realization of a three-dimensional topological insulator, Bi_2Te_3 *Science* **325** 178-181.

[45] Eremeev S V et al. 2012 Ideal Two-Dimensional Electron Systems with a Giant Rashba-Type Spin Splitting in Real Materials: Surfaces of Bismuth Tellurohalides *Phys. Rev. Lett.* **108** 246802.

[46] Ishizaka K et al. 2011 Giant Rashba-type spin splitting in bulk BiTeI *Nature Mater.* **10** 521-526.

[47] Krausz Z K, Hoque A M and Makk P 2020 Electrically Controlled Spin Injection from Giant Rashba Spin–Orbit Conductor BiTeBr *Nano Lett.* **20(7)** 4782-4791.

[48] Maaß H, Bentmann H, Seibel C *et al.* 2016 Spin-texture inversion in the giant Rashba semiconductor BiTeI *Nat. Commun.* **7** 11621.

[49] Xin J-Z et al. 2018 Synthesis and thermoelectric properties of Rashba semiconductor BiTeBr with intensive texture *Rare Metals* **37** 274-281.

[50] Jin M L et al. 2019 Pressure-induced superconductivity and quantum phase transitions in the Rashba material BiTeCl *J. Phys. Chem. Solids* **128** 211−217.

[51] Zhang X, Liu Q, Xu Q, Dai X and Zunger A 2018 Topological Insulators versus Topological Dirac Semimetals in Honeycomb Compounds *J. Am. Chem. Soc. B* **140,42** 13687–13694.

[52] Shao D, Guo Z, Wu X, Nie S, Sun J, Weng H and Wang Z 2021 Topological insulators in the NaCaBi family with large spin-orbit coupling gaps *Phys. Rev. Res.* **3** 13278.

[53] Sanjeev, Singh M, Kumar R, Srivastava S and Kumar T 2022 Structural, electronic and topological properties of NaCaBi and KBaBi compounds *J. Phys. Chem. Solids* **161** 110416.

[54] Teshome T and Datta A 2019 Topological phase transition in Sb_2Mg_3 assisted by strain *ACS Omega* **4** 8701-8706.

[55] Zhang X, Jin L, Dai X and Liu G 2017 Topological Type-II Nodal Line Semimetal and Dirac Semimetal State in Stable Kagome Compound Mg_3Bi_2 *J. Phys. Chem. Lett.* **8** 4814−4819.

[56] Chang T-R et al. 2019 Realization of a Type-II Nodal-Line Semimetal in Mg_3Bi_2 *Adv. Sci.* **6** 1800897.

[57] Mondal C, Barman C K, Kumar S, Alam A and Pathak B 2019 Emergence of topological insulator and nodal line semi-metal states in XX'Bi (X = Na, K, Rb; X' = Ca, Sr) *Sci. Rep.* **9** 527.

[58] Li R, Xie Q, Cheng X, Li D, Li Y, Chen X-Q 2015 First-principles study of the large-gap three-dimensional topological insulators M_3Bi_2 (M=Ca, Sr, Ba) *Phys. Rev. B* **92** 205130.

Accepted Manuscript

[59] Fu L and Kane C L 2007 Topological insulators with inversion symmetry *Phys. Rev. B* **76** 045302.

[60] Giannozzi P, Andreussi O and Brumme T et al. 2017 Advanced capabilities for materials modelling with QUANTUM ESPRESSO *J. Phys. Condens. Matter* **29** 465901.

[61] Giannozzi P, Baroni S, Bonini N et al. 2009 QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials *J. Phys. Condens. Matter* **21** 395502.

[62] Perdew J P, Burke K and Ernzerhof M 1996 Generalized Gradient Approximation Made Simple *Phys. Rev. Lett.* **77** 3865-3868.

[63] Mostofi A A, Yates J R, Lee Y-S, Souza I, Vanderbilt D and Marzari N 2008 Wannier90: a tool for obtaining maximally-localised Wannier functions *Comput. Phys. Commun.* **178** 685–699.

[64] Wu Q, Zhang S, Song H-F, Troyer M and Soluyanov A A 2018 WannierTools: an open-source software package for novel topological materials *Comput. Phys. Commun.* **224** 405-416.

Accepted Manuscript

The Journal of Supercomputing
https://doi.org/10.1007/s11227-023-05090-w

Trust factor-based analysis of user behavior using sequential pattern mining for detecting intrusive transactions in databases

Indu Singh[1] · **Rajni Jindal**[1]

Accepted: 2 February 2023
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

Abstract
Organizations today are employing databases on a large scale to store data essential for their functioning. Malicious access and modifications of the databases may lead to adverse financial and legal implications. In recent years, security researchers have focused on detecting abuse of access privileges by employees of an organization. Identifying threats from insiders is hard because they are aware of the organization of the database in addition to having authorised access privileges. To detect insider attacks effectively and efficiently, we present a novel approach to dynamically determine the malicious transactions using historical data. We propose Trust factor-based user behavior analysis using sequential pattern mining for database intrusion detection systems (TFUBID). Since, groups of users access the organizational database for similar purposes, we cluster user behavior vectors using fuzzy clustering and define a class of Integral Data Attributes using sequential pattern mining to model trust factor-based behavioral patterns of employees accessing the database assigning higher weight to critical elements and Directly Correlated Attributes. A comprehensive experimental evaluation on our synthetic dataset adhering to TPC-C standard benchmark revealed that TFUBID achieved an accuracy of 94% for detecting malicious transactions and outperforms competing state-of-the-art techniques on several performance measures.

Keywords Database intrusion detection · Database security · Sequential pattern mining · Data dependency mining · User behaviour analysis · Anomaly detection

✉ Indu Singh
indusingh@dtu.ac.in

Rajni Jindal
rajnijindal@dce.ac.in

[1] Department of Computer Science Engineering, Delhi Technological University, New Delhi, Delhi 110042, India

Published online: 21 February 2023

1 Introduction

The modern era is often referred as the "data economy" by virtue of the importance of data in almost all phases of business activity starting from product ideation to realization. Databases represent the most fundamental means to store large quantities of data for modern enterprises. However, with the proliferation of reported instances of unauthorized access to information and data theft, the security of databases has often come into question in recent times [4]. Protecting data from unauthorized access of employees of an organization is integral to maintaining its integrity and confidentiality, which represent two of the most important considerations for information security within an organization as per the confidentiality, integrity, availability (CIA) triad. Malicious updates to the database along with a breach in confidentiality may lead to loss of reputation, litigations and may lead to severe financial ramifications for organizations, and thus protecting data from malicious insiders is pivotal to their smooth functioning.

The Computer Emergency Response Team (CERT) report on insider threats [9] defined malicious insiders as employees and business partners who have authorized access to company database, but wilfully misuse their privileges to adversely impact the confidentiality, integrity and availability of an information system. Since insiders either have full or at least partial access to the organizational database, insider threats are harder to detect than threats from outsiders [15]. Traditional mechanisms such as access control and authentication are inadequate in meeting the security needs of modern organizations. Furthermore, most traditional security mechanisms are preventive in nature, however many insiders can easily corrupt the data stored in the database. Since most attacks by insiders and outsiders who have assumed the identity of legitimate insiders cannot be prevented, more advanced and effective data-driven mechanisms are needed.

Intrusion detection systems [12, 17, 18, 39] (IDSs) are a class of data driven security systems which dynamically monitor all transactions issued by employees of an organization and decide whether they are indicative of an attack or constitute a legitimate use the organizational database Since Denning's [12] seminal work, many intrusion detection systems have been proposed. While some IDSs learn regular patterns in database access and flag any deviation from the normal as an intrusion, some other IDSs detect malicious transactions with the database using well-known "signatures" of attacks [26, 33]. The former kind are referred as anomaly based IDSs [17, 18, 47] while the latter are called signature-based IDS. Upon detecting a malicious transaction, an IDS often calls a separate intrusion recovery system which handles the transaction as per organizational policy.

Recent studies on database intrusion detection systems (DIDS) have focused upon user behavior analysis to mine regular patterns of database access of users from their historical log data in the system [43]. Actions by users are then dynamically matched against their regular patterns to detect malicious behavior. However, modern IDSs need to address another important problem: organizations require numerous attributes to efficiently store the rapidly increasing quantities of data. As the number of attributes increases, database administrators may

find it extremely hard to keep track of malicious modifications on them [40]. The fact that insider threats can cause considerable damage to the system without being detected for extended periods of time so an intrusion detection system that evolves and adapts to threats encountered during operation is needed to maintain the integrity of data in such a complex environment. To this end, we consider a hierarchy of attribute in terms of their importance. More important attributes in a database such as the remuneration of employees, their personal identification numbers etc. are categorized as critical attributes.

In this paper, we develop an effective database IDS called TFUBID to prevent misuse of access privileges by insiders of an organization. We cluster vectors illustrative of the behaviour of each user in order to model general behavioral patterns rather than user specific patterns only. This is based on the assumption that groups of users access the organizational database for similar purposes. We cluster the user vectors using fuzzy clustering wherein each data point can belong to more than one cluster. Fuzzy clustering is used since employees may belong to different groups in the organization, and therefore their behavioral vectors may belong to more than one cluster. For each user we maintain a trust factor on the basis of their access pattern in the past. Trust factor is calculated by the rules which are generated using prefix span algorithm [30] of Sequential Pattern Mining. We update this score with every transaction and use this for quantifying the maliciousness of the user.

To summarize our paper has the following major contributions:

1. A Novel approach is presented for analyzing user behaviour based on upon trust factor of transactions committed by the user.
2. A new architecture for DIDS is designed that aims to prevent access privilege misuse, unauthorized access and modification of both critical and non critical attributes that are under direct relationship with critical attributes.
3. Modified Jensen-Shannon distance is used to give weights to data attributes and avoid the curse of dimensionality in dissimilarity calculation.
4. The notion of an "incredulity score" is introduced which quantifies the degree of anomalous behaviour exhibited by each user based on his previous transactions.

The rest of the paper is organized as follows. Section 2 presents related work, while Sect. 3 presents our novel approach TFUBID based on sequential pattern mining and fuzzy clustering. Results from our detailed experiments are discussed in Sect. 4, and we finally conclude the paper by summarizing our finding and exploring avenues of future work in Sect. 5.

2 Related work

Very few researches have been carried out in Database intrusion detection system as compared to Network based. Several systems [3, 27] have been devised for Intrusion Detection in operating systems and networks, but they fail to prevent database attacks. Database intrusion detection system are devised to work at query, transaction and user level (Table 1). Bertino et al. [5] discussed the need to tackle insider

threats and explained the roadblocks in ensuring confidentiality, integrity and availability. Kumar and Mohbey [23] gave a broad overview of the distinct approaches in parallel, distributed and scalable pattern mining in Big Data domain.

Denning [12] proposed a intrusion detection expert system (IDES) comprising of well-defined components including: profiles, subjects and objects, audit records and anomaly records, valid activity rules. The model works on the principle that violations in security can be effectively detected by monitoring the audit records of a system in order to identify and detect the anomalous patterns in system usage.

Broadly two techniques have been developed to cater to IDS in databases. The first technique is signature-based IDS [26, 33] which keeps a record of previous attacks in the form of attack signatures and are able to detect an incoming activity as malicious based on these signatures. Although, this technique has very low false alarm rates, it suffers from two major drawbacks. Firstly the database of signatures needs to be updated regularly and secondly the algorithm fails to detect novel attacks whose signatures are not known. To detect novel attacks anomaly based IDS's [17, 18, 47] were developed that measure the deviation in normal user pattern and assert the presence of insider threats. Research work in the detection of anomalous events [11] and event patterns illustrates that identification of anomalous events is based on user access patterns and user profiles. Furthermore, the detection of insider attacks [43] by auditing the user profiles at both parallel and incremental levels is done by learning a behaviour pattern for every user. Then probabilistic analysis is used for user evaluation against group patterns and routine patterns. The measure of deviation is then adopted as a score to justify and identify the existence of an anomaly. Thus we have devised a new approach based on user behaviour analysis that incorporates rule mining and fuzzy clustering to overcome above mentioned disadvantages.

2.1 Data dependency mining

Data dependency mining is a method for extracting useful patterns in a database through statistical methods. It finds dependencies among various data items in the database which can then be used for further analysis. We discuss some important papers utilizing this technique for intrusion detection.

Hu and Panda [17, 18] developed a data dependency miner for mining correlations from the database log. Data dependency rules were mined from transaction logs and queries that do not comply with these rules were deemed malicious. One of the limitations of this approach is that it treats all the attributes with same sensitivity, which is not always true in real world. In 2005, Srivastava et al. [40] proposed another data dependency approach which incorporated attribute sensitivity. They also used time pattern of transaction executions to identify normal or anomalous transactions, which reduced the number of false negatives.

Subudhi and Panigrahi [42] proposed an approach for detecting intrusive activities in databases using clustering and information fusion through ensemble learning. A transaction is passed through the clustering module for computing its cluster belongingness and an outlier factor that signifies its degree of outlierness.

Table 1 Overview of major database intrusion detection systems

Authors	Proposed technique	Approach for query evaluation	Scalability	User behaviour analysis	Anomaly detection	Drawbacks
Hu and Panda [17]	Integrated dependency with sequence alignment analysis	Syntax centric	Yes	No	Yes	Works only when data dependencies exist. Emphasizes on malicious modification of data, while malicious read operation cannot be detected
Srivastava et al. [41]	Integrated dependency with sequence alignment analysis	Syntax centric	No	No	Yes	Only update sequences are considered for malicious checks. Poses a problem in selection of appropriate value of support and confidence
Hashemi et al. [16]	Temporal mining	Syntax centric	No	No	Yes	Frequent association rule mining-based database IDSs ignore sequences that have frequency lower than minimum support threshold. So no rule is produced for the sequences of data items that are infrequent and of great interest
Sun et al. [43]	Behavioral Analysis	Data centric	Yes	Yes	Yes	In the anomaly measurement, the results highly rely on selection of the parameter k
Subudhi and Panigrahi [42]	Application of OPTICS and ensemble learning for database intrusion detection	Syntax centric	No	No	No	Integrated models have not yet been evaluated for refrigeration system fault diagnosis
Kim and Cho [22]	Optimizing CNN-LSTM neural networks with PSO for anomalous query access control	Syntax centric	No	No	Yes	The model provide little lexical interaction information for achieving desirable outcomes
Bu et al. [8]	Deep convolutional learning classifier system based on genetic algorithm	Syntax centric	Yes	No	No	The method is not trusted on different database schema

Depending on the outlier factor value, the transaction is classified as genuine or an outlier.

2.2 Sequential pattern mining

The concept of sequential pattern mining (SPM) has been extensively in the field of database security [17, 25] to discover frequently occurring sequences from audit logs, and subsequently form data dependency rules to characterize normal user behaviour.

For deriving association rules in large databases Agrawal and Srikant [1] presented new algorithms—Apriori and its variations. Though the Apriori algorithm does not scale well for large databases, Apriori Hybrid algorithm scales linearly with the size of the database. One major limitation encountered was setting appropriate values of support and confidence values for the algorithm. Further in the 1995 Agrawal and Srikant [2] proposed AprioriAll and AprioriSome for mining sequential patterns in a database. Both the algorithms scale linearly with size of database but suffer from the same problem of setting the support and confidence values. The limitations were overcome in 1996 when Agrawal and Srikant [39] presented generalized sequential pattern mining algorithm which was found to be more efficient than the Apriori algorithm. It is different from traditional pattern mining in the sense that it considers different significance of the items in real applications, with the weight being cost or profit of the item. Hu and Panda [17] in 2004, devised an efficacious data mining algorithm to recognize fraudulent transactions in database. It utilizes AprioriAll algorithm for generating sequential patterns which are applied to the rule generation algorithm to develop classification rules. The transactions that do not conform with these acquired rules are classified as malicious in nature. A shortcoming with this approach is that it considers that all the data items have same importance which is not always true in real world.

These limitation was overcome through weighted sequential pattern mining. The main focus of weighted sequential pattern mining was concerned with maintaining the downward closure property. Therefore, Yun and Leggett [46] in 2005 developed an approach that maintained downward closure property by using weight range and minimum weight constraints.

In 2014, Lan et al. [25] proposed an efficient method to extract weighted sequential patterns from sequential databases. A drawback with this model is that it is not able to handle dynamic insertion, deletion or modification of sequences.

Rahman et al. [31] in 2018 introduced the notion of weighted uncertain sequence mining through his algorithm that mined sequences with weight constraints from uncertain databases. The main challenge with this approach is to recognize the actual meaningful patterns as different patterns bear different relevance. This challenge was overcome when Rahman et al. [32] in 2019 proposed a new approach to generate sequences following both weight and support constraints in uncertain databases. Singh and Jindal [36] proposed a DIDS based upon expectation maximization clustering and sequential pattern mining (EMSPM) for detecting intrusive transactions. Kumar and Mohbey [24] proposed algorithm based on the Spark framework which uses a search space division strategy for data parallelism and effective task

allocation across cluster nodes. Kim et al. [20] proposed a personalized alarm model to detect frauds in online banking transactions using sequence pattern mining on each user's normal transaction log. Singh et al. [38] presented a hybrid metaheuristic clustering and closed sequential pattern mining-based DIDS for detecting malicious transactions. Jindal and Singh [19] proposed a unique method (FPGWWO) for DIDS build on frequent SPM and a modified metaheuristic hybrid clustering of Grey Wolf (GWO) and Whale optimization (WOA) algorithm.

2.3 User behaviour analysis

User behaviour analysis (UBA) in the field of intrusion detection comprises of mining normal user behaviour from the audit logs and comparing it with a new transaction to measure it's deviation. The new transaction is then deemed anomalous if the deviation exceeds a certain threshold value.

UBA has also been used in several other fields. Yang [45] used web usage data to construct user profiles using the concept of pattern strength and Euclidean distance metric. A major limitation of their approach is that user behaviour is dynamic in nature i.e., it can change with time, and this paper does not take this into account. Panell and Ashman [29] proposed a host-based network anomaly detection scheme which worked by profiling normal usage patterns of individual users. Detection was done by comparing a learned user profile against a possible intrusion. A drawback here is that an intruder can escape this IDS by mimicking the behaviour of a normal user. Hence a combination of user characteristics should be used for profiling users, which significantly improves the accuracy of system as well as decreases the time taken to detect the intruder. Ruijuan et al. [35] applied user behaviour analysis in a mobile cloud environment. They applied neural network clustering to form user profiles and use an information entropy layer in the network to account for weights of attributes. Similarity values of the new transaction are calculated and if it exceeds a certain threshold it is flagged as anomalous behaviour.

2.4 Anomaly detection

In IDSs anomaly detection is used to detect a malicious user or an intruder whose activity trends significantly deviates from that of a normal user. DEMIDS [10] generates user profiles based on audit trails. The profiles computed can be used to detect misuse and insider abuse. Through the notion of distance measure, it inputs information about the data structures and semantics of a particular database schema which in turn is used to obtain frequent itemsets describing the working scopes of users. Anomaly based intrusion detection system is able to determine role intruders [16], which finds the individuals behave differently than expected while holding a specific user role. This model proves to be extremely useful as it can help in protecting against insider threats. Ronao and Cho [34] proposed an anomaly detection technique for RBAC databases using PCA and Random Forest with weighted voting as feature selection techniques. This approach improves the false-positive and false-negative rates and reduces model building and execution time significantly. A

limitation of this paper is that confusion can arise between roles which have similar read write access rights. Kim and Cho [21] proposed a method of classifying user's role and authority using the CNN-LSTM neural networks by extracting features from SQL queries. In the anomaly detection method, CNN automatically extracts important features from database query and LSTM models the temporal information of the SQL sequence. The class activation map also identifies the SQL query features that affect the classification. CNN-LSTM neural networks surpass other state-of-the-art machine learning methods by great extent. Bu and Cho [7] built a convolutional neural-based learning classifier system (CN-LCS) that models the role of queries by combining conventional learning classifier system (LCS) with convolutional neural network (CNN) for a database intrusion detection system based on the RBAC mechanism is proposed. The modified Pittsburgh-style LCSs for the optimization of feature selection rules is combined with one-dimensional CNNs for modeling and classification. Singh et al. [37] proposed a database intrusion detection system (DIDS) using fractional distance-based anomaly detection with partially ordered dependency analysis to identify malicious transactions in the databases.

3 Proposed methodology

3.1 Basic notations

Large organizations deal with vast amounts of data whose security is of primary interest. Data in databases comprises of attributes describing real world objects called as entities. These attributes have varying levels of sensitivity, i.e., not all attributes are equally important to the integrity of database. As an example, signatures and other biometric information is highly sensitive for financial institutions, such as banks in comparison with other attributes like name, gender and age. Therefore, unauthorized access to these sensitive attributes is of greater concern and only certain employees may have access to such data elements and access by all others must be blocked to ensure the confidentiality and integrity of data.

To this end, we propose a two-phase intrusion detection and prevention model TFUBID *(Trust factor-based user behaviour analysis for database intrusion detection)* that clusters users based on similarity of their data access patterns and the types of queries submitted by them, i.e., our model tracks the access patterns of all users and classifies them as normal or malicious. The superiority of our model lies in its ability to prevent unauthorized retrieving and modification of most sensitive data elements referred to as IDAs or Integral Data Attribute. Our model also makes sure that the query pattern for access of IDAs is specific and fixed for a particular user to avoid data breaches, i.e., the user associates himself with his regular access behaviour. Any deviation from the usual may lead to a reduction in the user's confidence and may act as representative of user's malicious intent.

Some attributes can be used to indirectly infer IDAs and are therefore critical to the functioning of the organization. For instance, account number of a user may be used to access the signatures. Such attributes are referred as DCA (Directly Correlated Attributes). In our paper, the following terminologies are used:

Definition 1 *(Transaction)* Set of queries executed by a user. Each transaction is represented by a unique transaction ID and also carries the user ID. Hence $< U_{id}, T_{id} >$ act as unique identification key for each set of query patterns. Each Transaction T is denoted as:

$< U_{id}, T_{id}, < q_1, q_2, \ldots q_n >>$ where, q_i denotes the i_{th} query, $i \in [1 \ldots n]$

For example, suppose a user has ID 1001 and executes the following set of SQL queries:

```
q1: SELECT a,b,c FROM R1,R2 WHERE R1.A>R2.B
q2: SELECT P FROM R5 WHERE R5.P==10
```

Then this can be represented as a transaction of the form:
$t =< 1001, 67, < q1, q2 >>$

Definition 2 *(Query)* We define a query as a read or write request on an attribute of a relation. A query can represented as $< O(D_1), O(D_2), \ldots O(D_n) >$ where, $D_1, D_2, \ldots D_n \in Rs$ and Rs is the relation schema and D_i are the attributes. O represents the operations i.e., Read or write Operations. $O \in \{R, W\}$

For example, let us consider the following transaction:

```
start transaction

select balance from Account
where Account_Number='9001';

update Account set balance=balance-900
where Account_Number='9001';

commit; //if all sql queries succeed
rollback; //if any of Sql queries failed or error
```

The query corresponding to this transaction is:
<<(Account Number), R(balance)>, <(Account Number), R(balance), W(balance)> >

Definition 3 *(Read Sequence)* A read sequence is defined as $R(x_1), R(x_2), \ldots O(x_n)$ where O represents an operation i.e., Read or write Operation and thus $O \in \{R, W\}$. The Read sequence represents that the transaction may need to read all data items $x_1, x_2, \ldots, x_{n-1}$ before the transaction performs an operation $O \in \{R, W\}$ on data item x_n.

Definition 4 *(Write Sequence)* A write sequence is defined as $O(x_1), W(x_2), \ldots W(x_n)$ The write sequence is a representation of a transaction that needs to write all the data items $x_1, x_2, \ldots, x_{n-1}$ in this order after the operating upon the data item x_n.

Definition 5 *(Read Rules)* Read rules (RR) are the association rules generated from Read sequences whose confidence is greater than a user defined threshold (Ψ_{conf}). A read rule is represented as $R(x_1), R(x_2) ... \implies O(x)$. For all sequential patterns $< R(x_1), R(x_2), \ldots, R(X_{n-1}), O(x_n) >$ in read sequence set, the read rules are

generated as $R(x_1), R(x_2) \ldots \Longrightarrow O(x_n)$. If the confidence of the rule is greater than the minimum confidence (Ψ_{conf}), then it is included to the collection of read rules, which implies that before x_n, we need to read $x_1, x_2 \ldots x_{n-1}$. For example: The Read Rule corresponding to the read sequence

$< R(a), R(b), R(c), R(d), W(x) >$ is: $R(a), R(b), R(c), R(d) \implies W(x)$.

Definition 6 *(Write Rules)* Write rules (*WR*) are association rules generated from write sequences whose confidence is greater than a user defined threshold (Ψ_{conf}). A write rule is represented as

$O(x) \implies W(x_1), W(x_2) \ldots$

For all sequential patterns $< O(x), W(x_1), W(x_2)$, ...,$(x_k) >$ in the write sequence set, the write rules are generated as $O(x) \rightarrow W(x_1), w(x_2)$, ...,$w(x_k)$. If the confidence of a rule exceeds minimum confidence (Ψ_{conf}), then it is included in the set of write rules depicting that after updating x, data items x_1, x_2, ..., x_k should be updated as a part of the same transaction. For Example: The write rule corresponding to te write sequence $< W(x), W(y), W(z) >$ is $W(x) \implies W(y), W(z)$

Definition 7 *(Integral data attributes (IDA))* These data attributes are of prime significance and have direct correlation with the integrity of the system. In a vertically hierarchical organisation, these are the attributes accessed only by the top level management, and the access by lower levels of hierarchy is strictly protected. User access pattern for IDAs is used by our model as an indicative of user's behaviour which helps in detection of malicious activities in the system (Table 2).

Definition 8 *(Critical rules (CR))* A set of rules that contain a Integral Data Attributes in its antecedent or consequent.

$$\begin{aligned} CR = \zeta \mid (\zeta \in RR \vee \zeta \in WR) \\ \cap (x \in IDA \cap (R(x_1), R(x_2) \ldots \\ \implies O(x) \cup O(x) \implies W(x_1), W(x_2) \ldots)) \end{aligned} \tag{1}$$

We propose a method of user Access Pattern Recognition using the Critical Rules, which represent the user query pattern associated to the Integral Data Attributes.

Definition 9 *(Directly correlated attributes (DCA))* are attributes except IDAs, which are either part of antecedents or consequents of Critical Rules (Table 2).

$DCA = \{\mu_i | \mu_i \in CR \cap \mu_i \notin IDA\}$.

For example, if $R(b) \rightarrow R(a)$

$R(b), R(c) \rightarrow R()$ and if a is a IDA, then the set $\{b, c\}$ represents DCAs.

Table 2 Sensitivity level of different attributes

Type of attribute	Sensitivity level
Integral data attributes	Highest
Directly correlated attributes	Medium
Normal attributes	Low

The query patterns as perceived by our model TFUBID are explored using DCAs that represent the first level of access of the IDAs. The first level of access means that these elements are the connecting elements and are accessed before IDAs or IDAs are accessed via these elements. A user's behaviour is represented by a set of first-order statements (derived from queries) called attribute hierarchy encoded in first-order logic, which defines abstraction, decomposition and functional relationships between types of access arrangements. The unit-transactions accessing IDAs are decomposed into attribute hierarchy comprising of DCAs, which further represents the user's most sensitive retrieval pattern.

Definition 10 (*Incredulity Score*(ϕ)) Incredulity Score is indicative of the amount of deviation of the user's access pattern from his designated role. This score summarizes the user's historic malicious access attempts and attempts to quantify the extent of vulnerability that the organisation faces because of a particular user. Incredulity Score combined with the deviation of user's present query from his normal behaviour pattern, yields the output of the proposed IDS.

For our paper:

$0 <= \phi <= 1$.

Higher the Incredulity Score, more is the evidence of user deviating from the assigned role and the higher is the malicious intent.

Definition 11 (*Trust factor*) Trust factor is derived from the incredulity score and quantifies how credible a user is. It is calculated as

$TF = 1 - \phi$

Based on the values of the Trust factor (Table 3), we have proposed a division of Users among different levels, which suggests how likely a user, is going to execute a malicious query. A low trust level suggests that a user is more likely to execute a malicious query while a user with Medium trust level is less probable to do so as compared to the ones belonging to Low trust level. High Trust Level denotes the class of genuine users which will execute malicious queries with very less probability and thus can be trusted.

It must be noted that the Trust Factor ranges are subjective and may differ from one organisation to the other.

Definition 12 *(Incredulity Table)* In our approach, we also maintain a the incredulity scores of each user. The initial Incredulity scores for all users are set to 1. The initial

Table 3 Levels of trust factor

TF range	Trust level
0–0.3	Low
0.3–0.6	Medium
0.6–1	High

incredulity score is shown in Table 4 as ϕ_i. Initially a user is assumed to be malicious due to no previous record but with each query or update in future we determine whether the user is credible or not. Thus scores of all users are set to 1. The Incredulity table is updated each time a user executes query.

Consider a user having U_{id} 1001 who has a deviation from normal query equal to 0.81. Then the updated Incredulity table is as shown in Table 4 as $\sqrt{\text{ICS} * \phi_i}$. In the table ICS represents the deviation from normal query and ϕ_i denotes the Initial Incredulity score.

3.2 Learning phase

Our model aims to generate user-profiles from transaction-logs, and quantifies deviation from normal behaviour. In particular, this phase aims to recognise and characterise the user activity pattern on the basis of their queries arrangement. Figure 1 illustrates various components of the architecture of the proposed model.

3.2.1 Training data

A transaction log is a sequential record of all the changes which are made to the database and contains enough information to such that all the changes that are made to a data file by the transaction can be undone. The log records the beginning of a transaction, all the changes that are a part of it, and then the final commit or rollback of the transaction. Every database has to the minimum, one physical transaction log and one data file that is exclusive to the database for which it was made. Our initial input to the learning phase algorithm is the transaction log comprising of only authorised and consistent transactions. This data is free of any unauthorised activity and is used to form user profiles, role profiles etc based on normal user transactions. The logs are scanned to extract SQL queries and the User ID of the users who executed them.

3.2.2 SQL query parser

This module takes SQL queries as input, parses them and produces read and write sequences as output. The query parser also assigns a unique Transaction ID. The

Table 4 Initial incredulity score

U_{id}	ϕ_i	$\sqrt{\text{ICS} * \phi_i}$
1001	1	0.9
1002	1	1
1003	1	1
1004	1	1
1005	1	1

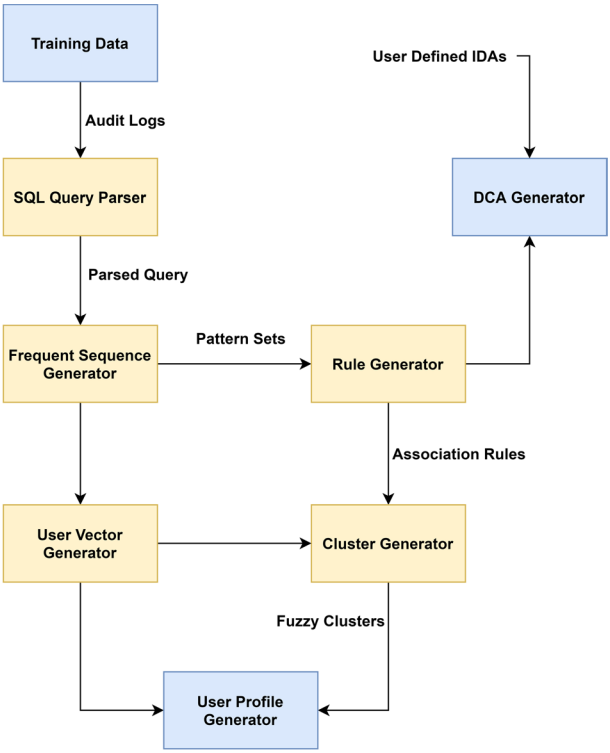


Fig. 1 Our proposed learning phase architecture

final output consists of: Transaction (T_{ID}), User ID U_{ID} and the read and write sequence generated by the parsing algorithm. As an example consider the 10 transactions shown in Table 5. For each transaction, the table comprises of the ID of the user executing the transaction, its transaction ID and its corresponding read-write sequence.

The operations $r(x)$ and $w(x)$ represent write and read operations, respectively. Data items are represented by English without loss of generality. Frequent Sequences are generated on the basis of minimum support as described in the next sub-section.

3.2.3 Frequent sequences generation

The read and write sequences generated by the SQL query parser are given as inputs to the frequent sequences generator. These sequences are first pre-processed where they are assigned weights according to type of data items, for instance the IDAs are given greater weight as compared to DCAs and other normal attributes. This module uses the prefix span algorithm [30] to generate frequent sequences out of the input sequences corresponding to each U_{ID}. Prefix span uses pattern growth

Table 5 User access sequences

User Id	Transaction Id	Transaction
2001	101	$< r(n), r(a), r(m), w(f), r(a), w(e) >$
2002	102	$< r(a), r(f), w(a), r(e), r(f), w(e) >$
2001	103	$< r(n), r(m), r(b), w(e), r(n), r(c), w(m), r(a), r(m), w(b), r(c), r(f), r(b), w(f) >$
2003	104	$< r(f), r(m), w(f), r(f), w(e) >$
2003	105	$< r(b), w(b), r(e), r(n), w(c), r(m), w(f), r(a), w(e) >$
2001	106	$< r(m), r(f), w(f), r(c), r(e), w(e), r(c), w(n) >$
2002	107	$< r(f), r(c), r(m), w(n) >$
2004	108	$< r(e), w(m) >$
2003	109	$< r(b), r(f), w(m) >$
2004	110	$< r(m), r(a), w(c), r(a), w(m), r(b), r(n), r(e), w(b) >$

method to extract frequent sequences, avoiding a huge candidate sequence generation leading to better memory utilization and execution time. Table 6 shows thirteen sequential patterns which fulfill the support constraint. For example, sequential pattern $< r(m), w(f), w(e) >$ is supported by transactions 1, 4, 3, and 6. An example of a sequence that does not fulfill minimal support is the sequence $< r(n), w(a), r(m) >$ that is only supported by transaction $< 1 >$. These frequent sequences are then used to mine sequential rules as shown in the next subsection.

3.2.4 Rule generation

The frequent sequences are given as inputs to the rule generator module which uses association rule mining to generate read rules and write rules from these frequent

Table 6 Sequential patterns where support exceeds threshold value

Sequential patterns
$r(c)$
$w(m)$
$r(a), w(e)$
$r(b), w(b)$
$r(b), w(n)$
$r(e), w(e)$
$r(f), w(e)$
$r(f), w(f)$
$r(m), r(f)$
$r(m), w(f), w(e)$
$r(n), r(m), r(a)$
$r(n), r(m), w(e)$
$r(n), r(m), w(f)$

sequences. Table 7 consists of read and write rules returned by the rule generation module.

By analyzing the sequential patterns which are mined in Table 5, a few examples can be utilized for finding data dependencies while others ought not be considered. To begin with, the sequential patterns like $< r(c) >$ just contains the operation on a datum, has no data dependency and they should not be considered. Second, some sequential patterns mined just contain read operations. Since we are concerned about malignant modifications to data items by user transactions, we focus on the transactions containing write operation(s). So patterns containing only read operations are likewise ignored. Data dependency rules are classified as read rules and write rules. The accompanying methodology is used to generate data dependency rules. For all sequential patterns $< r(x), r(y), ..., r(z), w(a) >$ in read sequence set, generate the read rules with the format $w(a) \rightarrow r(x), r(y), ..., r(z)$. In the event that confidence of the rule is larger than the minimum required confidence, it is added to the read rules set. Similarly for all sequential patterns $w(1)$, $w(2)$, $w(3)$, ..., $w(n)$ in the write sequence set, generate the write rules with the format $w(n) \rightarrow w(1), w(2), ..., w(3)$. If the confidence of the rule is larger than the minimum confidence, then it's added in the answer set of write rules. After the read and write rules are generated, they can be used for identifying malicious transactions by using the database log. The transactions that made modifications to the database without following data dependency rules are flagged as malicious transactions using the following procedure: for all database transaction that consist of write operations(s), confirm whether each write operation adheres to data dependency rule, i.e., whether this transaction read suitable data items before this update activity and whether it wrote proper data items after this update activity.

Table 7 Read and write rules

Read rule
$r(a)$ -> $w(e)$
$r(b)$ -> $w(b)$
$r(b)$ -> $w(n)$
$r(e)$ -> $w(e)$
$r(f)$ -> $w(e)$
$r(f)$ -> $w(f)$
$r(m)$ -> $r(f)$
$r(m)$ -> $w(f)$
$r(n)$, $r(m)$ -> $r(a)$
$r(n)$, $r(m)$ -> $w(e)$
$r(n)$, $r(m)$ -> $w(f)$
Write rule
$W(f)$ -> $w(e)$

Table 8 Classification of attributes

IDA	DCA	Normal
n, *e*	*a*, *b*, *f*, *m*	*c*

Table 9 Critical rules

Critical rules
r(*a*) -> *w*(*e*)
r(*b*) -> *w*(*n*)
r(*e*) -> *w*(*e*)
r(*f*) -> *w*(*e*)
r(*n*), *r*(*m*) -> *r*(*a*)
r(*n*), *r*(*m*) -> *w*(*e*)
r(*n*), *r*(*m*) -> *w*(*f*)

3.2.5 DCA generation

In our approach, we semantically define a class of data items known as Integral Data Attributes or IDAs. These IDAs and rules are given as input to our directly correlated attributes (DCA) generator which specifies all those elements as DCA which are present in either the antecedent or the consequent of those rules that involve at least one of the IDAs. As an example, Table 8 shows the distribution of data items into the aforementioned classes. Once the Integral Data Attributes are known, the model uses them to generate critical rules (Definition 8). Table 9 shows the critical rules mined using access pattern recognition.

Algorithm 1 DCA Generator

Data : IDA,
Set DCA = {},
RR = Set of Read Rules,
WR = Set of Write Rules

Result : The set of Directly Correlated Attributes DCA

Function: DCA Generator (IDA, RR, WR)

```
1  begin
2      Initialization
3      for Ω ∈ RR ∪ WR do
4          for α ∈ Ω do
5              if α ∈ IDA then
6                  while β ∈ Ω do
7                      DCA{} ← β
8                  end
9              end
10         end
11     end
12 end
```

3.2.6 User vector generation

Using the frequent sequences for the given audit period this module generates user vectors which are representative of a user's activity. A user vector is of the form

$$B_{ID} =< U_{ID}, w_1, w_2, w_3, ...w_n > ; \quad \text{where}, w_i = |O(a_i)| \tag{2}$$

$|O(a_i)|$ represents the total number of times user with the given U_{id} performs operation ($O \in \{R, W\}$) on the attribute a_i in the pre-defined audit period. An audit period τ refers to a period of time such as one year, a time window $\tau = [t_1, t_2]$ or the most recent 10 months. Each of these w_i would represent how frequently a user performs the operation on the particular data item.

$$\text{UV}_{ID} =< U_{ID}, < p(a_1), p(a_2), \ldots, p(a_n) > ; \quad \text{where}, p(a_k) = \frac{w_k}{\sum_{w_j \in B_i} w_j} \tag{3}$$

$p(a_k)$ is defined as the probability of accessing the attribute a_k.

Value of close to 1 would mean that the user accesses the given attribute frequently. As an example, consider Table 10 which gives a count of operations

Table 10 User vector generator

User Id	w_a	w_b	w_c	w_e	w_f	w_m	w_n
2001	3	3	4	4	5	5	3
2002	2	0	1	2	3	1	1
2003	1	3	1	3	5	3	1
2004	2	2	1	2	0	3	1

performed on data item by a specific user. This table is then used as input to calculate the probability of access which in turn helps in generating each user's profile vector as shown in Table 11.

3.2.7 Cluster generation

Cluster generator takes user vectors and rules as input and generates fuzzy clusters. Users are clustered into different fuzzy clusters based on the similarity of their user vectors using the Fuzzy c-means [6, 13]. clustering algorithm A cluster profile would include $C_i =< C_{ID}, \{R\} >$ where, C_{ID} represents the cluster centroid, and $\{R\}$ is a set of rules which is formed by taking the union of all the rules that the members of the given fuzzy cluster follow. Each user belongs to a cluster to a certain degree w_{ij}, where w_{ij} represents the membership coefficient of the i_{th} user (u_i) with the j_{th} cluster. The center of a cluster (α) is the mean of all points, weighted by their membership coefficients. Mathematically,

$$\begin{aligned} w_{ij} &= \frac{1}{\sum_{k=1}^{C} \| \frac{u_i - \alpha_j}{u_i - \alpha_k} \|^{\frac{2}{i-1}}}; \\ \text{where, } \alpha_k &= \frac{\sum_u w(u)^m u}{\sum_u w(u)^m} \end{aligned} \tag{4}$$

The objective function that is minimized to create clusters is: [28]

$$\text{argmin} \sum_{i=1}^{n} \sum_{j=1}^{C} w_{ij}^m \|u_i - \alpha_j\|^2 \tag{5}$$

where n is the total number of users, C is the number of clusters, and m is the fuzzifier. The dissimilarity or distance function used in the formation of fuzzy clusters is the modified Jenson Shannon distance [14] which is given as:

Table 11 User profile vector

<2001, <0.11, 0.11, 0.15, 0.15, 0.185, 0.185, 0.11>>
<2002, <0.2, 0, 0.1, 0.2, 0.3, 0.1, 0.1>>
<2003, <0.058, 0.176, 0.068, 0.176, 0.29, 0.176, 0.058>>
<2004, <0.2, 0.2, 0.1, 0.2, 0, 0.3, 0.1>>

Table 12 Profile of user U_{1001}

Input					Output
c1	c2	c3	c4	User vector	User profile
0.2	0.2	0.2	0.4	<U_{1001},0.2,0.109,0.9, 0.6>	<U_{1001},<0.2,0.1,0.9,0.6>, <0.2,0.2,0.2,0.4 >>

Given two user vectors,

$$\mathrm{UV}_x =< U_x, < p_x(a_1), p_x(a_2), p_x(a_3), \ldots p_x(a_n) >> \tag{6}$$

and

$$\mathrm{UV}_y =< U_y, < p_y(a_1), p_y(a_2), p_y(a_3), \ldots p_y(a_n) >> \tag{7}$$

of equal length n, the modified Jensen Shannon distance is computed as in equation (7),

$$A = (1 + p_x(a_i) * w(a_i)) \log_2 \frac{1 + p_x(a_i) * w(a_i)}{1 + p_y(a_i) * w(a_i)} \tag{8}$$

$$B = (1 + p_y(a_i) * w(a_i)) \log_2 \frac{1 + p_y(a_i) * w(a_i)}{1 + p_x(a_i) * w(a_i)} \tag{9}$$

$$D(\mathrm{UV}_p \| \mathrm{UV}_q) = \sum_{i=1}^{n} \frac{A + B}{2} \tag{10}$$

where, $w(a_i)$ is the semantic weight associated with the $a_{i^{th}}$ attribute.

3.2.8 User profile generator

The User profile generation module takes user vectors and the cluster profiles as input and generates user profiles of the form:

$U_i =< U_{ID}, < p(a_1), p(a_2), p(a_3) \ldots p(a_k) >, < c_1, c_2, \ldots c_C >>$ Where U_{ID} is a unique ID given to each user, $< p(a_1), p(a_2), p(a_3), \ldots (a_n) >$ is a vector containing the probability of the user accessing a particular attribute, and $< c_1, c_2, \ldots c_C >$ is a vector representing the membership coefficients of the given user for C different clusters. For instance, consider a system with 4 fuzzy clusters and 4 attributes, the given Table 12 illustrates the profile of user U_{1001}.

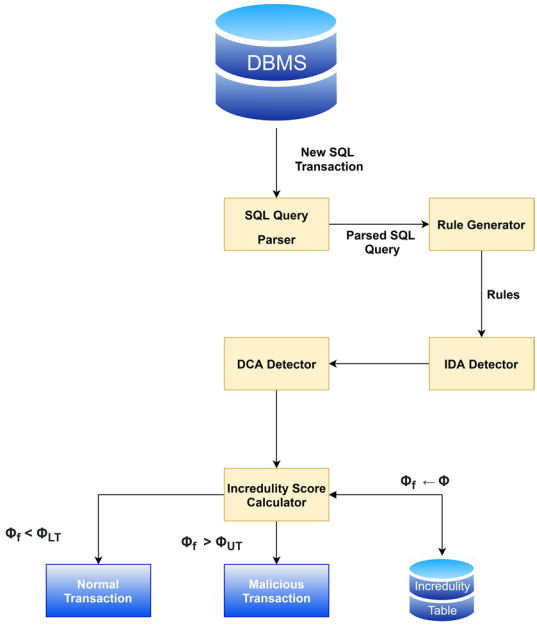


Fig. 2 Our proposed testing phase architecture

3.3 Testing phase

In this phase, we test the efficacy of the trained model for detecting malicious queries. A user query is obtained as input and it is compared with the user's historical access patterns to see if the test transaction is malicious. The model first checks whether the user is trying to access a IDA. If yes, the transaction is allowed only if the given user has accessed that IDA before. Next, it checks if any DCA is being accessed. A user can perform write operation on a DCA only if it was previously written by the same user, otherwise the transaction is termed as malicious. Next, we check if the transaction abides by the rules that are generally followed by similar users. Figure 2 illustrates the architecture of the testing phase.

3.3.1 Rule generator

This module takes the read-write sequences generated by the SQL query parser and outputs rules followed by the input transaction. It uses prefix span algorithm [30] for sequential pattern mining to generate rules. This can be a read rule or a write rule and indicates the operations done by the user, data attributes accessed by the user and the order in which they are accessed. Now this rule is used to check for deviation from the regular behaviour.

3.3.2 IDA detector

The semantically critical elements referred to in our approach as IDAs are detected in this module. The read-write rule corresponding to the incoming transaction is checked for the presence of IDAs.

Algorithm 2 IDA Detector

Data : Set of rules (γ) from test transaction, Set χ_{IDA}, U_{ID}, User Profile(Θ)

Result : Checks whether the test transaction is malicious or normal with respect to IDA

```
1  begin
2      Initialization
3      for ρ ∈ γ do
4          for λ ∈ ρ do
5              if λ ∈ χ_IDA then
6                  if w(λ) ∈ ρ & Θ[U_ID][w(λ)] == 0
                    then
7                      Raise Alarm;
8                  end
9                  if r(λ) ∈ ρ & Θ[U_ID][r(λ)] == 0 &
                    Θ[U_ID][w(λ)] == 0 then
10                     Raise Alarm;
11                 end
12             end
13         end
14     end
15 end
```

If the rule being checked for intrusion contains a IDA, then it is dealt with using the following policy:

1. If read operation has been performed on any IDA, i.e., *r*(IDA) is present in the rule and UV[*i*][*r*(IDA)] = 0 and UV[*i*][*w*(IDA)] = 0 for the given user, then the transaction is termed as malicious.
2. If write operation has been performed on any IDA i.e., w(IDA) is encountered and UV[*i*][*w*(IDA)] = 0 for the given user, then the transaction is termed as malicious.

Algorithm 3 DCA Detector

Data : Set of rules (γ) from test transaction, Set χ_{DCA}, U_{ID}, User Profile(Θ)

Result : Checks whether the test transaction is malicious or normal with respect to DCA

```
1  begin
2      Initialisation
3      for ρ ∈ γ do
4          for λ ∈ ρ do
5              if λ ∈ χ_DCA then
6                  if w(λ) ∈ ρ & Θ[U_ID][w(λ)] == 0
                    then
7                      Raise Alarm;
8                  end
9              end
10         end
11     end
12 end
```

3.3.3 DCA detector

This module addresses the issue of inference attacks on IDAs. As discussed earlier, certain data elements can be used to access IDAs, i.e., first order inference. This module uses the rules mined in the learning phase to determine which elements can be used to directly infer the IDAs. Our system seeks to prevent inference attacks by closely monitoring write operations on DCAs. If write operation has been performed on any DCAs i.e., w(DCA) is present in the rule to be checked and UV[i][w(DCA)] = 0 for the given user, then the transaction is termed as malicious.

3.3.4 Incredulity score calculator and analyzer

If the transaction has not been found malicious in the previous two modules, we check if the transaction is malicious based on the previous history of the user and the behaviour pattern of all similar users. To do so, we maintain a record of action of all users using their Incredulity Scores.

The deviation of a user's new transaction with his normal access pattern is referred to as incredulity, and the corresponding measure of incredulity is the incredulity score. A user who is a potential threat tends to have a high incredulity score. We use ICS to keep a track of the maximum similarity of the given rule. We combine ICS with ϕ_i

(Initial incredulity score) to get the final measure of incredulity score ϕ_f for the given user. We define 2 thresholds ϕ_{LT} and ϕ_{UT}. ϕ_{UT} represents the upper limit for the incredulity score of a non-malicious user whereas ϕ_{LT} denotes the lower limit. This means that if ϕ_f for a user comes out to be greater than ϕ_{UT}, the user is malicious. On the other hand, ϕ_f value less than ϕ_{LT} denotes a benign user. In particular, if the incoming rule (R_1) is a write rule, then the consequent of the incoming rule is matched with the corresponding rules in the user's cluster. A user is said to be the part of the i_{th} cluster if $\mu_i > \delta$.

where, μ_i is the fuzzy membership coefficient of the given user for the i_{th} cluster and δ is a user defined weight. Otherwise, if the incoming rule (R_1) is a read rule, then the antecedent of the incoming rule is matched with the corresponding rules in the cluster of which a user is as part.

In order to quantitatively measure the similarity between two rules, we use weighted Jaccard distance:

$$\mathrm{JD} = 1 - \frac{\delta_1(R_1 \cap R_2) - \delta_2(R_1 \cup R_2 - R_1 \cap R_2)}{R_1 \cup R_2} \tag{11}$$

Algorithm 4 Weighted Jaccard Distance

Data : Rules R_1, R_2, δ_1, δ_2,
Set χ_{R_1}, χ_{R_2}
Result : Distance between the two rules (τ)
Function: jcDistance (R_1, R_2)

```
begin
    for ρ ∈ γ do
        for Ω ∈ R1 do
            χR1 ← Ω
        end
        for Ω′ ∈ R2 do
            χR2 ← Ω′
        end
        τ = 1 − (δ1(χR1 ∩ χR2) − δ2(χR1 ∪ χR2 − χR1 ∩ χR2)) / (χR1 ∪ χR2)
        return τ;
    end
end
```

$\forall R_2, \mu_i > \delta$ and $i \in [1, k]$ The minimum value of JD is regarded as ICS. ϕ_i is fetched directly from incredulity table. Final incredulity score for the given user is calculated as:

$$\phi_f = \sqrt{\mathrm{ICS} * \phi_i} \tag{12}$$

1. If $\phi_f < \phi_{\text{LT}}$, the transaction is termed as non-malicious. In this case, the current incredulity score in the incredulity table for the given user is reduced by a factor known as "amelioration factor ($\mathring{A}$)". Thus, ϕ_i is updated as $\phi_i = \mathring{A}\phi_i$
2. If $\phi_{\text{UT}} < \phi_f \geq \phi_{\text{LT}}$, the transaction is termed as non-malicious and the incredulity table entry for the given user is updated with ϕ_f.
3. If $\phi_f \geq \phi_{\text{UT}}$ the transaction is termed as malicious.

As a concrete example let us consider the initial incredulity score values of various users in the organization as shown in Table 13.

In addition, the minimum value of ICS for each user is shown in Table 14. Then the incredulity score can be calculated using Equation number 9 as shown in Table 15.

If we consider ϕ_{LT} as 0.3 and ϕ_{UT} as 0.6, the nature of each transaction by the users and the updated incredulity scores ϕ_f are summarized in Table 16.

Table 13 Initial incredulity score of each table

U_{id}	ϕ_i
1001	0.9
1002	0.8
1003	0.2
1004	0.6
1005	0.7

Table 14 Minimum incredulity score corresponding to each user

U_{id}	ICS
1001	0.2
1002	0.3
1003	0.2
1004	0.6
1005	0.3

Table 15 incredulity score for users

U_{id}	$\phi_f = \sqrt{\text{ICS} * \phi_i}$
1001	0.42
1002	0.49
1003	0.2
1004	0.6
1005	0.46

Table 16 Final classification corresponding to each user

U_{id}	ϕ_f	Nature of transaction	Updated ϕ_f
1001	0.42	Non-malicious	0.42
1002	0.49	Non-malicious	0.49
1003	0.2	Non-malicious	0.198
1004	0.6	Malicious	0.6
1005	0.46	Non-malicious	0.46

4 Experimentation

In this section we describe the testing methodology for our proposed algorithm. First, we discuss the dataset used for evaluation in the experiments. Second, we evaluate the performance of our approach on different accuracy metrics and compare our performance with the pertinent works.

4.1 Description of dataset

4.1.1 Synthetic dataset

The primal requirement for the performance evaluation of an algorithm in the field of intrusion detection is the presence of a standard dataset. However, a dataset which consists of transactions occurring in a pragmatic setting does not exists, as organisations are bound under confidentiality constraints. Hence, we created our own synthetic dataset, mimicking real world transactions by using TPC-C standard benchmark [44].

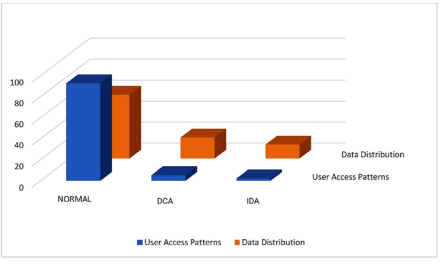
To assess our approach on a wide gamut of attacks, we generated two different sets of malicious transactions. First set of transactions were generated randomly with the assumption that the attacker may be unaware of normal database dependencies. This set is generated by replacing attributes randomly of legitimate queries, which constitute the usual behavior and by replacing legitimate transactions of each.

The second set of transactions were generated such that they do not conform to the average user activity and are indicators of users trying to perform transactions outside of their scope and permissions.

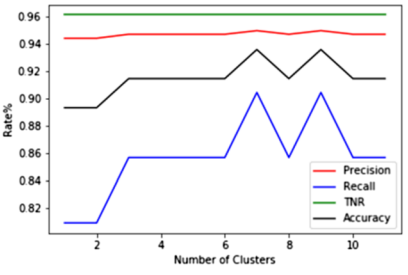
4.1.2 Normal transactions

The performance of the algorithm is inspected by conducting out several tests on a credit card company dataset adhering to the TPC-C benchmark [44]. The TPC-C schema comprises of a combination of both read only and read and write transactions that try to replicate the activities found in complex OLTP application environment. The database schema, data population and implementation rules are devised in a manner so as to replicate modern OLTP systems.

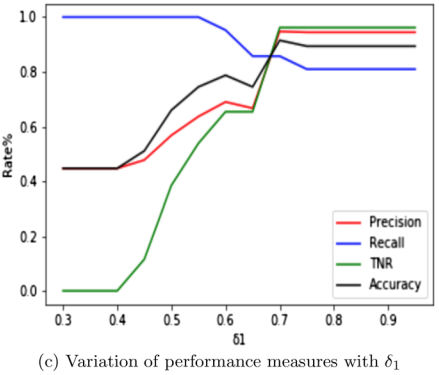
We use two audit logs: training log, testing log. The training log comprises of legitimate user transactions and testing log consists of a mixture of legitimate as well as malicious user transactions. To mimic the presence of unusual records in the real-world dataset, we have also injected some anomalies for detection. The injected anomalies are altered in a particular manner from legitimate transactions making them non-trivial to be detected. In conclusion, about 20,000 transactions were used where, about 99% of data was non-malicious while less than 1% of data was malicious. The reason for this imbalance can be accounted to the fact that we are trying to replicate the original database in a banking organization. The percentage of malicious transaction in a real-time database is very small when compared to the non-malicious transactions. Figure 3a highlights the distribution of the dataset used for analysis. The details of IDAs, DCAs and Normal data items has already been given in Sect. 3 and examples have been discussed in Sect. 5. The access pattern data hereby shows that IDAs are seldom examined, that too only by a few user roles and hence, protection of IDAs from malicious access is of major importance as compared to DCAs and Normal data elements.



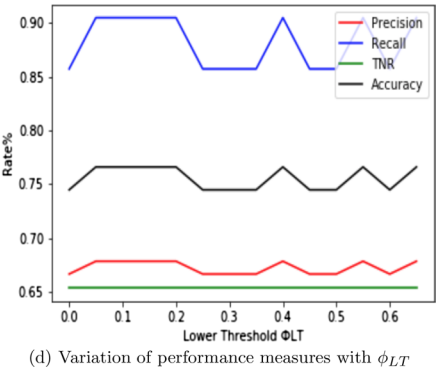
(a) Frequency of data items and their access frequency



(b) Variation of performance with number of clusters



(c) Variation of performance measures with δ_1



(d) Variation of performance measures with ϕ_{LT}

Fig. 3 Variations in performance measures with different parameters

4.2 Cluster analysis

When the number of user(s) roles go beyond a given threshold, it becomes infeasible for an IDS to keep track of individual user access patterns for detecting anomalies. Therefore, to improve the efficacy of the IDS while minimizing the performance loss, we have used clustering-based approach.

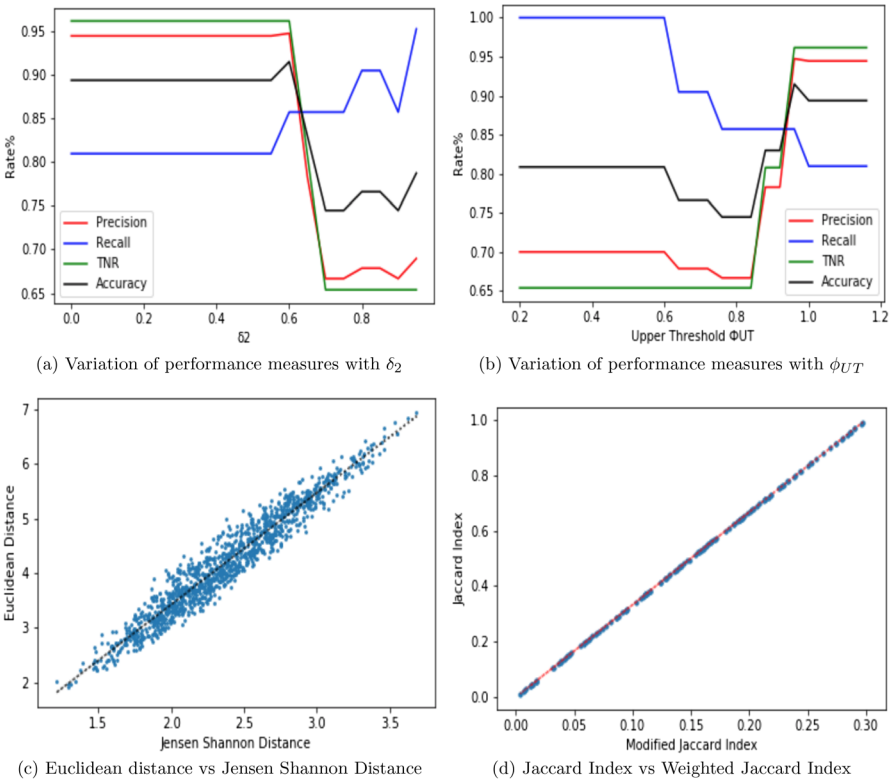
We have selected Fuzzy clustering over other clustering techniques. Fuzzy clustering is a method of clustering in which each data item can belong to more than one cluster; whereas, in non-fuzzy clustering, data is distributed into distinct clusters and each data point belongs to exactly one cluster. Membership grades are allocated to each of the data point in fuzzy clustering. These membership grades denote the extent to which data points belong to each cluster. While evaluating the performance by varying the number of clusters formed, it is observed that a certain number of clusters results stands out in performance. Figure 3b depicts variation in precision, recall, TNR, accuracy with change in number of clusters. It can be observed that TNR and Precision remain almost constant and are invariant with respect to the number of clusters. On the other hand recall and accuracy have similar trends with minute change in the magnitudes. It can be seen that there is an overall increasing trend in recall and accuracy as number of clusters increases. Both recall and accuracy attain their optimal values for number of clusters 7 and 9.

4.3 Distances and thresholds

In Sect. 3.2, we have described Modified Jensen-Shannon distance as a measure to calculate distance between two user vectors of same length. In applied math and statistics, the Jensen-Shannon divergence is a technique of estimating the similarity between two distributions. It's additionally called info radius (IRad) or total divergence to the average. It's supported the Kullback–Leibler divergence, with some notable (and useful) variations, together with that it's comparable and it's continually a finite value. The root of the Jensen–Shannon divergence is a metric typically known as Jensen-Shannon distance. We preferred to use modified Jensen-Shannon distance to give weights to data attributes and avoid the curse of dimensionality. The variation of modified Jensen–Shannon distance with Euclidean distance is shown in Fig. 4c.

In Sect. 3.2.1, we have defined weighted Jaccard distance to quantitatively measure the similarity between two rules. The Jaccard index is a statistical measure used for comparing the diversity and similarity of sample sets. The Jaccard coefficient measures the similarity between finite sample sets, and is computed as the size of the intersection by the size of the union of the sample sets. The variation of weighted Jaccard index with Jaccard index is shown in Fig. 4d.

The variation of precision, recall, TNR, accuracy with the various thresholds, namely δ_1, δ_2, ϕ_{UT}, ϕ_{LT} (defined in Sect. 3) is shown in Figs. 3c to 4b. Figure 3c represents the variation of Precision, recall, TNR, accuracy with δ_1. It can be observed from the graph that Precision, TNR and Accuracy increase with the



(a) Variation of performance measures with δ_2

(b) Variation of performance measures with ϕ_{UT}

(c) Euclidean distance vs Jensen Shannon Distance

(d) Jaccard Index vs Weighted Jaccard Index

Fig. 4 Variations in performance measures with respective variations in distances

increase in value of δ_1, while the value of Recall decreases with increase in value of δ_1. The variation of Precision, recall, TNR, accuracy with δ_2 has been outlined in Fig. 4a. It can be inferred from the graph that the value of Precision, TNR and Accuracy starts decreasing when the value of δ_2 increases beyond a certain value. Recall, on the other hand, increases for higher values of δ_2. Figure 4b further highlights the variation of precision, recall, TNR, accuracy with ϕ_{UT}. It can be concluded that the value of Precision first decreases and then exponentially increases with the increase in value of ϕ_{UT}. An identical trend is followed by Accuracy. Somewhat similar trend is followed by TNR except that it does not decrease initially. On the contrary, the value of Recall decreases with the increase in value of ϕ_{UT}. Figure 3d shows the variation of Precision, recall, TNR, accuracy with ϕ_{LT}. It can be observed from the graph that the values of all the parameters fluctuate a little but remain more or less constant with the increase in value of ϕ_{LT} .

With regards to the dataset we have used, following inferences can be drawn from the graphs:

1. Value of δ_1 should be close to 0.65 for optimum performance.

2. Value of δ_2 should be close to 0.55 for optimum performance.
3. Value of ϕ_{UT} should be close to 0.59 for optimum performance.
4. Value of ϕ_{LT} should be close to 0.2 for optimum performance.

4.4 Comparison with related methods

Precision is defined as the ratio of correctly identified malicious transactions known as true positives (TP) to the total number of transactions classified as malicious from the database logs, a combination of false positives (FP) and true positives (TP). Recall (TPR) is defined as the ratio of correctly identified malicious transactions known as true positives (TP) to the total number of malicious transactions existing in the database logs, a combination of false negatives (FN) and true positives (TP). F1-score is defined as the harmonic mean of precision and recall. Accuracy is defined as the ratio of correctly identified transactions i.e., true positives and true negatives (TP + TN) to total number of transactions (TP + FP + TN + FN). Using these performance measures, we will also compare our approaches with other related works. Our different approaches are as follows:

Approach 1. Our approach using modified Jenson–Shanon distance and weighted Jaccard index.

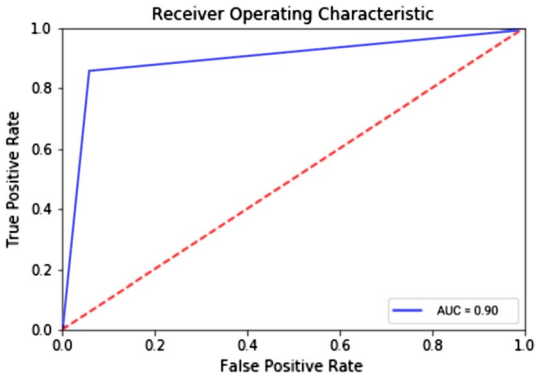
Approach 2. Using unweighted Jaccard index with Jenson–Shanon distance.

Approach 3. Using Euclidean distance with unweighted Jaccard index.

Table 17 shows the comparison of our proposed approaches in terms of Precision (TPR), Accuracy (ACC) and F1 score. It can be observed that approach 1 outperforms other approaches in all the three aspects. The main reason behind the effectiveness of Approach 1 lies in its fundamental idea of using modified Jenson–Shanon Distance and weighted Jaccard Index both of which are able to capture fine characteristics of the user behaviour. It also provides a comparative analysis of our approach with prior methodologies based on the True Positive Rate, Accuracy and F1 Score. The performance of all three approaches was

Table 17 Comparison of our approaches with related works

Sensitivity measures	TPR	ACC	*F*1 score
Approach 1	0.96	0.94	0.94
Approach 2	0.73	0.80	0.85
Approach 3	0.74	0.83	0.85
Hu and Panda [17]	0.73	0.81	0.79
Kim and Cho [21]	0.88	0.93	0.90
Hashemi et al. [16]	0.71	0.84	0.82
Majumdar et al. [40]	0.70	0.80	0.78
Sun et al. [43]	0.91	0.93	0.92
Subudhi and Panigrahi [42]	0.92	0.92	0.93
Bu and Cho [7]	0.93	0.92	0.92
Kim and Cho [22]	0.93	0.94	0.94
Bu et al. [8]	0.85	0.88	0.86



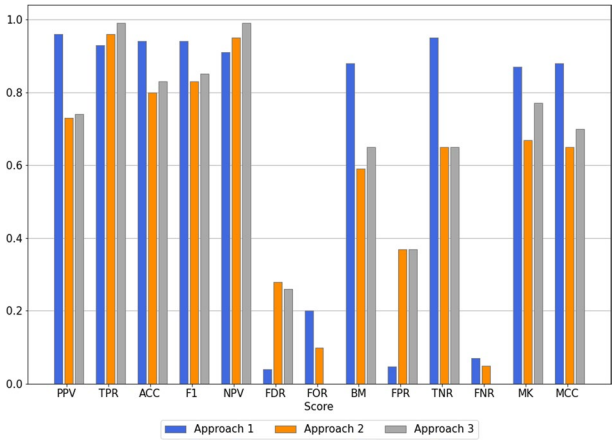


Fig. 5 ROC curve

Fig. 6 Comparison of approaches on basis of various sensitivity metrics

assessed using the aforementioned metrics. The evaluation results show that our algorithm outperforms all the compared methodologies. This can be ascribed to the low sensitivity of our algorithm to changes in user patterns since we incorporate user behaviour analysis in our approach.

It is interesting to note that Approach 3, although slightly, is able to perform better as compared to Approach 2. Approach 3 employed Euclidean distance whereas Approach 2 used Jenson–Shanon Distance which is known to deliver better results over Euclidean distance yet the former performed better. This shows that the modification introduced in Jenson–Shanon Distance is imperative and has enhanced the overall performance of the model (Fig. 5).

Figure 6 shows a detailed bar graph representing a comparison of three proposed approached on the basis of the magnitudes of various performance measures generated during the testing phase.

Our approach consolidates the advantages of statistical Integral Data Attribute analysis and user behaviour analysis-based anomaly based detection methods, resulting in a reduction in both false positive and false negative errors.

5 Conclusion and future work

We have devised Trust factor-based user behavior intrusion detection system (TFUBID), a two-phase database intrusion detection system that clusters users based on their attribute access patterns and the types of queries that they implement. The efficacy of our model lies in its ability to prevent illegitimate retrieval and modification of most sensitive data elements called integral data attributes (IDA) and the attributes that are directly correlated with them. To avoid data breaches, TFUBID uses sequential pattern mining to ensure that each user follows a particular query pattern for accessing IDAs, i.e., the user associates himself with his regular access behaviour. Any deviation from the regular arrangement may lead to depreciation of trust factor of the user's profile, indicating user's malicious intent. Our extensive evaluation shows that TFUBID is able to outperform other approaches in terms of various performance measures. One of the main applications of our research work is within organizations, where the system can be used to monitor the behavior of employees and detect any deviation from their regular access patterns that may indicate a malicious intent. This can help organizations maintain the confidentiality and integrity of their data, and prevent financial and reputational damage. Additionally, TFUBID can also be used in other settings such as financial and medical sectors to protect customer sensitive data from unauthorized access and malicious updates. The system can also be integrated with other security systems such as firewalls to provide a more comprehensive security solution.

In future, for making our model more robust, we would like to incorporate the temporal dimension of data access along with patterns mined from inter-query dependencies. We are also planning to introduce more rigorous measures to account for the sensitivity of operations which has been proven to directly impact the performance.

Author contributions IS: Conceptualization, Methodology, Writing, Validation, Review and Editing. RJ: Conceptualization, Supervision, Methodology, Validation, Review and Editing.

Funding The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Data availability The datasets generated and/or analzsed during the current study are available in http://www.tpc.org/tpcc/default.asp.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with human participants performed by any of the authors.

Informed consent This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Agrawal R, Srikant R (1994) Fast algorithms for mining association rules in large databases. In: Proceedings of the 20th International Conference on Very Large Data Bases, VLDB '94. Morgan Kaufmann Publishers Inc, San Francisco, pp 487–499
2. Agrawal R, Srikant R (1995) Mining sequential patterns. In: Proceedings of the Eleventh International Conference on Data Engineering, ICDE '95. IEEE Computer Society, pp 3–14
3. Aljawarneh S, Aldwairi M, Yassein MB (2018) Anomaly-based intrusion detection system through feature selection analysis and building hybrid efficient model. J Comput Sci 25:152–160
4. Bertino E, Sandhu R (2005) Database security—concepts, approaches, and challenges. IEEE Trans Depend Secure Comput 2(1):2–19
5. Bertino E, Terzi E, Kamra A, Vakali A (2005) Intrusion detection in RBAC-administered databases. In: 21st Annual Computer Security Applications Conference (ACSAC'05), pp 10–182
6. Bezdek JC, Ehrlich R, Full W (1984) FCM: the fuzzy c-means clustering algorithm. Comput Geosci 10(2):191–203
7. Bu SJ, Cho SB (2020) A convolutional neural-based learning classifier system for detecting database intrusion via insider attack. Inf Sci 512:123–136
8. Bu S-J, Kang H-B, Cho S-B (2022) Ensemble of deep convolutional learning classifier system based on genetic algorithm for database intrusion detection. Electronics 11(5):745
9. Cappelli D, Moore A, Trzeciak R, Shimeall T (2006) Common sense guide to prevention and detection of insider threats 3rd edition-version 3.1
10. Chung CY, Gertz M, Levitt K (2000) DEMIDS: a misuse detection system for database systems. Springer, Boston, pp 159–178
11. Corney M, Mohay GM, Clark AJ (2011) Detection of anomalies from user profiles generated from system logs. In: Proceedings of the Ninth Australasian Information Security Conference, Volume 116 of AISC '11. Australian Computer Society, Inc., pp 23–32
12. Denning DE (1987) An intrusion-detection model. IEEE Trans Softw Eng 13(2):222–232
13. Dunn JC (1973) A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters. J Cybern 3(3):32–57
14. Fuglede B, Topsoe F (2004) Jensen–Shannon divergence and Hilbert space embedding. In: International Symposium on Information Theory, 2004. ISIT 2004. Proceedings, p 31
15. Furnell S (2004) Enemies within: the problem of insider attacks. Comput Fraud Secur 2004(7):6–11
16. Hashemi S, Yang Y, Zabihzadeh D, Kangavari M (2008) Detecting intrusion transactions in databases using data item dependencies and anomaly analysis. Expert Syst 25(5):460–473
17. Hu Y, Panda B (2004) A data mining approach for database intrusion detection. In: Proceedings of the 2004 ACM Symposium on Applied Computing, SAC '04. Association for Computing Machinery, New York, pp 711–716
18. Hu Y, Panda B (2005) Design and analysis of techniques for detection of malicious activities in database systems. J Netw Syst Manag 13(3):269–291
19. Jindal R, Singh I (2022) Detecting malicious transactions in database using hybrid metaheuristic clustering and frequent sequential pattern mining. Cluster Comput 25(6):3937–3959
20. Kim J, Jung H, Kim W (2022) Sequential pattern mining approach for personalized fraudulent transaction detection in online banking. Sustainability 14(15):9791
21. Kim T-Y, Cho S-B (2019) CNN-LSTM neural networks for anomalous database intrusion detection in RBAC-administered model. In: Gedeon T, Wong KW, Lee M (eds) Neural information processing. Springer, Cham, pp 131–139
22. Kim TY, Bae Cho S (2021) Optimizing CNN-LSTM neural networks with PSO for anomalous query access control. Neurocomputing 456:666–677
23. Kumar S, Mohbey KK (2022) A review on big data based parallel and distributed approaches of pattern mining. J King Saud Univ Comput Inf Sci 34(5):1639–1662
24. Kumar S, Mohbey KK (2022) A utility-based distributed pattern mining algorithm with reduced shuffle overhead. IEEE Trans Parallel Distrib Syst 34(1):416–428
25. Lan G-C, Hong T-P, Lee H-Y (2014) An efficient approach for finding weighted sequential patterns from sequence databases. Appl Intell 41(2):439–452
26. Lee VCS, Stankovic JA, Son SH (2000) Intrusion detection in real-time database systems via time signatures. In: Proceedings Sixth IEEE Real-Time Technology and Applications Symposium. RTAS 2000, pp 124–133

27. Lunt T, Tamaru A, Gilham F, Jagannathan R, Jalali C, Neumann P, Javitz H, Garvey T (1992) A real-time intrusion-detection expert system
28. Mangalampalli A, Pudi V (2009) Fuzzy association rule mining algorithm for fast and efficient performance on very large datasets, pp 1163–1168
29. Pannell G, Ashman H (2010) Anomaly detection over user profiles for intrusion detection. In: 8th Australian Information Security Management Conference
30. Pei J, Han J, Mortazavi-Asl B, Wang J, Pinto H, Chen Q, Dayal U, Hsu MC (2004) Mining sequential patterns by pattern-growth: the prefixspan approach. IEEE Trans Knowl Data Eng 16(11):1424–1440
31. Rahman MM, Ahmed CF, Leung CK, Pazdor AGM (2018) Frequent sequence mining with weight constraints in uncertain databases. In: Proceedings of the 12th International Conference on Ubiquitous Information Management and Communication, IMCOM '18. Association for Computing Machinery, New York
32. Rahman MM, Ahmed CF, Leung CK-S (2019) Mining weighted frequent sequences in uncertain databases. Inf Sci 479:76–100
33. Roesch M (1999) Snort: lightweight intrusion detection for networks. In: Proceedings of the 13th USENIX Conference on System Administration, LISA '99. USENIX Association, New York, pp 229–238
34. Ronao CA, Cho S-B (2016) Anomalous query access detection in RBAC-administered databases with random forest and PCA. Inf Sci 369:238–250
35. Ruijuan Z, Jing C, Mingchuan Z, Junlong Z, Qingtao W (2016) User abnormal behavior analysis based on neural network clustering. J China Univ Posts Telecommun 23(3):29–44
36. Singh I, Jindal R (2021) Expectation maximization clustering and sequential pattern mining based approach for detecting intrusive transactions in databases. Multim Tools Appl 80(18):27649–27681
37. Singh I, Manuja M, Mathur R, Goswami M (2020) Detecting intrusive transactions in databases using partially-ordered sequential rule mining and fractional-distance based anomaly detection. Int J Intell Eng Inform 8(2):138–171
38. Singh I, Mehraj B, Gupta N (2022) Intrusion detection system for databases: a hybrid metaheuristic clustering and closed sequential pattern mining approach. In: 2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS), vol 1, pp 38–45
39. Srikant R, Agrawal R (1996) Mining sequential patterns: generalizations and performance improvements. In: Advances in Database Technology—EDBT '96. Springer, Berlin, pp 1–17
40. Srivastava A, Sural S, Majumdar AK (2006) Weighted intra-transactional rule mining for database intrusion detection. In: Ng W-K, Kitsuregawa M, Li J, Chang K (eds) Advances in knowledge discovery and data mining. Springer, Berlin, pp 611–620
41. Srivastava A, Sural S, Majumdar A (2006) Database intrusion detection using weighted sequence mining. JCP 1:8–17
42. Subudhi S, Panigrahi S (2019) Application of optics and ensemble learning for database intrusion detection. J King Saud Univ Comput Inf Sci 34:972–981
43. Sun Y, Haoran X, Bertino E, Sun C (2016) A data-driven evaluation for insider threats. Data Sci Eng 1(2):73–85
44. Transaction Processing Performance Council (1996) TPC benchmark C standard specification revision 5.2. http://www.tpc.org/tpcc/spec/tpcc_current.pdf
45. Yang Y (2010) Web user behavioral profiling for user identification. Decis Support Syst 49:261–271
46. Yun U, Leggett JJ (2005) WFIM: weighted frequent itemset mining with a weight range and a minimum weight. In: Proceedings of the 2005 SIAM International Conference on Data Mining. SIAM, pp 636–640
47. Zamanian Z, Feizollah A, Anuar NB, Kiah LBM, Srikanth K, Kumar S (2019) User profiling in anomaly detection of authorization logs. In: Alfred R, Lim Y, Ibrahim AAA, Anthony P (eds) Computational science and technology. Springer, Singapore, pp 59–65

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.

Communications in Theoretical Physics

PAPER

Viscous cosmology in holographic dark energy with Granda–Oliveros cut-off

To cite this article: Simran Kaur and C P Singh 2023 *Commun. Theor. Phys.* **75** 025401

View the article online for updates and enhancements.

You may also like

- Klein–Gordon oscillator with magnetic and quantum flux fields in non-trivial topological space-time
Faizuddin Ahmed

This content was downloaded from IP address 14.139.251.98 on 13/02/2023 at 06:41

© 2023 Institute of Theoretical Physics CAS, Chinese Physical Society and IOP Publishing Printed in China and the UK

Commun. Theor. Phys. **75** (2023) 025401 (13pp) https://doi.org/10.1088/1572-9494/aca651

Viscous cosmology in holographic dark energy with Granda–Oliveros cut-off

Simran Kaur* **and C P Singh**

Department of Applied Mathematics, Delhi Technological University, Bawana Road, Delhi-110 042, India

E-mail: skjuneja.maths@gmail.com and cpsphd@rediffmail.com

Received 27 September 2022, revised 6 November 2022
Accepted for publication 28 November 2022
Published 6 February 2023

Abstract

We study the effect of bulk viscosity on holographic dark energy (HDE) with Granda–Oliveros infra-red cut-off. We focus on a generalized form of the bulk viscous coefficient which successfully describes the present-day evolution of the Universe. We observe that the model with bulk viscosity provides an elegant description of the early and late-time evolution of the Universe. We constrain the model through the combined observational data of the Strong Lensing System (SLS), measurements of Hubble parameter, Type Ia supernova data of Pantheon sample and local H_0 measured by SH0ES. Using best-fit values obtained from two different combinations of data, we estimate the present value of Hubble's constant $H_0 = 71.567^{+1.448}_{-0.848}$ and $H_0 = 69.197^{+1.563}_{-1.924}$, the deceleration parameter $q_0 = -0.535^{+0.016}_{-0.016}$ and $q_0 = -0.536^{+0.016}_{-0.016}$, and the equation of state $\omega_0 = -0.690^{+0.010}_{-0.010}$ and $\omega_0 = -0.691^{+0.011}_{-0.011}$, respectively. The HDE model with bulk viscosity exhibits the phase transition from decelerated epoch to accelerated epoch. A comparison of the model with the standard ΛCDM model is discussed with the statefinder and cosmographic parameters.

Keywords: viscosity, holographic dark energy, cosmology

(Some figures may appear in colour only in the online journal)

1. Introduction

Modern cosmology has plunged into a data-driven era. According to various accumulated observational evidence [1–9], the Universe is under an accelerated expansion which gives rise to the concept of dark energy (DE). It has been observed that the Universe contains 70% dark energy components and the remaining 30% includes baryons and cold dark matter. In the context of general relativity, a very simple candidate for this DE is the cosmological constant Λ, known as the Λ-cold-dark-matter (ΛCDM) model which is verified by recent highly precise Planck cosmic microwave background measurements [10]. However, it mainly suffers from two problems, the cosmological constant problem and the cosmic coincidence problem [11]. As we know very little about the realistic nature of DE, many cosmologists have proposed either modified theories or dynamical DE models to resolve these problems.

The holographic dark energy (HDE) is a dynamical DE model and the holographic principle (HP) is the foundation of this model [12, 13]. Being a proposal of quantum gravity, the HP states that the entropy of any system is not related to its volume but to its surface area. According to the principle, the maximum entropy limit set by the system should not be greater than the entropy of a black hole of the same size. This gives a connection between the short distance (ultra-violet) cut-off and long distance (infra-red) cut-off. Mathematically, in accordance with the HP, for a system with size L and an ultra-violet (UV) cut-off Λ, the vacuum energy density ρ_h should relate to the boundary surface of a system in a way $\rho_h \leqslant M_P{}^2 L^{-2}$. The energy density of HDE at the saturation of the system is given as $\rho_h = 3b^2 M_P{}^2 L^{-2}$ where b is a numerical constant and M_P denotes the reduced Planck constant. The HDE energy density depends on the choice of infra-red (IR) cut-off which represents the large length scale of the Universe. Hsu [14] considered the Hubble horizon as an IR cut-off and found that the HDE model could not drive the observable late-time expansion of the Universe. In a paper, Li [15] found that assuming the event horizon as an IR cut-off

* Author to whom any correspondence should be addressed.

could drive the accelerated expansion of the Universe but it faced the causality problem. Some other HDE models with different IR cut-offs have been proposed in the literature, e.g., agegraphic dark energy [16, 17], Ricci scale dark energy [18]. Granda and Oliveros [19] came up with a new IR cut-off which is a combination of the Hubble parameter and its time derivative to solve the causality problem. This cut-off has been studied by many authors [20–33] to explain the present-day evolution of the Universe.

The imperfect fluids including shear and bulk viscosity play a key role in the dynamics of the Universe. The viscous term arises when the fluid expands (or contracts) rapidly and ceases to be in thermodynamic equilibrium. It is a measure of pressure required to restore the thermodynamic equilibrium. Eckart [34] proposed the general theory for the relativistic imperfect fluid by considering the first-order deviation from equilibrium. Later on, this theory was modified by Landau and Lifshitz [35]. This theory suffers some problems in its formulation, e.g., equilibrium states are unstable and non-causal [36]. In fact, the neglected second-order terms are responsible for this problem. In order to remove such a problem, a second-order theory in a relativistic framework was developed by Israel and Stewart [37], which was stable and causal. However, the Eckart theory can be recovered from Israel-Stewart (IS) theory when the relaxation time becomes zero. Therefore, in the limit of vanishing relaxation time, the Eckart theory is a good approximation to IS theory. Due to the simple form of Eckart's theory and considering that the relaxation time goes to zero in the late-time evolution of the Universe, it has been studied by many authors to observe the effect of bulk viscous fluid in the late-time evolution of the Universe.

In the case of a homogenous and isotropic Universe, assuming the cosmological principle, the dissipative process within a thermodynamical approach can be modeled as a bulk viscosity. The bulk viscous pressure is characterized by the bulk viscous coefficient ξ. There is various parametrization of ξ available in the literature. The simplest parametrization of the bulk viscous coefficient is considered to be constant, i.e., $\xi = \xi_0$. The other parametrization has been considered as $\xi = \xi_1 H$, where ξ_1 is a constant and H is the Hubble parameter, and a linear combination form $\xi = \xi_0 + \xi_1 H$. In recent years, many authors [38–61] have discussed viscous cosmology to show the late-time evolution of the Universe by assuming these parameterized forms of the bulk viscous coefficient. Some authors [62–66] have studied the viscous cosmological model with bulk viscous coefficient depending on both expansion and acceleration rates, i.e., $\xi = \xi_0 + \xi_1 H + \xi_2 \ddot{a}/\dot{a}$, where ξ_2 is a constant, a denotes the scale factor, and the dot represents the differentiation with respect to cosmic time.

It has already been discovered that a non-viscous HDE model with a Granda–Oliveros IR cut-off does not show the observable evolution of the Universe [60]. In this paper, we examine the viscous effect in a dynamical HDE model with the new IR cut-off suggested by Granda and Oliveros [19] to achieve the present day evolution of the Universe. In this dynamical HDE model, we consider that our Universe is filled with dark matter consisting of bulk viscosity and HDE. We study the evolution of the Universe by considering a more generalized form of bulk viscous coefficient as proposed in the papers [62, 63]. We constrain the model to get the best-fit values of model parameters by using different observational data such as Type Ia supernova (Pantheon), Hubble data, strong lensing data, and the local Hubble value of SH0ES. We distinguish the viscous model from the standard ΛCDM model by studying the statefinder parameters and cosmographic parameters.

This paper is divided into the following sections: section 2 presents the basic equations of the HDE model with bulk viscosity and the solutions for the Hubble parameter, along with some other main cosmological parameters. In section 3, we estimate the best-fit values of parameters using the latest observational data. We use the best-fit values to discuss the evolution of the different cosmological parameters in section 4. A discussion on statefinder parameters and cosmographic parameters is presented in section 5. The conclusion is in the last section 6.

2. Bulk viscous HDE model

For presenting the bulk viscous HDE, we start from an isotropic and homogeneous flat Friedmann–Lemaitre–Robertson–Walker (FLRW) metric which is defined by the line element

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + a^2(t)[\mathrm{d}r^2 + r^2\mathrm{d}\Omega^2], \tag{1}$$

where $\mathrm{d}\Omega^2 = \mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2$ and $a(t)$ is the scale factor of the Universe. Throughout we use units $8\pi G = c = 1$.

It has been observed that the HDE model with Granda–Oliveros as the IR cut-off does not show the phase transition [60]. However, the observations indicate a transition phase from deceleration to acceleration during the evolution of the Universe. The recent works on bulk viscosity [62–66] show that this dissipative fluid plays an important role in describing the late-time evolution. Therefore, it will be worthy to discuss the HDE model with bulk viscosity by using the Granda–Oliveros IR cut-off to explain phase transition.

In a homogeneous and isotropic Universe, the only dissipative process allowed is the bulk viscosity. It is well known that bulk viscosity plays an important role in the Universe dynamics at the background level because it satisfies the cosmological principle. The viscous process fundamentally changes the equation of motion of relativistic fluids through the addition of new terms of hydrodynamics in energy-momentum tensor.

Hence, in dissipative cosmology with bulk viscosity, Einstein's field equations modify to $R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = T^{\mathrm{eff}}_{\mu\nu}$, where $T^{\mathrm{eff}}_{\mu\nu}$ is the effective energy-momentum tensor of the cosmic fluids including the bulk viscosity. This tensor is considered as $T^{\mathrm{eff}}_{\mu\nu} = (\rho + P)u_\mu u_\nu + P g_{\mu\nu}$, where $\rho = \rho_m + \rho_h$ is the total energy density of matter and HDE, and P is the effective fluid pressure which may be defined as $P = p_m + \Pi + p_h$. Here p_m, Π and p_h represent the dark matter

pressure, bulk viscous pressure and HDE pressure, respectively.

The first approach to studying the relativistic bulk viscosity process in thermodynamics systems is based on the Eckart theory [32]. It is known that this theory is unstable and non-causal against perturbations around the thermodynamical equilibrium state [36], i.e., it describes that all the equilibrium states are unstable and the signals can propagate through the fluids faster than the speed of light, i.e., with superluminal velocities. These problems could be traced back to their restriction to first-order deviations from equilibrium. In order to solve these problems, Israel and Stewart [37] proposed a second-order full causal theory in a relativistic framework. Despite instability and non-causality, Eckart's theory is still suitable for cosmological investigations while dealing with the accelerating Universe with bulk viscous fluids. This is due to the fact that this theory is a good approximation to the IS theory in the limit when the relaxation time vanishes in the late time. Another reason is that Eckart's theory is less complicated than the IS theory, so by taking advantage of the equivalence of both theories at this limit, it has been widely used by many authors to characterize the bulk viscous fluid in describing the late-time acceleration when the relaxation time goes to zero.

In Eckart formalism [34], the viscous pressure Π is considered as $\Pi = -\xi u^{a}_{;a} = -3\xi H$, where ξ is the bulk viscous coefficient, u_a is the four-velocity vector and H is the Hubble parameter. The bulk viscous coefficient ξ is conventionally chosen to be a positive quantity on thermodynamical grounds.

Let us consider the FLRW Universe (1) dominated by pressureless dark matter with the viscous term and the energy of the HDE. The non-vanishing equations of Einstein's field equations are

$$3H^2 = \rho_m + \rho_h \tag{2}$$

and

$$(\dot{\rho_m} + \dot{\rho_h}) + 3(\rho_m + \rho_h + p_h - 3\xi H)H = 0, \tag{3}$$

where an overdot denotes the derivative with respect to cosmic time t and $H = \dot{a}/a$ is the Hubble parameter.

Assuming that there is no interaction between the dark matter with bulk viscosity and HDE, the conservation equation (3), therefore, conserves separately, which is given by

$$\dot{\rho}_m + 3H(\rho_m - 3H\xi) = 0 \tag{4}$$

and

$$\dot{\rho}_h + 3H(1 + \omega_h)\rho_h = 0, \tag{5}$$

where $p_h = \omega_h \rho_h$ is the equation of state for HDE. Here, ω_h is the equation of state (EoS) parameter for HDE.

In the HDE model, as discussed in section 1, the UV cut-off is related to the vacuum energy, and the IR cut-off is related to the large scale of the Universe, for example, the Hubble horizon, future event horizon or particle horizon. Taking L as the size of the current Universe, for instance, using the Hubble scale, the resulting energy density is comparable to the present-day DE. Hsu [14] studied the HDE model with the Hubble horizon as the IR cut-off and found that the evolution of DE is the same as that of the dark matter (dust matter). Therefore, it cannot drive the Universe to accelerated expansion. The same appears if one chooses the particle horizon of the Universe as the length scale L. However, Li [15] studied the HDE model with the event horizon as the IR cut-off and found that the holographic DE not only gives the observation value of DE in the Universe but also can drive the Universe to an accelerated expansion phase. In that case, however, an obvious drawback concerning causality appears in this proposal. An event horizon is a global concept of spacetime and the existence of the event horizon of the Universe depends on the future evolution of the Universe. Therefore, the event horizon exists only for a Universe with forever accelerated expansion. Granda–Oliveros [19] proposed a new IR cut-off for HDE, which is a combination of the Hubble parameter and its time derivative. This model depends on local quantities and avoids the problem of causality which appears using the event horizon area as the IR cut-off.

The new IR cut-off for HDE as proposed by Granda and Oliveros [19] is given by

$$\rho_h = 3(\alpha H^2 + \beta \dot{H}), \tag{6}$$

where α and β are the dimensionless parameters to be computed by the current observational data.

Granda and Oliveros [19] argued that since the underlying origin of the holographic DE is still unknown, the inclusion of the time derivative of the Hubble parameter may be expected as this term appears in the curvature scalar, and has the correct dimension. This kind of density may appear as the simplest case of more general $f(H, \dot{H})$ holographic density in the FLRW background. Comparing (6) with the HDE density $\rho_h = 3b^2 M_P^2 L^{-2}$ shows that the corresponding IR cut-off for model (6) is

$$L = H^{-1}\left(1 + \frac{\beta}{\alpha}\frac{\dot{H}}{H^2}\right)^{-1/2} \tag{7}$$

which depends on local quantities and avoids the causality problem. Using the above-mentioned energy density for HDE, equations (2), (4) and (5) give the evolution equation for the Hubble function as

$$\dot{H} + \frac{3(1 + \alpha\omega_h)}{(2 + 3\beta\omega_h)}H^2 = \frac{3\xi H}{2 + 3\beta\omega_h}. \tag{8}$$

Using $x = \ln a$, the above equation can be transformed into

$$h' + \frac{3(1 + \alpha\omega_h)}{(2 + 3\beta\omega_h)}h = \frac{3\xi}{(2 + 3\beta\omega_h)H_0}, \tag{9}$$

where $h = H/H_0$ is the dimensionless Hubble parameter and $h' = dh/dx$ i.e. the prime indicates the derivative with respect to $\ln a$. The above evolution equation can analytically be solved provided to assume a specific form of bulk viscous coefficient ξ.

In an expanding Universe, the bulk viscous coefficient may depend on both the velocity and acceleration. The most logical form can be a linear combination of three terms: the

first term is a constant ζ_0, the second term is proportional to the Hubble parameter, which characterizes the dependence of the bulk viscosity on velocity, and the third is proportional to $\ddot{a}/\dot{a}$, characterizing the effect of acceleration on the bulk viscosity. Thus, we consider the parameterized bulk viscous coefficient as a combination of three terms which is given by [63]

$$\xi = \xi_0 + \xi_1\frac{\dot{a}}{a} + \xi_2\frac{\ddot{a}}{\dot{a}}. \tag{10}$$

In terms of the Hubble parameter, this can be written as

$$\xi = \xi_0 + \xi_1 H + \xi_2\left(\frac{\dot{H}}{H} + H\right), \tag{11}$$

where ξ_0, ξ_1 and ξ_2 are constants. Defining the dimensionless bulk viscous parameters $\zeta_0 = \xi_0/H_0$, $\zeta_1 = \xi_1 + \xi_2$, $\zeta_2 = \xi_2$ and $\zeta = \xi/H_0$, equation (11) reduces to

$$\zeta = \zeta_0 + \zeta_1 h + \zeta_2 h'. \tag{12}$$

Using (12) into (9), we finally get the evolution equation as

$$h' + \frac{3(1+\alpha\omega_h-\zeta_1)}{(2+3\beta\omega_h-3\zeta_2)}h = \frac{3\zeta_0}{2+3\beta\omega_h-3\zeta_2}. \tag{13}$$

Integrating the above equation to get the solution of the dimensionless Hubble parameter as

$$h = \frac{\zeta_0}{1+\alpha\omega_h-\zeta_1} + \left(1-\frac{\zeta_0}{1+\alpha\omega_h-\zeta_1}\right) \times a^{-\frac{3(1+\alpha\omega_h-\zeta_1)}{(2+3\beta\omega_h-3\zeta_2)}}, \tag{14}$$

which can further be simplified by considering a normalized relationship between the scale factor and redshift, $a = (1+z)^{-1}$ to get the solution for the Hubble parameter as

$$H = H_0\left[\frac{\zeta_0}{1+\alpha\omega_h-\zeta_1} + \left(1-\frac{\zeta_0}{1+\alpha\omega_h-\zeta_1}\right) \times (1+z)^{\frac{3(1+\alpha\omega_h-\zeta_1)}{(2+3\beta\omega_h-3\zeta_2)}}\right]. \tag{15}$$

It should be noted that for ζ_0, ζ_1 and ζ_2 equal to zero, the Hubble parameter gives $H = H_0(1+z)^{3(1+\alpha\omega_h)/(2+3\beta\omega_h)}$, which corresponds to the power-law solution of the HDE Universe. Further, $\alpha = 0$ and $\beta = 0$ reduce the model to the matter-dominated Universe $H = H_0(1+z)^{3/2}$, whose solution gives the power-law expansion of the Universe. However, equation (15) gives the solution of scale factor which has power-law form in early times and an exponential expansion in the late-time evolution of the Universe. Therefore, we obtain a solution that shows a phase transition from deceleration epoch to acceleration epoch during the evolution. The transition from decelerated to accelerated phase can be further explained by defining the deceleration parameter q as $q = -\frac{a\ddot{a}}{\dot{a}^2} = -1 - \frac{a}{h}\frac{\mathrm{d}h}{\mathrm{d}a}$. Thus, from equation (14), we get

$$q = -1 + \frac{3(1+\alpha\omega_h-\zeta_1)}{2+3\beta\omega_h-3\zeta_2} \times \frac{\left(1-\frac{\zeta_0}{1+\alpha\omega_h-\zeta_1}\right)(1+z)^{\frac{3(1+\alpha\omega_h-\zeta_1)}{2+3\beta\omega_h-3\zeta_2}}}{\left[\frac{\zeta_0}{1+\alpha\omega_h-\zeta_1} + \left(1-\frac{\zeta_0}{1+\alpha\omega_h-\zeta_1}\right)(1+z)^{\frac{3(1+\alpha\omega_h-\zeta_1)}{(2+3\beta\omega_h-3\zeta_2)}}\right]} \tag{16}$$

which depends on cosmic time and hence shows the phase transition. We observe that $q(z)$ approaches to -1 in the late time (negative redshift). The transition redshift z_{tr} can be obtained by substituting $q = 0$ in the above equation, which is obtained as

$$z_{\mathrm{tr}} = -1 + \left[1 + \frac{(1+\alpha\omega_h-\zeta_1)(1+3(\alpha-\beta)\omega_h-3(\zeta_0+\zeta_1-\zeta_2))}{\zeta_0(2+3\beta\omega_h-3\zeta_2)}\right]^{-\frac{2+3\beta\omega_h-3\zeta_2}{3(1+\alpha\omega_h-\zeta_1)}}. \tag{17}$$

From (16), the present value of $q(z)$ corresponds to $z = 0$ is

$$q_0 = -1 + \frac{3(1+\alpha\omega_h-\zeta_1-\zeta_0)}{2+3\beta\omega_h-3\zeta_2}. \tag{18}$$

Let us derive one more parameter to discuss the evolution of the Universe. This parameter is known as the effective equation of state (EoS) parameter, which is defined as $w_{\mathrm{eff}} = -1 - \frac{2a}{3h}\frac{\mathrm{d}h}{\mathrm{d}a}$. For this viscous HDE model we get

$$w_{\mathrm{eff}} = -1 + \frac{2(1+\alpha\omega_h-\zeta_1-\zeta_0)}{2+3\beta\omega_h-3\zeta_2}\frac{(1+z)^{\frac{3(1+\alpha\omega_h-\zeta_1)}{2+3\beta\omega_h-3\zeta_2}}}{h}. \tag{19}$$

The present value of w_{eff} at $z = 0$ for our model is

$$\omega_{\mathrm{eff}}(z=0) = -1 + \frac{2(1+\alpha\omega_h-\zeta_1-\zeta_0)}{2+3\beta\omega_h-3\zeta_2}. \tag{20}$$

3. Data samples and methodology

In this section, following the derivation of $H(z)$ obtained in (15), we constrain the space parameters (H_0, ζ_0, ζ_1, ζ_2, α, β, ω_h) of viscous HDE model using the two different combinations of the following mentioned recent data sets.

3.1. Strong lensing system

In this paper, we use new data from the Strong Lensing System (SLS) to constrain the model parameters. Due to the presence of gravity, light rays that pass near matter are bent in accordance with the general theory of relativity, and this bending of light is strong enough to produce multiple images of the source. This phenomenon is known as strong gravitational lensing [67]. For elliptical galaxies acting as lenses,

image separation depends on the mass of the lens and also on the angular diameter distance between the lens and the source, and also between the observer and the lens.

The chi-square function for 204 SLS in the redshift $0.063 < z_l < 0.950$ for the lens and $0.196 < z_s < 3.595$ for the source is given by [68]

$$\chi^2_{\rm SLS} = \sum_{i=1}^{204} \frac{[D^{\rm th}(z_1, z_s, \Theta) - D^{\rm obs}(\theta_{\rm e}, \sigma^2)]^2}{(\delta D^{\rm obs})^2}, \tag{21}$$

where Θ is the free space parameter and $D^{\rm obs} = c^2\theta_{\rm e}/4\pi\sigma^2$. The uncertainty of each $D^{\rm obs}$ measurement is calculated by

$$\delta D^{\rm obs} = D^{\rm obs}\sqrt{\left(\frac{\delta\theta_{\rm e}}{\theta_{\rm e}}\right)^2 + 4\left(\frac{\delta\sigma}{\sigma}\right)^2}, \tag{22}$$

where $\delta\theta_{\rm e}$ and $\delta\sigma$ are the error reported for the Einstein radius and velocity dispersion respectively.

In the above equations, $\theta_{\rm e}$ is the Einstein radius of the lens which is defined by a Singular Isothermal Sphere (SIS) as

$$\theta_{\rm e} = 4\pi\frac{\sigma^2_{\rm SIS} D_{ls}}{c^2 D_{\rm s}}, \tag{23}$$

where D_{ls} is the angular diameter distances between the lens galaxy and the source and D_s is the angular diameter distances between the observer and the source, σ_{SIS} is the velocity dispersion of the lens galaxy and c is the speed of light.

Now, D_s, in terms of redshift can be defined as

$$D_s(z) = \frac{c}{(1+z)}\int_0^{z_s}\frac{dz'}{H(z')} \tag{24}$$

and D_{ls} is given by

$$D_{ls}(z) = \frac{c}{(1+z)}\int_{z_l}^{z_s}\frac{dz'}{H(z')}. \tag{25}$$

Therefore, the theoretical distance $D^{\rm th}$ in equation (21) is defined by the ratio $D^{\rm th} = D_{ls}/D_{\rm s}$.

3.2. Hubble data

We use the updated collection of 57 Hubble $H(z)$ data points between the redshift range $0.07 \leqslant z \leqslant 2.36$ that consists of 31 points collected from the differential age (DA) technique and 26 data points received through line-of-sight *BAO* and other methods. The complete list of data is presented in table 1 with respective references, compiled by Sharov and Vasiliev [69]. The minimized chi-squared function for determining the best-fit values of the model is

$$\chi^2_{\rm H(z)}(\Theta) = \sum_{i=1}^{57}\frac{[H(z_i, \Theta) - H_{\rm obs}(z_i)]^2}{\sigma^2_{\rm H(z_i)}} \tag{26}$$

where $H(z_i, \Theta)$ represents the theoretical values of the Hubble parameter with model parameters, $H_{\rm obs}(z_i)$ is the observed values of the Hubble parameter and σ_i represents the standard deviation measurement uncertainty in $H_{\rm obs}(z_i, \Theta)$.

3.3. Type Ia supernovae (Pantheon data)

The confidence in Type Ia supernovae (SNe) as standard candles has been growing rapidly over the last two decades.

Table 1. H(z) data in $Km\ s^{-1}$ Mpc^{-1} consisting of 57 points (source: Sharov and Vasiliev, 2018).

DA Method (31 points)

z	$H(z)$	$\sigma_{\rm h(z)}$	*Ref.*	z	$H(z)$	$\sigma_{\rm h(z)}$	*Ref.*
0.070	69	19.6	[70]	0.4783	80	99	[74]
0.90	69	12	[71]	0.480	97	62	[70]
0.120	68.6	26.2	[70]	0.593	104	13	[72]
0.170	83	8	[71]	0.6797	92	8	[72]
0.1791	75	4	[72]	0.7812	105	12	[72]
0.1993	75	5	[72]	0.8754	125	17	[72]
0.200	72.9	29.6	[73]	0.880	90	40	[70]
0.270	77	14	[71]	0.900	117	23	[71]
0.280	88.8	36.6	[73]	1.037	154	20	[72]
0.3519	83	14	[72]	1.300	168	17	[71]
0.3802	83	13.5	[74]	1.363	160	33.6	[76]
0.400	95	17	[71]	1.430	177	18	[71]
0.4004	77	10.2	[74]	1.530	140	14	[71]
0.4247	87.1	11.2	[74]	1.750	202	40	[71]
0.4497	92.8	12.9	[74]	1.965	186.5	50.4	[76]
0.470	89	34	[75]				

BAO and other method (26 points)

z	$H(z)$	$\sigma_{\rm h(z)}$	*Ref.*	z	$H(z)$	$\sigma_{\rm h(z)}$	*Ref.*
0.24	79.69	2.99	[77]	0.52	94.35	2.64	[79]
0.30	81.7	6.22	[78]	0.56	93.34	2.3	[79]
0.31	78.18	4.74	[79]	0.57	87.6	7.8	[83]
0.34	83.8	3.66	[77]	0.57	96.8	3.4	[84]
0.35	82.7	9.1	[80]	0.59	98.48	3.18	[79]
0.36	79.94	3.38	[79]	0.60	87.9	6.1	[82]
0.38	81.5	1.9	[81]	0.61	97.3	2.1	[81]
0.40	82.04	2.03	[79]	0.64	98.82	2.98	[79]
0.43	86.45	3.97	[77]	0.73	97.3	7.0	[82]
0.44	82.6	7.8	[82]	2.30	224	8.6	[85]
0.44	84.81	1.83	[79]	2.33	224	8	[86]
0.48	87.79	2.03	[79]	2.34	222	8.5	[87]
0.51	90.4	1.9	[81]	2.36	226	9.3	[88]

The first strong indication of the accelerated expansion of the Universe was also given by the SNe observations. The latest compilation of Pantheon sample SNe has been used, which consists of 1048 data points from SNLS, SDSS, Pan-STARRS1, and a HST survey in the redshift range of $0.014 \leqslant z \leqslant 2.3$ [89].

We can minimize the χ^2 function for SNe as

$$\chi^2_{\rm SNe} = \mu\ C^{-1}\ \mu^{\rm T}, \tag{27}$$

where $\mu = \mu_i^{\rm obs} - \mu^{\rm th}$ and C is the covariance matrix of $\mu^{\rm obs}$ given in [90]. The observed distance modulus is given in Ref. [91]. Also, the theoretical distance modulus is defined as

$$\mu^{\rm th} = 5\log_{10}[d_L(z)/10pc] + M, \tag{28}$$

where M is the nuisance parameter and $d_{L(z)}$ is the dimensionless luminosity distance defined as [89]

$$d_{\rm L}(z) = (1+z)c\int_0^z\frac{{\rm d}z'}{H(z', \Theta)}, \tag{29}$$

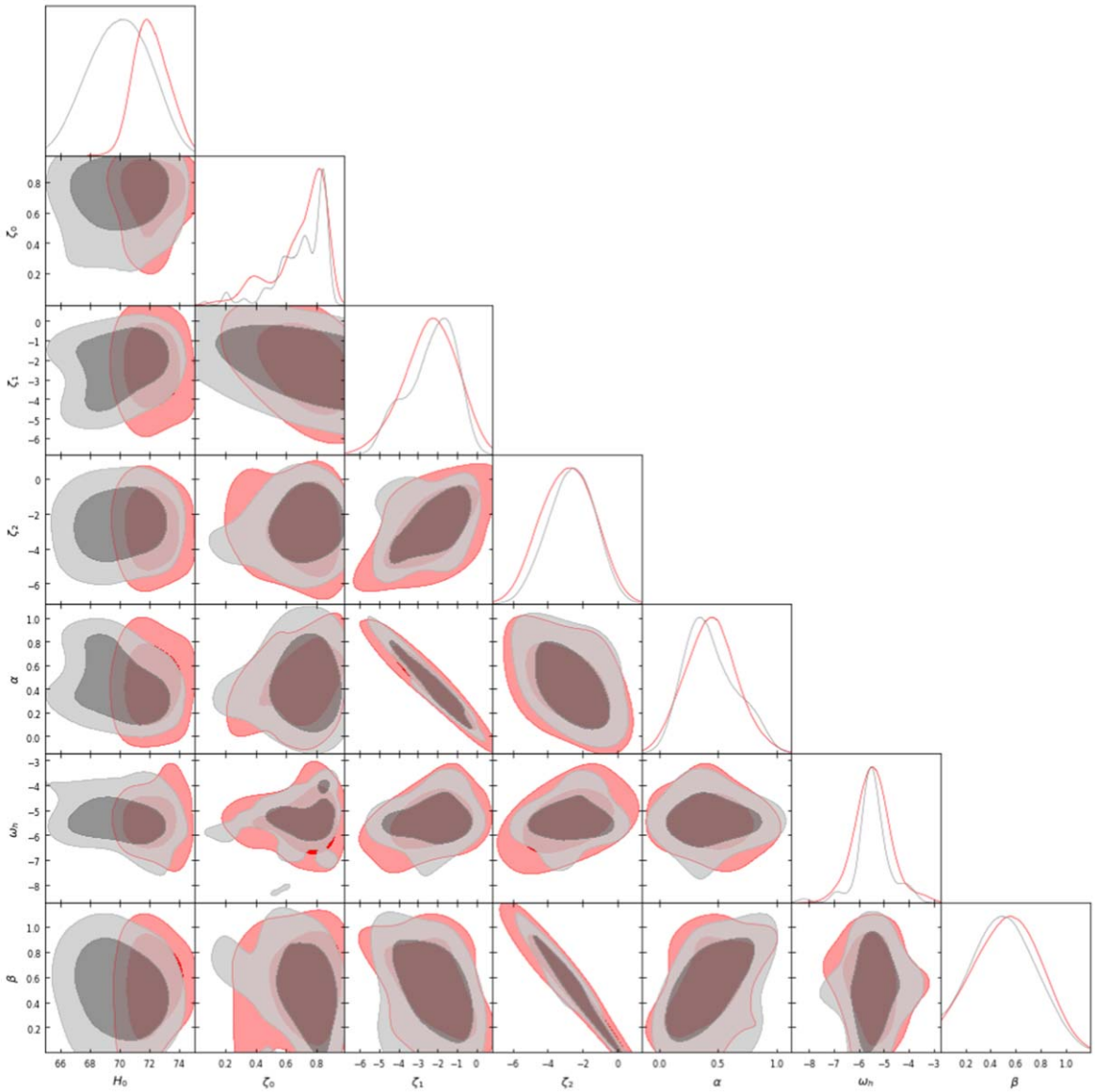


Figure 1. The likelihood contours at 68.3% CL and 95.4% CL for viscous HDE model correspond to DS1(red color) and DS2 (grey color) datasets.

Table 2. The best-fit values of free parameters of HDE and ΛCDM models with errors from DS1 and DS2 datasets.

	*DS*1		*DS*2	
Model	HDE	Λ*CDM*	HDE	Λ*CDM*
H_0	$71.567^{+1.448}_{-0.848}$	$70.061^{+0.775}_{-0.852}$	$69.197^{+1.563}_{-1.924}$	$68.944^{+0.738}_{-0.530}$
ζ_0	$0.758^{+0.097}_{-0.221}$	—	$0.652^{+0.146}_{-0.220}$	—
ζ_1	$-2.690^{+1.335}_{-1.650}$	—	$-2.458^{+1.327}_{-2.056}$	—
ζ_2	$-2.835^{+1.410}_{-1.608}$	—	$-2.940^{+1.447}_{-1.906}$	—
α	$0.460^{+0.261}_{-0.275}$	—	$0.462^{+0.294}_{-0.266}$	—
ω_h	$-5.556^{+0.917}_{-0.782}$	—	$-5.576^{+0.336}_{-0.525}$	—
β	$0.517^{+0.294}_{-0.238}$	—	$0.565^{+0.301}_{-0.263}$	—
$\Omega_{\rm m}$	$0.308^{+0.121}_{-0.093}$	$0.286^{+0.004}_{-0.007}$	$0.26^{+0.122}_{-0.103}$	$0.288^{+0.004}_{-0.006}$
Ω_Λ	$0.668^{+0.112}_{-0.103}$	$0.711^{+0.010}_{-0.006}$	$0.73^{+0.122}_{-0.123}$	$0.707^{+0.009}_{-0.005}$
$\chi^2_{\rm min}$	525.033	537.330	525.473	530.538

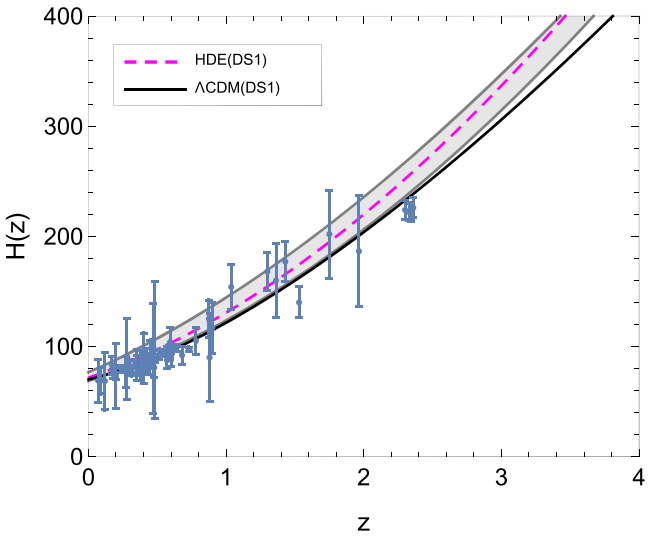


Figure 2. Plot of the Hubble function as a function of redshift for HDE model with bulk viscosity for DS1 dataset over H(z) points and its comparison with ΛCDM model (solid black line). The 57 $H(z)$ data points are also shown with error bars in blue dots. The band corresponds to the error at the 95.4% confidence level.

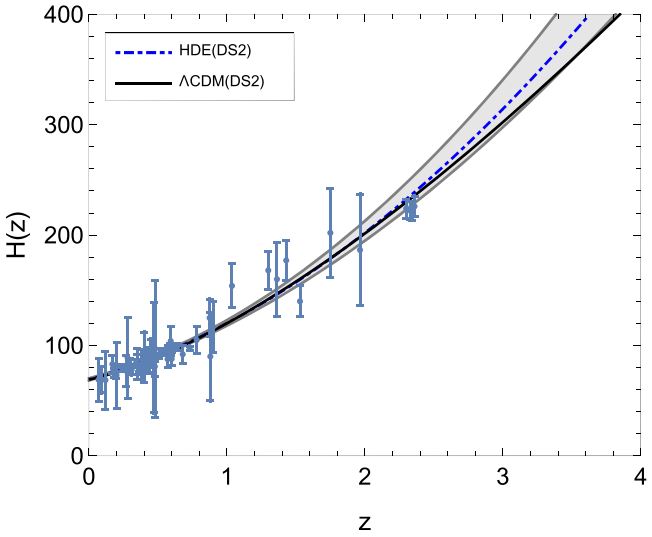


Figure 3. Plot of the Hubble function as a function of redshift for viscous HDE model with bulk viscosity for DS2 dataset over H(z) points and its comparison with ΛCDM model (solid black line). The 57 $H(z)$ data points are also shown with error bars in blue dots. The band corresponds to the error at the 95.4% confidence level.

where Θ is the space parameters of the model and c is the speed of light.

3.4. Local Hubble constant

In addition, we take the recently measured local Hubble constant H_0 as $H_0 = 73.5 \pm 1.4$ km s^{-1}Mpc^{-1} by SH0ES as mentioned in [92].

3.5. Methodology

We perform fitting to determine the best-fit values of model parameters using the Markov chain Monte Carlo (MCMC) method in the EMCEE library [93]. We consider two combinations of the latest datasets of *SNe*(*Pantheon*), $H(z)$, *SLS* and local H_0, namely, DS1: $\chi^2_{\rm DS1} = \chi^2_{\rm SNe} + \chi^2_{\rm H(z)} + \chi^2_{\rm SLS} + \chi^2_{\rm H_0}$ and DS2: $\chi^2_{\rm DS2} = \chi^2_{\rm SNe} + \chi^2_{\rm H(z)} + \chi^2_{\rm SLS}$. In our statistical analysis, we minimize the function χ^2 for these data sets. The best-fit values of space parameters ($\Theta = H_0$, ζ_0, ζ_1, ζ_2, α, ω_h, β) of viscous HDE and ($\Theta = H_0$, Ω_Λ, Ω_m) of ΛCDM models obtained by both the combinations are provided in table 2. The likelihood contour at 1σ(68.3%) and 2σ(95.4)% confidence levels for both data sets is represented in figure 1. The data set DS1 fitted is shown by the red contours while the greyish contours are derived from the DS2 data set. In the following section, we present and discuss the results obtained from the above-mentioned data sets.

4. Results and discussion

We listed the best-fit values of model parameters from two different joint combinations of the DS1 and DS2 datasets in table 2. In what follows, we study the observational parameters, namely the Hubble constant, deceleration parameter, equation of state parameter, and the age of the Universe to describe the global dynamics of the Universe.

Figures 2 and 3 show the evolution of the Hubble parameter $H(z)$ as a function of the redshift using the best-fit values of free parameters constrained from both the datasets DS1 and DS2. We have also traced the trajectory of the ΛCDM model to compare it to the evolution of the viscous HDE model. The bars stand for the observational data of $H(z)$ as mentioned in [69] as well as given in table 1. It is observed that the fit obtained from DS2 is consistent with observational $H(z)$ data points.

The present values of the Hubble constant obtained from DS1 and DS2 are $H_0 = 71.567^{+1.448}_{-0.848}$ and $H_0 = 69.197^{+1.563}_{-1.924}$, respectively. The first value of H_0 is slightly lower than the value obtained by SH0ES project $H_0 = 73.5 \pm 1.4$ km s^{-1}Mpc^{-1} [92] whereas the second result of H_0 is slightly higher than Planck result [94], where $H_0 = 67.7 \pm 0.46$ km s^{-1}Mpc^{-1}. The respective χ^2 with DS1 and DS2 data sets are 525.033 and 525.473.

The reduced χ^2 statistics are very beneficial in the goodness of fit testing. It is defined as $\chi^2_{\rm red} = \chi^2_{\rm min}/\nu$, where ν is known as the degree of freedom (dof) and is defined as the difference between the total number of combined data points used and the number of estimated free model parameters. We have the number of data points for DS1 as $N = 1310$ (1048 data of SNe, 57 data of H(z), 204 data of SLS and 01 data of local H_0) and for DS2 it is $N = 1309$ (1048 data of SNe, 57 data of H(z) and 204 data of SLS). The viscous HDE has seven free parameters whereas the ΛCDM has 3. Thus, the $\chi^2_{\rm red}$ for ΛCDM comes out to be $\chi^2_{\rm red} = 0.411$ and $\chi^2_{\rm red} = 0.406$ whereas for the viscous HDE model, these are $\chi^2_{\rm red} = 0.402$ and $\chi^2_{\rm red} = 0.403$, respectively, which is less than unity with each dataset, showing that both the models fit well and data sets are compatible with the considered model.

Table 3. Values of z_{tr}, q_0, $\omega_{eff}(z=0)$ and t_0 (Gyr) for different combinations of data sets.

	*DS*1		*DS*2	
Model	HDE	ΛCDM	HDE	ΛCDM
z_{tr}	$0.666^{+0.426}_{-0.310}$	$0.707^{+0.223}_{-0.205}$	$0.777^{+0.234}_{-0.130}$	$0.699^{+0.143}_{-0.125}$
q_0	$-0.535^{+0.016}_{-0.016}$	$-0.568^{+0.013}_{-0.012}$	$-0.536^{+0.016}_{-0.016}$	$-0.563^{+0.012}_{-0.012}$
$\omega_{eff}(z=0)$	$-0.690^{+0.010}_{-0.010}$	$-0.615^{+0.006}_{-0.006}$	$-0.691^{+0.011}_{-0.011}$	$-0.611^{+0.068}_{-0.068}$

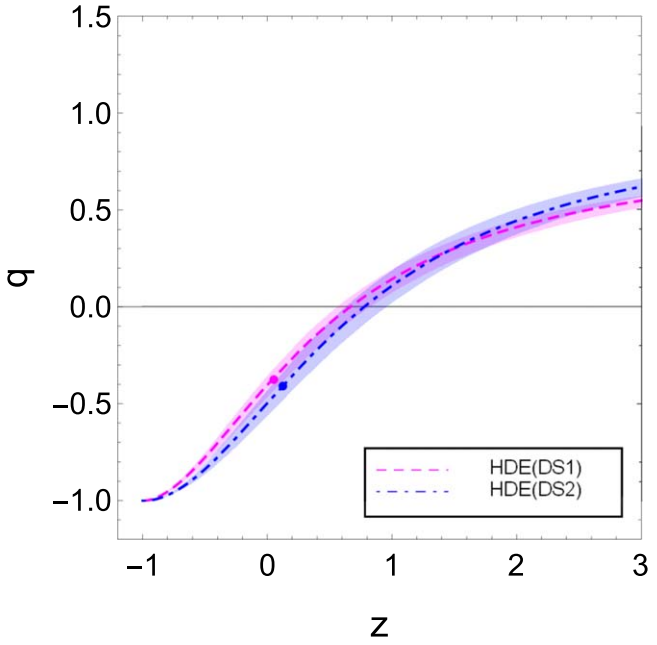


Figure 4. Plot of deceleration parameter q versus redshift z for best-fit values of free parameters obtained from DS1 and DS2 data sets. The current value q_0 is shown by a dot on the trajectory. The band corresponds to the error at the 95.4% confidence level.

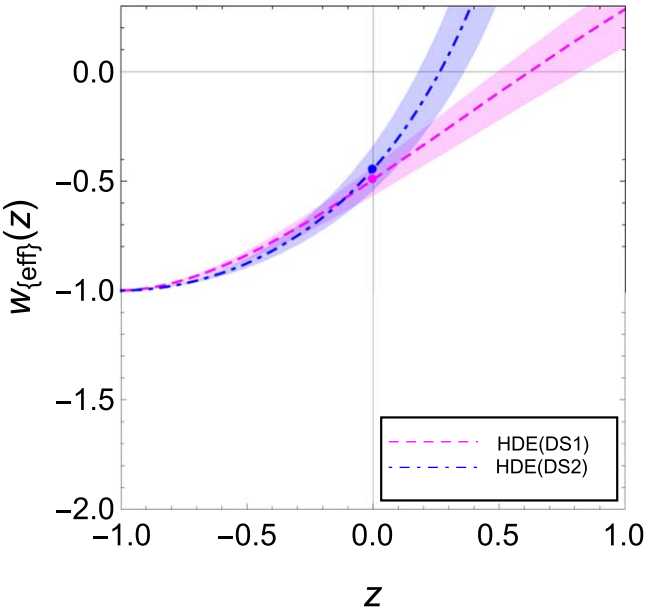


Figure 5. Plot of Effective equation of state parameter ω_{eff} versus redshift z for best-fit values of free parameters obtained from DS1 and DS2 data sets. The current $\omega_{eff}(z=0)$ is shown by a dot on the trajectory. The band corresponds to the error at the 95.4% confidence level.

Table 3 presents the values of transition redshift, the present value of q and ω_{eff} of viscous HDE and ΛCDM models. Figure 4 shows the evolution of the deceleration parameter defined in equation (16) for best-fit values of free parameters with the joint datasets. The deceleration parameter shows the transition from $q>0$ to $q<0$ with both data sets *DS*1 and *DS*2. In this cosmological scenario, the signature flipping occurs at the transition redshift $z_{tr}=0.666^{+0.426}_{-0.310}$ and $z_{tr}=0.777^{+0.234}_{-0.130}$ corresponding to the DS1 and DS2 datasets, which are very close to the ΛCDM model's transition values. These results are also consistent with the results reported in [95]. Thus, the Universe has a transition from an early decelerated era to the current observable accelerated era. The present value of $q(z)$ is found to be about $q_0=-0.535^{+0.016}_{-0.016}$ and $q_0=-0.536^{+0.016}_{-0.016}$, respectively, which are comparable with ΛCDM and $q_0=-0.55\pm0.01$ of the Planck spacecraft data [94]. It is to be noted that Capozziello *et al* [96] have obtained $q_0=-0.56\pm0.04$.

Next, we analyze the cosmic expansion using the effective equation of state parameter ω_{eff}. We have plotted the trajectory of ω_{eff} versus redshift z in figure 5, using the best-fit values of free parameters of both the combinations of datasets. The present values of effective EoS parameter are $\omega_{eff}(z=0)=-0.690^{+0.010}_{-0.010}$ and $\omega_{eff}(z=0)=-0.691^{+0.010}_{-0.010}$ for both; the combination of data sets and their comparison with ΛCDM is listed in table 3. These values are comparatively larger than $\omega_0=-0.93$ which was predicted by joint analysis of $WMAP+BAO+H(z)+SNe$. The model behaves like quintessence DE. Note that $\omega_{eff}\to-1$ in the future time of evolution which implies that the viscous holographic dark energy model approaches to the de Sitter model in late time. Finally, we estimate the age of the Universe as $t_0\approx13.26$ *Gyr* and $t_0\approx13.56$ *Gyr*, respectively, which are comparable to the value reported in [94] with $t_0=13.79\pm0.02$ *Gyr*.

5. Diagnostic parameters

In this section, we discuss diagnostic parameters, like statefinder parameters and cosmography parameters, like jerk, snap, lerk, and m parameters to discriminate the proposed viscous HDE model with the dark energy model, like the ΛCDM model.

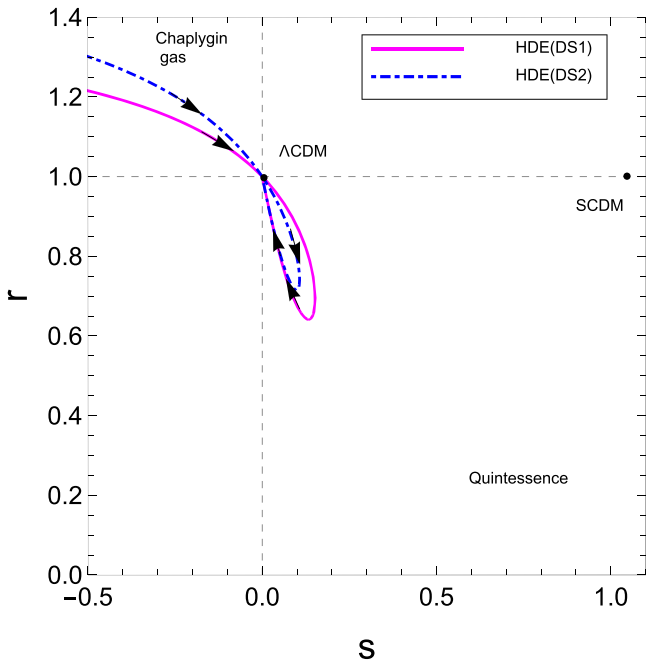


Figure 6. The evolution of $\{r, s\}$ in $s-r$ plane corresponding to best-fit values of free parameters obtained from DS1 and DS2. The direction of the evolution is shown by the arrows on each trajectory.

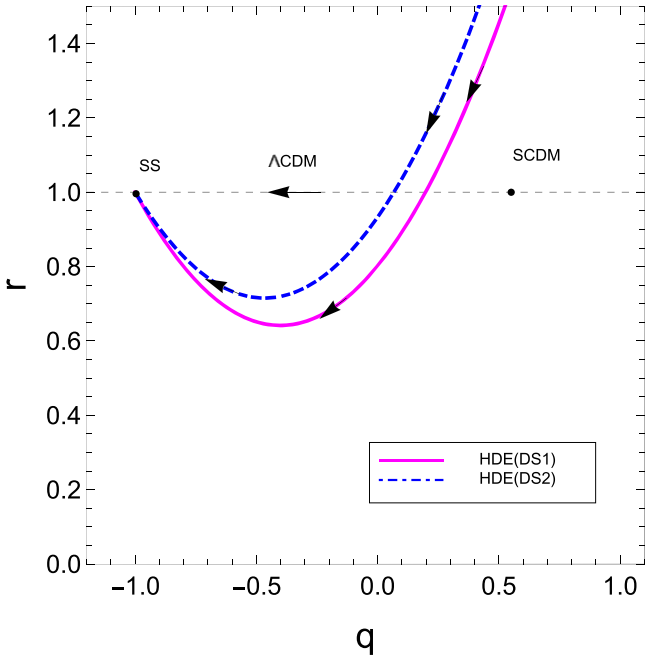


Figure 7. The evolution of $\{r, q\}$ in $q-r$ plane corresponding to best-fit values of free parameters obtained from DS1 and DS2.

Sahni *et al* [97] and Alam *et al* [98] introduced a new pair of geometrical diagnostic, known as statefinder parameters $\{r,\ s\}$ to distinguish the DE models. The pair, using the higher order derivatives of the scale factor, describe the expansion dynamics of the Universe and are defined as

$$r = \frac{1}{aH^3}\frac{\mathrm{d}^3 a}{\mathrm{d}t^3} = \frac{\ddot{H}}{H^3} - 3q - 2, \quad s = \frac{r-1}{3(q-1/2)}. \tag{30}$$

Now, substituting the values of the scale factor and its derivatives from (15) in the above equation, we get

$$\begin{aligned} r = {} & \frac{9(1+\alpha\omega_h-\xi_0-{\xi_1}^2)}{2+3\beta\omega_h-3{\xi_2}^2}\mathrm{e}^{-\frac{6\xi_0(t-t_0)}{2+3\beta\omega_h-3\xi_2}} \\ & - 9\frac{(1+\alpha\omega_h-\xi_0-\xi_1)(1-\alpha\omega_h+3\beta\omega_h+\xi_1-3\xi_2)}{2+3\beta\omega_h-3{\xi_2}^2}\mathrm{e}^{-\frac{3\xi_0(t-t_0)}{2+3\beta\omega_h-3\xi_2}} \\ & + (2+3\beta\omega_h-3\xi_2), \end{aligned} \tag{31}$$

and

$$\begin{aligned} s = {} & \frac{1}{(2+3\beta\omega_h-3\xi_2)\{2(1+\alpha\omega_h-\xi_0-\xi_1)-\mathrm{e}^{-\frac{3\xi_0(t-t_0)}{2+3\beta\omega_h-3\xi_2}}(2+3\beta\omega_h-2\xi_2)\}} \\ & \times [2\mathrm{e}^{-\frac{6\xi_0(t-t_0)}{2+3\beta\omega_h-3\xi_2}}(1+\alpha\omega_h-\xi_0-\xi_1) \\ & \times \{1+\alpha\omega_h-\xi_0-\xi_1-\mathrm{e}^{-\frac{6\xi_0(t-t_0)}{2+3\beta\omega_h-3\xi_2}}(1+\xi_1-3\xi_2-\alpha\omega_h+3\beta\omega_h)\}]. \end{aligned} \tag{32}$$

Figure 6 shows the $s-r$ plane trajectories of the viscous holographic dark energy model for best-fit values of parameters achieved through the DS1 and DS2 data. It can be observed that the $\{r,\ s\}$ evolutions start from a region of chaplygin gas where $r>1,\ s<0$ and tends to $\{r,\ s\}\rightarrow\{1,\ 0\}$ in future, a value of the ΛCDM model. In figure 6, the horizontal line $r=1$ belongs to the ΛCDM model. The present values of the statefinder parameters are $\{r_0 = 0.6415, s_0 = 0.132\}$ and $\{r_0 = 0.716, s_0 = 0.094\}$ for DS1 and DS2, respectively, which show a deviation from the ΛCDM model. An alternate way to check the differences between the models is through the $\{q,\ r\}$ trajectory. In figure 7 we plot such a $q-r$ trajectory for both data sets. The arrow represents the

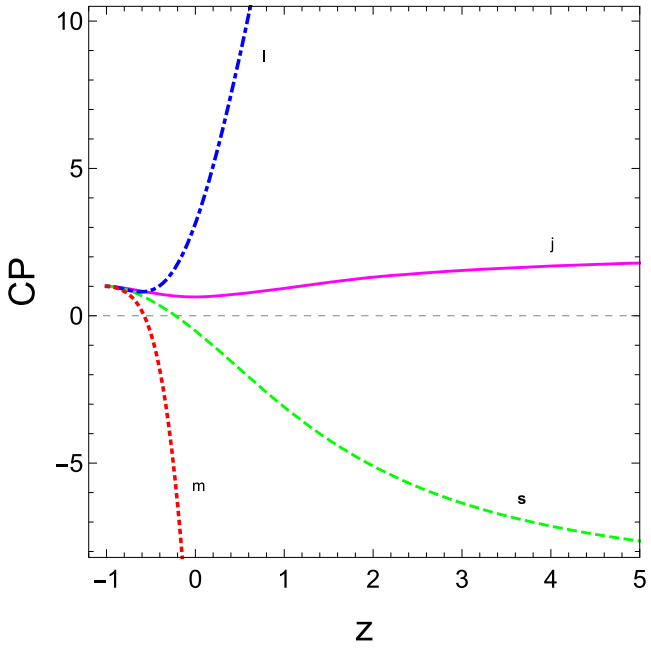


Figure 8. The evolutions of *jerk*, *lerk*, *snap* and *m* parameters corresponding to best-fit values of model parameters obtained from DS1 and DS2. The direction of the evolution of each trajectory is shown by the arrow.

direction of the evolution of the model. The trajectories evolute from deceleration phase to accelerated epoch. We find that $\{q,\ r\}$ converges to a steady state (SS) model $\{-1,\ 1\}$ in the future ($z \to -1$) with both data sets. In figure 7, the horizontal line at $r = 1$ belongs to the time evolution of the ΛCDM model.

In addition to the above, we report the high order cosmography parameters (*jerk*, *snap*, *lerk*, *m*) for best-fit values of parameters of the viscous HDE model. These parameters are defined as (see, [99] for historical nomenclature)

$$j = \frac{1}{aH^3}\frac{d^3a}{dt^3},\ \mathbf{s} = \frac{1}{aH^4}\frac{d^4a}{dt^4},$$
$$l = \frac{1}{aH^5}\frac{d^5a}{dt^5},\ m = \frac{1}{aH^6}\frac{d^6a}{dt^6} \tag{33}$$

The jerk parameter provides us the information about the dynamics of the DE. Although the *snap*, *lerk* and *m* parameters do not have a distinguished physical meaning, they are still an important part of the Taylor series of the Hubble parameter and give us more precision in the preferred model. They help to measure the rate of cosmic expansion more accurately.

The above cosmographic parameters in terms of redshift are defined as

$$j = -q + (1+z)\frac{dq}{dz} + 2q(1+q), \tag{34}$$

$$\mathbf{s} = j - 3j(1+q) - (1+z)\frac{dj}{dz}, \tag{35}$$

$$l = \mathbf{s} - 4\mathbf{s}(1+q) - (1+z)\frac{d\mathbf{s}}{dz}, \tag{36}$$

$$m = l - 5l(1+q) - (1+z)\frac{dl}{dz}. \tag{37}$$

Here, *r* and *j* are the same but the *s* parameter defined in (30) is different than the one defined in (35). Using the best-fit values of model parameters, the evolution of these CP for our model has been shown graphically against redshift, for DS1, and the same can be plotted for the data set DS2.

The current values of these parameters are found to be $\{j_0 = 0.6415,\ \mathbf{s}_0 = -0.5082,\ l_0 = 3.1301$ and $m_0 = -14.614\}$ and $\{j_0 = 0.716,\ \mathbf{s}_0 = -0.323,\ l_0 = 2.852$ and $m_0 = -12.490\}$ for DS1 and DS2 respectively. It can be observed from figure 8 that all the cosmological parameters $j, \mathbf{s}, l$ and m tend to unity as $z \to -1$, i.e., our model is in good agreement with the observations of the standard ΛCDM model in late-time evolution. We find that *jerk* and *lerk* parameters have trajectories in the same direction while **s** and *m* have the same evolution but are different than *j* and *l* and both transit from initial negative values to later positive ones. It is also noted that the positive *jerk* and *lerk* imply that the Universe has undergone a transition from deceleration to acceleration.

6. Conclusion

In this paper, we have studied the effect of bulk viscosity in the HDE model with Granda–Oliveros as an IR cut-off in the framework of FLRW space-time. The viscous term has a negative pressure, therefore, it has been studied to observe the early and late-time accelerated expansion of the Universe. It is well known that the HDE model with the Granda–Oliveros IR cut-off does not show the phase transition which contradicts the present observable Universe. Instead of assuming the other cut-off, we have included a viscous term in the HDE model with the same IR cut-off to observe the phase transition. We have assumed a more general form of the bulk viscous coefficient ξ in terms of H and $\dot{H}$ to demonstrate how the bulk viscosity explains the accelerating Universe. We have obtained the solution of different cosmological parameters such as the Hubble parameter, deceleration parameter and equation of state parameter.

We have constrained the space parameters of the viscous HDE model using the latest observational data of the Strong Lensing system containing 204 data points, $H(z)$ data comprising 57 points, SNe (pantheon data) of 1048 points and local H_0. We have performed two joint analyses, namely DS1 and DS2, comprising $SLS + SNe + H(z) + H_0$ and $SLS + SNe + H(z)$, respectively to find the best-fit values of parameters. We have also assumed the ΛCDM model as a concordance model to compare the results of the viscous HDE model. The best-fit values for both models are listed in table 2. Using the best-fit values, we have investigated the effects of bulk viscosity on the evolution of the Universe by plotting the trajectory of different cosmological parameters. We have found that the viscous HDE model fits well to both data sets. We have found a good agreement to data sets according to the $\chi^2_{\rm red}$ value. We have also obtained the

evolutions of cosmographic parameters and statefinder parameters. In the following, we summarize the main results.

We have plotted and analyzed the trajectory of each of the main observable cosmological parameters such as $H(z)$, $q(z)$ and $\omega_{\rm eff}(z)$ using the best-fit values of model parameters. In figures 2 and 3, the Hubble function obtained analytically has been confronted by error bars of Hubble data with best-fit values and compared with ΛCDM. We have found that the model predicts a better fit along with the ΛCDM model. From figures 4 and 5, the evolution of the deceleration and EoS parameters show that the model transits from deceleration phase to the accelerated phase. In late time, we have found that the parameters tend towards -1. The present value of deceleration and EoS parameters are very near to the values of ΛCDM model as mentioned in table 3. The ages of the Universe predicted by the viscous HDE model from two data sets are slightly lower than the value of the ΛCDM model. We have also studied the diagnostic parameters, namely statefinder and cosmographic parameters, for the viscous HDE model, and compared them with the ΛCDM model. Using the best-fit values, we have plotted the evolutions of $\{r, s\}$, $\{r, q\}$ in $(s-r)$, and $(q-r)$ planes, respectively, in figures 6 and 7 and discussed the behavior accordingly. In figure 8, we have plotted the evolution of cosmographic parameters, which explain the dynamics of the model. It has been observed that these parameters tend to 1 as z tends towards -1, that is, the viscous HDE model is in good agreement with the observation of the concordance model in late-time evolution.

The presence of bulk viscosity in dark energy models is an interesting and significant approach to show the phase transition of the Universe. The present model gives a good alternative to explain the accelerating phenomena of the Universe.

Acknowledgments

The authors are very thankful to the reviewers for their constructive suggestions to improve the manuscripts. SK, one of the authors, is acknowledged for providing SRF (Senior Research Fellowship) from the Council of Scientific and Industrial Research, New Delhi, India.

References

[1] Riess A G *et al* 1998 Observational evidence from supernovae for an accelerating universe and a cosmological constant *Astron. J.* **116** 1009
[2] Perlmutter S *et al* 1999 Measurements of ω and Λ from 42 high redshift supernovae *Astrophys. J.* **517** 565
[3] Tegmark M *et al* 2004 Cosmological parameters from SDSS and WMAP *Phys. Rev.* D **69** 103501
[4] Adelman-McCarthy J K *et al* 2006 The fourth data release of the sloan digital sky survey: preprint *Astrophys. J. Suppl.* **162** 38
[5] Riess A G *et al* 2007 New Hubble space telescope discoveries of type Ia supernovae at $z > 1$: narrowing constraints on the early behavior of dark energy *Astrophys. J.* **659** 98
[6] Spergel D N *et al* 2007 Wilkinson microwave anisotropy probe (WMAP) three year results: implications for cosmology *Astrophys. J. Suppl.* **170** 377
[7] Kessler R *et al* 2009 First-year sloan digital sky survey-ii supernova results: hubble diagram and cosmological parameters *Astrophys. J. Suppl.* **185** 32
[8] Abbott T M C *et al* 2019 First cosmology results using type Ia supernovae from the dark energy survey: constraints on cosmological parameters *Astrophys. J.* **872** L30
[9] Brout D 2019 First cosmology results using Type Ia supernovae from the dark energy survey: analysis, systematic uncertainties and validation *Astrophys. J.* **874** 150
[10] Ade P *et al* 2016 Planck 2015 results XIII. Cosmological parameters *Astron. Astrophys.* **594** A13
[11] Weinberg S 1989 The cosmological constant problem *Rev. Mod. Phys.* **61** 1
[12] Hooft G t 1993 Dimensional reduction in quantum gravity arXiv:gr-qc/9310026
[13] Susskind L 1995 The world as a hologram *J. Math. Phys.* **36** 6377
[14] Hsu S D 2004 Entropy bounds and dark energy *Phys. Lett.* B **594** 13
[15] Li M 2004 A model of holographic dark energy *Phys. Lett.* B **603** 1
[16] Cai R G and Dark A 2007 Energy model characterized by the age of the universe *Phys. Lett.* B **657** 228
[17] Wei H and Cai R G 2008 A new model of agegraphic dark energy *Phys. Lett.* B **660** 113
[18] Gao C, Wu F Q, Chen X and Shen Y G 2009 Holographic dark energy model from Ricci scalar curvature *Phys. Rev.* D **79** 043511
[19] Granda L N and Oliveros A 2008 Infrared cut-off proposal for the holographic density *Phys. Lett.* B **669** 275
[20] Granda L N and Oliveros A 2009 New infrared cut-off for the holographic scalar fields models of dark energy *Phys. Lett.* B **671** 199
[21] Karami K and Fehri J 2010 Holographic dark energy in a non-flat universe with granda-oliveros cut-off *Int. J. Theor. Phys.* **49** 1118
[22] Yu F *et al* 2010 A more general interacting model of holographic dark energy *Phys. Lett.* B **688** 263
[23] Wang Y and Xu L 2010 Current observational constraints to the holographic dark energy model with a new infrared cutoff via the Markov chain Monte Carlo method *Phys. Rev.* D **81** 083523
[24] Malekjani M, Khodam-Mohammadi A and Nazari-pooya N 2011 Cosmological evolution and statefinder diagnostic for new holographic dark energy model in non flat universe *Astrophys. Space Sci.* **332** 515
[25] Oliveros A and Acero M A 2015 New holographic dark energy model with non-linear interaction *Astrophys. Space Sci.* **357** 12
[26] Abdollahi Zadeh M, Sheykhi A and Moradpour H 2017 Holographic dark energy with the sign-changeable interaction term *Int. J. Mod. Phys.* D **26** 1750080
[27] Singh C P and Kumar P 2017 Holographic dark energy in brans-dicke theory with logarithmic form of scalar field *Int. J. Theor. Phys.* **56** 3297
[28] Kumar P and Singh C P 2017 New agegraphic dark energy model in Brans-Dicke theory with logarithmic form of scalar field *Astrophys. Space Sci.* **362** 52
[29] Akhlagi A, Malekjani M, Basilakos S and Haghi H 2018 Model selection and constraints from holographic dark energy scenarios *Mon. Not. Roy. Astron. Soc.* **477** 3659

[30] Srivastava M and Singh C P 2018 New holographic dark energy model with constant bulk viscosity in modified $f(R, T)$ gravity theory *Astrophys. Space Sci.* **363** 117

[31] Singh C P and Srivastava M 2018 Viscous cosmology in new holographic dark energy model and the cosmic acceleration *Eur. Phys. J.* C **78** 190

[32] Koussour M, Filali H, Shekh S H and Bennai M 2022 Holographic dark energy in Gauss-Bonnet gravity with Granda-Oliveros cut-off *Nucl. Phys.* B **978** 115738

[33] Oliveros A, Sabogal M A and Acero M A 2022 Barrow holographic dark energy with Granda-Oliveros cut-off *Eur. Phys. J. Plus* **137** 783

[34] Eckart C 1940 The thermodynamics of irreversible processes III: relativistic theory of the simple fluid *Phys. Rev.* **58** 919

[35] Landau L D and Lifshitz E M 1987 *Fluid Mechanics* (Stoneham: Butterworth Heinemann)

[36] Hiscock W A and Lindblom L 1985 Generic instabilities in first-order dissipative relativistic fluid theories *Phys. Rev.* **31** 725

[37] Israel W and Stewart J M 1979 Transient relativistic thermodynamics and kinetic theory *Ann. Phys.* **118** 341

[38] Grøn Ø 1990 Viscous inflationary universe models *Astrophys. Space Sci.* **173** 191

[39] Brevik I and Odintsov S D 2002 Cardy-Verlinde entropy formula in viscous cosmology *Phys. Rev.* D **65** 067302

[40] Brevik I and Gorbunova O 2005 Dark energy and viscous cosmology *Gen. Relativ. Gravit.* **37** 2039

[41] Ren J and Meng X H 2006 Cosmological model with viscosity media (dark fluid) described by an effective equation of state *Phys. Lett.* B **633** 1

[42] Hu M G and Meng X H 2006 Bulk viscous cosmology: Statefinder and entropy *Phys. Lett.* B **635** 186

[43] Fabris J C, Goncalves S V B and de S Ribeiro R 2006 Bulk viscosity driving the acceleration of the Universe *Gen. Relativ. Grav.* **38** 495

[44] Setare M R 2007 Interacting holographic generalized Chaplygin gas model *Phys. Lett.* B **654** 1

[45] Cruz N, Lepe S and Peña F 2007 Dissipative generalized Chaplygin gas as phantom dark energy *Phys. Lett.* B **646** 177

[46] Meng X H, Ren J and Hu M G 2007 Friedmann cosmology with a generalized equation of state and bulk viscosity *Commun. Theor. Phys.* **47** 379

[47] Singh C P, Kumar S and Pradhan A 2007 Early viscous universe with variable gravitational and cosmological constants *Class. Quantum Grav.* **24** 455

[48] Singh C P 2008 Bulk viscous cosmology in early universe *Pramana J. Phys.* **71** 33

[49] Feng C J and Li X Z 2009 Viscous Ricci dark energy *Phys. Lett.* B **680** 355

[50] Li B and Barrow J D 2009 Does bulk viscosity create a viable unified dark matter model? *Phys. Rev.* D **79** 103521

[51] Meng X H and Dou X 2009 Singularities and entropy in bulk viscosity dark energy model *Commun. Theor. Phys.* **52** 377

[52] Avelino A and Nucamendi U 2009 Can a matter-dominated model with constant bulk viscosity drive the accelerated expansion of the universe? *J. Cosmol. Astropart. Phys.* JCAP04(2009)006

[53] Avelino A and Nucamendi U 2010 Exploring a matter-dominated model with bulk viscosity to drive the accelerated expansion of the Universe *J. Cosmol. Astropart. Phys.* JCAP08(2010)009

[54] Xu Y D, Huang Z G and Zhai X H 2012 Generalized Chaplygin gas model with or without viscosity in the w-w' plane *Astrophys. Space Sci.* **337** 493

[55] Singh C P and Kumar P 2014 Friedmann model with viscous cosmology in modified $f(R, T)$ gravity theory *Eur. Phys. J.* C **74** 3070

[56] Brevik I, Elizalde E, Odintsov S D and Timoshkin A V 2017 Inflationary universe in terms of a van der Waals viscous fluid *Int. J. Geom. Meth. Mod. Phys.* **14** 1750185

[57] Brevik I, Grøn Ø, de Haro J, Odintsov S D and Saridakis E N 2017 Viscous cosmology for early- and late-time universe *Int. J. Mod. Phys.* D **26** 1730024

[58] Odintsov S D, Oikonomou V K, Timoshkin A V, Saridakis E N and Myrzakulov R 2018 Cosmological fluids with logarithmic equation of state *Ann. of Phys.* **398** 238

[59] Singh C P and Kumar A 2018 Ricci dark energy model with bulk viscosity *Eur. Phys. J. Plus* **133** 312

[60] Singh C P and Srivastava M 2018 Viscous cosmology in new holographic dark energy model and the cosmic acceleration *Eur. Phys. J.* C **78** 190

[61] Singh C P and Kaur S 2020 Probing bulk viscous matter dominated model in Brans-Dicke Theory *Astrophys. Space Sci.* **365** 2

[62] Ren J and Meng X-H 2007 Dark viscous fluid described by a unified equation of state in cosmology *Int. J. Mod. Phys.* D **16** 1341

[63] Avelino A, Garcia-Salcedo R, Gonzalez T, Nucamendi U and Quiros I 2013 Bulk viscous matter-dominated universes: asymptotic properties *J. Cosmol. Astropart. Phys.* JCAP08(2013)012

[64] Sasidharan A and Mathew T K 2015 Bulk viscous matter and recent acceleration of the universe *Eur. Phys. J.* C **75** 348

[65] Sasidharan A and Mathew T K 2016 Phase space analysis of bulk viscous matter dominated universe *J. High Energy Phys.* JHEP06(2016)138

[66] Singh C P and Kumar A 2019 Observational constraints on viscous Ricci dark energy model *Astrophys. Space Sci.* **364** 94

[67] Narayan R and Bartelmann M 1996 Lectures on gravitational lensing arXiv:astro-ph/9606001

[68] Amante M H *et al* 2020 Testing dark energy models with a new sample of strong-lensing systems *Mon. Not. Roy. Astron. Soc.* **498** 6013

[69] Sharov G S and Vasiliev V O 2018 How predictions of cosmological models depend on Hubble parameter data sets *Math. Model. Geom.* **6** 1

[70] Stern D *et al* 2010 Cosmic chronometers: constraining the equation of state of dark energy. I: H(z) measurements *J. Cosmol. Astropart. Phys.* JCAP02(2010)008

[71] Simon J, Verde L and Jimenez R 2005 Constraints on the redshift dependence of the dark energy potential *Phys. Rev.* D **71** 123001

[72] Moresco M *et al* 2012 Improved constraints on the expansion rate of the Universe up to $z \sim 1.1$ from the spectroscopic evolution of cosmic chronometers *J. Cosmol. Astropart. Phys.* JCAP08(2012)006

[73] Zhang C *et al* 2014 Four new observational H(z) data from luminous red galaxies in the Sloan Digital Sky Survey data release seven *Res. Astron. Astrophys.* **14** 1221

[74] Moresco M *et al* 2016 A 6% measurement of the Hubble parameter at z 0.45: direct evidence of the epoch of cosmic re-acceleration *J. Cosmol. Astropart. Phys.* JCAP05(2016)014

[75] Ratsimbazafy A L *et al* 2017 Age-dating luminous red galaxies observed with the Southern African Large Telescope *Mon. Not. R. Astron. Soc.* **467** 3239

[76] Moresco M 2015 Raising the bar: new constraints on the Hubble parameter with cosmic chronometers at z $\approx$ 2 *Mon. Not. R. Astron. Soc. Lett.* **450** L16

[77] Gaztaaga E *et al* 2009 Clustering of luminous red galaxies—IV. Baryon acoustic peak in the line-of-sight direction and a direct measurement of H(z) *Mon. Not. R. Astron. Soc.* **399** 1663

[78] Oka A *et al* 2014 Simultaneous constraints on the growth of structure and cosmic expansion from the multipole power

spectra of the SDSS DR7 LRG sample *Mon. Not. R. Astron. Soc.* **439** 2515
[79] Wang Y *et al* 2017 The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: tomographic BAO analysis of DR12 combined sample in configuration space *Mon. Not. R. Astron. Soc.* **469** 3762
[80] Chuang C H and Wang Y 2013 Modelling the anisotropic two-point galaxy correlation function on small scales and single-probe measurements of H(z), DA(z) and f(z)s8(z) from the Sloan Digital Sky Survey DR7 luminous red galaxies *Mon. Not. R. Astron. Soc.* **435** 255
[81] Alam S *et al* 2017 The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: cosmological analysis of the DR12 galaxy sample *Mon. Not. R. Astron. Soc.* **470** 2617
[82] Blake C *et al* 2012 The WiggleZ dark energy survey: joint measurements of the expansion and growth history at $z < 1$ *Mon. Not. R. Astron. Soc.* **425** 405
[83] Chuang C H *et al* 2013 The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: single-probe measurements and the strong power of f(z)Sigma(8)(z) on constraining dark energy *Mon. Not. R. Astron. Soc.* **433** 3559
[84] Anderson L *et al* 2014 The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: baryon acoustic oscillations in the Data Releases 10 and 11 Galaxy samples *Mon. Not. R. Astron. Soc.* **441** 24
[85] Busca N G *et al* 2013 Baryon acoustic oscillations in the Lyα-Forest of BOSS quasars *Astron. Astrophys.* **552** A96
[86] Bautista J E *et al* 2017 Measurement of baryon acoustic oscillation correlations at $z = 2.3$ with SDSS DR12 Lyα-Forests *Astron. Astrophys.* **603** A12
[87] Delubac T *et al* 2015 Baryon acoustic oscillations in the Lyα forest of BOSS DR11 quasars *Astron. Astrophys.* **574** A59
[88] Font-Ribera A *et al* 2014 Quasar-Lyman α forest cross-correlation from BOSS DR11: baryon acoustic oscillations *J. Cosmol. Astropart. Phys.* JCAP05(2014)027
[89] Scolnic D M *et al* 2018 The complete light-curve sample of spectroscopically confirmed sne ia from pan-STARRS1 and cosmological constraints from the combined pantheon sample *Astrophys. J.* **859** 101
[90] Conley A *et al* 2011 Supernova constraints and systematic uncertainties from the first three years of the supernova legacy survey *Astrophys. J. Suppl. Ser.* **192** 1
[91] Betoule M *et al* 2014 Improved cosmological constraints from a joint analysis of the SDSS-II and SNLS supernova samples *Astron. Astrophys.* **568** A22
[92] Reid M J, Pesce D W and Riess A G 2019 An improved distance to ngc 4258 and its implications for the hubble constant *Astrophys. J.* **886** L27
[93] Foreman-Mackey D, Hogg D, Lang D and Goodman J 2012 emcee: The MCMC Hammer *Publ. Astron. Soc. Pac.* **125** 306
[94] Aghanim N(Planck Collaboration) *et al* 2020 Planck 2018 results. VI. Cosmological parameters *Astron. Astrophys.* **641** A6
[95] Mamon A A 2018 Constraints on a generalized deceleration parameter from cosmic chronometers *Mod. Phys. Lett.* A **33** 1850056
[96] Capozziello S, Agostine R D and Luongo O 2020 High-redshift cosmography: auxiliary variables versus Padé polynomials *Mon. Not. Roy. Astron. Soc.* **494** 2576
[97] Sahni V *et al* 2003 Statefinder-A new geometrical diagnostic of dark energy *JETP Lett.* **77** 201
[98] Alam U *et al* 2003 Exploring the expanding universe and dark energy using the statefinder diagnostic *Mon. Not. R. Astron. Soc.* **344** 1057
[99] Dunajski M and Gibbons G 2008 Cosmic jerk, snap and beyond *Class. Quantum Gravity* **25** 235012