SCHOLARLY PUBLICATIONS

A CURRENT AWARENESS BULLETIN
OF RESEARCH OUTPUT

@DTU

(109th Edition)

JANUARY 2022

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

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

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

This issue contains the information published during **January, 2022**. 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 Life Cycle Approach to Environmental Assessment of Wastewater and Sludge Treatment Processes, ***6.Kulvendra Patel***, DTU

2. Novel method to Design FIR Digital Filter Using Whale Optimization, Suman Yadav, ***3.Manjeet Kumar***, Richa Yadav and Ashwni Kumar, Electronics, DTU

3. A Numerical Study of Analog Parameter of Negative Capacitance Field Effect Transistor with Spacer, ***6.Yash Pathak***, ***3.Bansi Dhar Malhotra*** and ***3.Rishu Chaujar***, Physics and Biotechnology, DTU

4. A review on Finger vein based Recognition, ***3.Ajai Kumar Gautam*** and ***3.Rajiv Kapoor***, Electronics, DTU

5. A visual atlas of genes' tissue-specific pathological roles, Priyadarshini Rai, Atishay Jain2, Neha Jha, Divya Sharma, Shivani Kumar, Abhijit Raj, ***6.Apoorva Gupta***, Sarita Poonia, Smriti Chawla, Angshul Majumdar, Tanmoy Chakraborty, Gaurav Ahuja and Debarka Sengupta, Biotechnology, DTU

6. An approach for automated descriptive answer verification, Abesh Jha and ***8.Nikhil Mishra***, Computer, DTU

7. An optimized weighted voting based ensemble model for DDoS attack detection and mitigation in SDN environment, ***6.Aastha Maheshwari***, ***8.Burhan Mehraj***, ***8.Mohd Shaad Khan*** and ***8.Mohd Shaheem Idrisi***, CSE, DTU

8. Approximation of Jain Operators by Statistical Convergence, Nurhayat Ispir, ***3.Naokant Deo*** and Neha Bhardwaj, Mathematics, DTU

9. Assessment of Gestational Age by New-born Joint Angles, Arnab Nandy, Tanushree Mondal, ***8.Divyoshanu Mondal Ivan***, Tapti Sengupta, Sankar Kumar Das, D. Goldar, Avijit Hazra and Rakesh Mondal, Biotechnology, DTU

10. Assessment of the effect of the judicial prohibition on firecracker celebration at the Diwali festival on air quality in Delhi, India, ***6.Shailendra Kumar Yadav***, ***3.Rajeev Kumar Mishra***, Bhola Ram Gurjar, Environmental, DTU

11. Augmentative role of Piriformospora indica fungus and plant growth promoting bacteria in mitigating salinity stress in Trigonella foenum-graecum, ***7.Sanskriti Bisht***, ***7.Shatrupa Singh***, ***7.Madhulika Singh***, ***3.Jai Gopal Sharma***, Biotechnology, DTU

12. Bézier variant of summation-integral type operators, ***6.Neha***, ***3.Naokant Deo*** and Ram Pratap, Applied Mathematics, DTU

13. Blockchain-Based Medical Certificate Generation and Verification for IoT-based Healthcare Systems, Suyel Namasudra, ***6.1.Pratima Sharma***, Ruben Gonzalez Crespo and Vimal Shanmuganathan, CSE, DTU

14. CanarDeep: a hybrid deep neural model with mixed fusion for rumour detection in social data streams, Deepak Kumar Jain, Akshi Kumar and ***7.Akshat Shrivastava***, CSE, DTU

15. Closed Loop Control Using RSAPS Algorithm for 5-Level CHB Multilevel Inverter, ***6.Praveen Bansal*** and ***3.Alka Singh***, Electronics, DTU

16. Coefficient Estimates for Certain Analytic Functions, ***6.Surya Giri*** and ***3.S. Sivaprasad Kumar***, Applied Mathematics, DTU

17. Combating Workplace Toxicity with the Help of Workplace Spirituality: A Study of Indian Public Sector Oil and Gas Companies, ***3.Naval Garg***, ***7.Ashi Rohtagi***, ***7.1.Chaitanya Mittal*** and ***7.Deepali Sharma***, DSM, DTU

18. Comparison of TiO2 catalysis and Fenton's treatment for rapid degradation of Remazol Red Dye in textile industry effluent, ***7.Harsh Pipil***, ***6.Shivani Yadav***, ***6.Harshit Chawla***, ***6.Sonam Taneja***, ***6.Manisha Verma***, ***8.Nimisha Singla*** and ***3.A. K. Haritash***, Environmental, DTU

19. CONCENTRATED SOLAR POWER (CSP) HYBRIDISATION: TRANSFORMING THE ENERGY SECTOR OF INDIA, ***8.Sourabh Yadav***, ***8.Shubham Sharma***, ***8.Shubham Maurya***, ***8.Sandeep Yadav*** and ***3.Dr. J.P Kesari***, Mechanical, DTU

20. Cybersecurity training for Web Applications through Serious Games, Devottam Gaurav, Yash Kaushik, Santhoshi Supraja, Abhi Khandelwal, ***8.Karsheet Negi***, Manmohan Prasad Gupta and Manmohan Chaturvedi, Design, DTU

21. Data Congestion Prediction in Sensors Based IoT Network, ***3.Rajesh K Yadav***, ***6.Aastha Maheshwari*** and Prem Nath, CSE, DTU

22. Delineation of groundwater recharge potential zones for its sustainable development utilizing GIS approach in Kabul basin, Afghanistan, ***3.Santosh Kumar Singh*** and ***6.Ali Reza Noori***, Environmental, DTU

23. DepthNet: A Monocular Depth Estimation Framework, ***6.Anunay***, ***6.Pankaj*** and ***3.Chhavi Dhiman***, Mechanical, CSE and Electronics, DTU

24. Design and Analysis of MIMO FSO System and WDM FSO System for Leh (Ladakh), India under Worst Weather Conditions, ***6.Rekha Rani*** and ***3.Gurjit Kaur***, Electronics, DTU

25. Down-shifting photoluminescence studies of thermally stable Dy3+ ions doped borosilicate glasses for optoelectronic device applications, ***6.Rajat Bajaj***, Aman Prasad, A. V. S. Yeswanth, ***6.Pooja Rohilla***, ***6.Sumandeep Kaur*** and ***3.A. S. Rao***, Applied Physics, DTU

26. Effect of Gate Oxide Material Variability on The Analog Performance of T-Gate GaN-MOS-HEMT with Graded Buffer, ***6.Megha Sharma***, ***6.Bhavya Kumar*** and ***3.Rishu Chaujar***, Applied Physics, DTU

27. Effect of Plasma Control Parameters on the Growth of Nitrogen-Doped Nanocone-Vertical Graphene Hybrid: Theoretical Investigations, ***6.Shruti Sharma*** and ***3.Suresh C. Sharma***, Applied Physics, DTU

28. Effect of synthetic and aromatic amine antioxidants on oxidation stability, performance, and emission analysis of waste cooking oil biodiesel, ***6.Khushbu Yadav***, ***3.Naveen Kumar*** and ***3.Rajiv Chaudhary***, Mechanical, DTU

29. Electric Vehicle Charging Standards in India and Safety Consideration, ***7.Sakshi Chalia***, ***6.Aakash Kumar Seth*** and ***3.Mukhtiar Singh***, Electrical, DTU

30. Electrostatic Analysis of Ferroelectric and High Dielectric Layer Assisted MOSFET, ***6.Rashi Mann*** and ***3.Rishu Chaujar***, Physics, DTU

31. Emotion Recognition Using Rational Dilation Wavelet Transform For Speech Signal, ***6.Ravi*** and ***3.Sachin Taran***, Electronics, DTU

32. Energy dissipation of flow over the type-B Piano Key Weir, ***8.Deepak Singh*** and ***3.Munendra Kumar***, Civil, DTU

33. Fireworks induced quasi-ultrafine particle number concentration and size-resolved elemental distribution in megacity Delhi, ***6.Shailendra Kumar Yadav***, ***3.Rajeev Kumar Mishra*** and Bhola Ram Gurjar, Environmental, DTU

34. Folksonomy Based Fuzzy Filtering Recommender System, ***3.Vinod Kumar***, ***8.1.Ayush Tanwar***, ***8.1.Arhan Relan*** and ***8.1.Chitraksh Grover***, CSE, DTU

35. Generalized Durrmeyer operators based on inverse Pólya–Eggenberger distribution, Minakshi Dhamija, ***3.Naokant Deo***, ***6.Ram Pratap*** & Ana Maria Acu, Maths, DTU

36. HR factors for the successful implementation of Industry 4.0: A systematic literature review, ***6.Anju Verma*** and Murugesan Venkatesan, DTU

37. HRV and GSR as Viable Physiological Markers for Mental Health Recognition, ***8.Shikha***, ***Aryan***, ***8.Lovish Arya*** and ***3.Divyashikha Sethia***, CSE and Electronics, DTU

38. Hydraulic Design and Analysis of Piano Key Weirs: A Review, ***8.Deepak Singh*** and ***3.Munendra Kumar***, Civil, DTU

39. Implementation of Dual Tree Complex Wavelet Transform with Mean Energy Features to detect Alzheimer's Disease, ***8.Neha Garg*** and ***3.Mahipal Chaudhary***, Electronics, DTU

40. Implementation of LSTM for Prediction of Diabetes using CGM, ***6.Sunny Arora***, ***3.Shailender Kumar*** and Pardeep Kumar, CSE, DTU

41. Magnification Dependent Histopathology Breast Cancer Image Binary Classification Using Different Levels Of EfficientNet, ***6.Saman Saadi Zadeh***, Electronics, DTU

42. Mapping of global, regional and national incidence, mortality and mortality-to-incidence ratio of lung cancer in 2020 and 2050, ***3.Rajesh Sharma***, USME, DTU

43. Measuring Human Auditory Attention with EEG, ***6.Chirag Ahuja*** and ***3.Divyashikha Setia***, CSE, DTU

44. Multi-functional Fractional-order Shadow Filter using OTA, ***6.Garima Varshney***, ***3.Neeta Pandey*** and ***3.Rajeshwari Pandey***, ECE, DTU

45. Nanomaterials for Remediation of Environmental Pollutants, Arpita Roy, ***8.Apoorva Sharma***, ***8.Saanya Yadav***, Leta Tesfaye Jule and Ramaswamy Krishnaraj, Biotechnology, DTU

46. New Multicarrier Modulation Scheme for Harmonics Mitigation of T-Type Solar Multilevel Inverter, Sanjay Upreti, Shivam Kumar Yadav, Bhim Singh and ***3.Narendra Kumar***, Electrical, DTU

47. Newly Realized Grounded Capacitance Multiplier Using Single CFDITA, ***6.Dheeraj Singh***, ***3.Deva Nand*** and Atul Kumar, Electronics, DTU

48. Numerical simulation of analog metrics and parasitic capacitances of GaAs GS-GAA FinFET for ULSI switching applications, ***6.Bhavya Kumar*** and ***3.Rishu Chaujar***, Applied Physics, DTU

49. Numerical study of passive mixing in a 3-dimensional helical micromixer with two inlets at offset, ***7.Akash Sinha*** and ***3.M. Zunaid***, Mechanical, DTU

50. Obstacle avoidance for a swarm of unmanned aerial vehicles operating on particle swarm optimization: a swarm intelligence approach for search and rescue missions, ***3.Girish Kumar***, ***8.1.Arham Anwar***, ***8.1.Abhinav Dikshit***, ***8.1.Abhirup Poddar***, Umang Soni and Weon Keun Song, Mechanical, DTU

51. ON A SUBFAMILY OF STARLIKE FUNCTIONS RELATED TO HYPERBOLIC COSINE FUNCTION, ***7.MRIDULA MUNDALIA*** AND ***3.S. SIVAPRASAD KUMAR***, APPLIED MATHEMATICS, DTU

52. ON POWER VALUES OF SUM OF DIVISORS FUNCTION IN ARITHMETIC PROGRESSIONS, ***6.SAI TEJA SOMU*** AND ***8.VIDYANSHU MISHRA***, APPLIED MATHEMTICS, DTU

53. Optimal Feature Selection for Imbalanced Text Classification, ***3.Anshu Khurana*** and ***3.Om Prakash Verma***, CSE and Electronics, DTU

54. Optimum techno-eco performance requisites for vacuum annulus tube collector–assisted double-slope solar desaltification unit integrated modified parabolic concentrator, ***6.Ashok Kumar Singh*** and ***3.Samsher Gautam***, Mechanical, DTU

55. Plastic waste management practices pertaining to India with particular focus on emerging technologies, Sampriti Kataki, ***6.Krithika Nityanand***, Soumya Chatterjee, Sanjai K. Dwivedi and Dev Vrat Kamboj, DTU

56. Prediction of an optimum engine response based ondi erent input parameters on common rail direct injection diesel engine: A response surface methodology approach, ***6.M. Kumar***, ***3.N.A. Ansari***, A. Sharma, ***6.V.K. Singh***, ***3.R. Gautam*** and Y. Singh, Mechanical, DTU

57. ReactiveFnJ: A Choreographed Model for Fork-Join Workflow in Serverless Computing, Urmil Bharti, ***3.1.Anita Goel*** and S. C. Gupta, DTU

58. Reflectocardiography (RCG): A leadless, optical reflection based facile measurement of human cardiac rhythm, Arindam Kushagra, Uddipan Dasgupta, Srishti Roychowdhury, ***7.Srijani Samanta*** and Shivani Srivastava, Biotechnology, DTU

59. Resolving stress of university teacher: exploring role of workplace spirituality and constructive workplace deviance, ***3.Naval Garg***, Sarika Kumari and B.K. Punia, USME, DTU

60. Review on fabrication methodologies and its impacts on performance of dye-sensitized solar cells, Geetam Richhariya, Bhim Charan Meikap and ***3.Anil Kumar***, Mechanical and Environment, DTU

61. REVIEW RESEARCH PAPER NEUROMORPHIC COMPUTING AND APPLICATIONS, ***8.Hitesh Dureja***, ***8.Yash Garg***, ***3.Dr. Rishu Chaujar*** and ***6.Bhavya Kumar***, Applied Physics, DTU

62. SHARP BOUNDS OF THIRD HANKEL DETERMINANT FOR A CLASS OF STARLIKE FUNCTIONS AND A SUBCLASS OF q-STARLIKE FUNCTIONS, ***6.SHAGUN BANGA*** AND ***3.S. SIVAPRASAD KUMAR***, APPLIED MATHEMATICS, DTU

63. SHORT TERM EFFECT OF PHYSICAL EXERCISE ON SELECTIVE ATTENTION USING EEG AND STROOP TASK, ***3.Divyashikha Sethia***, ***8.1.Taksh Kamlesh***, ***8.1.Mrigank Singh*** and Sonia Baloni Ray, CSE, DTU

64. Single OTRA-Based Multifunctional Filter Configuration, ***6.Garima Mann***, ***3.Pragati Kumar*** and ***3.D. R. Bhaskar***, Electrical, DTU

65. Software-Defined Networking: A Traffic Engineering Approach, ***6.Shanu Bhardwaj*** and ***3.Ashish Girdhar***, CSE, DTU

66. Solution of Transportation Problem Under Spherical Fuzzy Set, ***8.Vineet Kumar***, ***3.Anjana Gupta*** and ***3.H.C. Taneja***, Applied Mathematics, DTU

67. Spatio-temporal variations of quality of rainwater and stormwater and treatment of stormwater runoff using sand–gravel filters: case study of Delhi, India, ***6.Harsh Pipil***, ***3.A. K. Haritash*** and Krishna R. Reddy, Environmental, DTU

68. Static Performance Assessment of Junctionless Accumulation Mode Gate Stack Gate All Around (JAM-GS-GAA) FinFET Under Severe Temperature, ***6.Bhavya Kumar***, ***6.Megha Sharma*** and ***3.Rishu Chaujar***, Applied Physics, DTU

69. Stumpy-Paced Green Power Generation System, Ashok Kumar Singh, P. Kumar, R. Kishore, ***3.Phool Singh***, ***6.Neeraj Kant***, ***3.Pushpendra Singh***, Gopal Nandan, A. Jaiswal and R. Prasad, Mechanical, DTU

70. Synthesis optimisation and efficiency enhancement in Eu3+ doped barium molybdenum titanate phosphors for w-LED applications, ***6.Pooja Rohilla*** and ***3.A.S. Rao***, Applied Physics, DTU

71. Systematic Realization of CFOA Based Rössler Chaotic System and Its Applications, ***3.Kriti Suneja***, ***3.Neeta Pandey*** and ***3.Rajeshwari Pandey***, DTU

72. Thermal Performance Improvement of Solar Parabolic Dish System Using Modified Spiral Coil Tubular Receiver, Rajkumar Malviya, Prashant V. Baredar and ***3.Anil Kumar***, Mechanical and Environment, DTU

73. Third Order Low Frequency Sinusoidal Oscillators Employing CFOAs and Grounded Capacitors, ***7.Ajishek Raj***, ***3.Pragati Kumar*** and ***3.D. R. Bhaskar***, Electrical, DTU

74. Transfer learning for cancer diagnosis in histopathological images, Sandhya Aneja, ***7.1.Nagender Aneja***, Pg Emeroylariffion Abas and Abdul Ghani Naim, CSE, DTU

75. Watermarking of ECG signals compressed using Fourier decomposition method, Prashant Mani Tripathi, Ashish Kumar, Rama Komaragiri and ***3.Manjeet Kumar***, Electronics, 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*

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/357955446

A Life Cycle Approach to Environmental Assessment of Wastewater and Sludge Treatment Processes

Article *in* Water and Environment Journal · January 2022

DOI: 10.1111/wej.12774

CITATIONS
0

READS
17

2 authors, including:



Kulvendra Patel
Delhi Technological University
3 PUBLICATIONS **0** CITATIONS

SEE PROFILE

All content following this page was uploaded by Kulvendra Patel on 20 January 2022.

The user has requested enhancement of the downloaded file.

Patel Kulvendra (Orcid ID: 0000-0003-3068-6157)

Title: A Life Cycle Approach to Environmental Assessment of Wastewater and Sludge Treatment Processes

Author 1

Kulvendra Patel

Department of Environmental Engineering, Delhi Technological University, Delhi, 110042, India
ORCID No. 0000-0003-3068-6157

Author 2

S.K. Singh
Professor & Head, Department of Environmental Engineering, Delhi Technological University, Delhi, 110042, India
ORCID No. 0000-0001-8215-8562

Corresponding Author: Kulvendra Patel
Department of Environmental Engineering, Delhi Technological University, Delhi, 110042, India.
Email: - ravipatel71496@gmail.com

No. of Words in abstract – 150
No. of Words in main text – 5800
No. of Words in references – 1668
No. of Figures - 5
No. of Tables - 3

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/wej.12774

This article is protected by copyright. All rights reserved.

A Life Cycle Approach to Environmental Assessment of Wastewater and Sludge Treatment Processes

Abstract

Life Cycle Assessment (LCA) is a tool for determining a product or a process's environmental effects. In this study, LCA evaluates the critical sources of environmental impacts in wastewater treatment plants (WWTPs). The aim of the study is to conduct a comparative LCA of four wastewater treatment (WWT) methods: activated sludge process, sequential batch reactor, constructed wetlands, and up-flow anaerobic sludge blanket, including the sludge treatment methods.: Anaerobic Digestion (AD), AD with Pre-treatment, Lime Stabilization (LS), and LS with energy recovery. Environmental impacts are analyzed using the IMPACT 2002+ approach and Simapro 9.1 software using the Ecoinventv3.6 database. The study result shows that the sequential batch reactor has the highest environmental effects due to its high energy consumption. Electricity used for treatment is the primary contributor to environmental impacts. The study helps in evaluating the potential environmental impacts corresponding to a WWTP and identifying the hotspots associated with the process.

Keywords: - Anaerobic Digestion, Global Warming Potential**,** Life Cycle Assessment, Sequential Batch Reactor, Simapro

Accepted Article

This article is protected by copyright. All rights reserved.

INTRODUCTION

Sustainability is a significant concern in many industries, especially regarding essential broad-spectrum resources like water (Beery & Repke, 2010). Wastewater treatment plants (WWTPs) are essential for achieving sustainable development goals (SDG) by reclaiming valuable resources like water and nutrients (Tabesh, et al., 2018). In many countries, wastewater generation is increasing as a result of population growth and improved living standards. Wastewater treatment plants are built and operated with the goal of reducing wastewater pollution and minimizing emissions and harmful effects on the environment and human health (Wang, et al., 2012; Aenab & Singh, 2015). Despite continuing attempts to support the implementation of wastewater treatment systems worldwide, there is still no access to improved sanitation for about 2500 million people (WHO & UN-Water, 2014). The procurement of technologies for wastewater treatment is another problem facing municipal authorities. Only 20 percent of worldwide wastewater production obtains effective treatment in accordance with the Fourth World Water Development Report (UNESCO, 2012).

It was necessary to choose a regional scenario to focus on during the LCA process. An Indian scenario was adopted as a reference in this case, but the qualitative information obtained is of global significance. In India, Municipal wastewater production in urban areas is about 61,754 MLD. Currently, there is only 22,963 MLD of capacity available for treating municipal wastewater. Due to a capacity deficit in WWTs, almost 62 percent of overall wastewater is drained directly into water bodies, leaving a significant gap in urban WWT (Kamble, et al., 2019). By 2050, it is expected that 132 billion liters of wastewater would be produced every day, expanding the gap even further (Bhardwaj, 2005). Thus, an overall analysis shows that in the coming years, there will be a double-edged challenge to cope with owing to increasing population and industrialization: reduced freshwater supply and increasing wastewater

This article is protected by copyright. All rights reserved.

generation. In India, wastewater is currently utilized for irrigation, gardening, flushing, and industrial process water. There are currently 1469 Sewage Treatment Plants (STPs) in India, out of which 1093 are operational, 102 are non-operational, and 274 are under construction. The compliance status of 900 STPs is available from 1093 operational STPs, and only 578 STPs with a combined capacity of 12,200 MLD are found to comply with the State Pollution Control Board's approved standards (CPCB, 2021). The activated sludge process is the most often used technology in class 1 cities (populations of 100,000 or more) in India, accounting for 59.5 % of total installed capacity. It is followed by Up-flow anaerobic sludge blanket technology, accounting for 26 % of total installed capacity. In addition, 28 % of the plants use the Waste Stabilization Ponds (Kaur, et al., 2012; Gupta, et al., 2018).

LCA is a methodology for analyzing processes and product's inputs, outputs, and environmental impacts for their entire life cycle which is known as “cradle-to-grave” analysis. The analysis covers the life-cycle of a product, process, or system, including raw material extraction and processing, manufacture, transportation and distribution, usage, reuse, and final disposal (Consoli, et al., 1993). Typical system boundaries included in LCA are cradle-to-gate (from resource extraction to the factory gate), gate-to-gate (from one defined point to second point), and gate-to-grave (from factory gate to use and disposal phase). In the past, LCA has proven to be a valuable method for calculating a wastewater treatment method's environmental footprint. Municipal wastewater and sludge include rich resources such as water, organic matter, energy, and nutrients (phosphorus and nitrogen) that may be recovered for various economic, social, and environmental uses (Mateo-Sagasta, et al., 2015). LCA has been used in various researches globally to discuss the environmental footprint and resource use of WWTPs. LCA has been used in various researches globally to discuss the environmental footprint and resource use of WWTPs. Tabesh, et al., (2018) evaluated the environmental impacts of Tehran’s WWTP using LCA. The results depicts that the use of biogas instead of natural gas

This article is protected by copyright. All rights reserved.

alleviate the effects of WWTP. The analysis also shows that from an eutrophication perspective the usage of treated wastewater for irrigation was more environmental friendly. Singh & Kazmi, (2018) analysed the 16 decentralized WWTP based on different technologies adopted in northern India. The technologies considered were – sequential batch reactor, moving bed bioreactor, rotating biological contractor. The result shows that the moving bed bioreactors yielded best effluent quality, also the aerobic treatment systems were more efficient than anaerobic systems. Cartes, et al., (2018) evaluated the environmental impacts of sludge treatment scenarios in Chile. The methods included were anaerobic digestion (with and without pre-treatment), and chemical stabilisation (with and without energy recovery facility). The analysis shows that the anaerobic digestion process with pre-treatment have exhibits lower impacts than chemical stabilisation methods. The goal of Awad, et al., (2019) was to determine the environmental impacts for improvement of a conventional municipal WWTP in developing countries. WWTP of Gamasa, Egypt was taken as a case study (plant consist of grit chambers, settling tanks, aeration tanks and sludge thickeners and drying beds). The study comprises of four scenarios; scenario 1 was the existing WWTP in Gamasa; scenario 2 Gamasa WWTP + anaerobic digester; scenario 3 Gamasa WWTP + tertiary treatment; and scenario 4 Gamasa WWTP + anaerobic digester + tertiary treatment. The result shows that the scenario 4 was very beneficial for all impact categories due to improvement of water saving and energy efficiency. Hao, et al., (2019) analysed the environmental impacts of a WWTP in China and compared with potential resource recovery options, particularly the thermal energy recovery. The findings clearly showed that traditional WWTPs do not improve the ecosystem as a whole, but the proposed scheme does due to resource/energy recovery-based procedures. Thermal energy recovery was critical in achieving a net-zero LCA analysis. Another study by Morsy, et al., (2020) compares the two scenarios; Scenario 1, the existing WWTP at Abu-Rawash providing only primary treatment and Scenario 2, adopting the secondary treatment unit. A number of

This article is protected by copyright. All rights reserved.

LCA studies analysed practical application as case studies including water and wastewater treatment systems, water recycling facilities, and sludge treatment plants (Tillman, et al., 1998; Pasqualino, et al., 2009; Singh, et al., 2016; Alyaseri & Zhou, 2017). These studies shows that the environmental impacts were mainly due to the consumption of energy and recommended to use more renewable form of energy. However, these do not usually involve treatment of sludge generated. Each technology generates sludge of differing consistency and quantities, necessitating a range of sludge management options. The wastewater treatment technologies must be evaluated in conjunction with appropriate sludge disposal methods. A WWT plant is considered efficient if it can harvest the potential energy contained in the organic matter and nutrients found in the wastewater. The plant's sludge treatment unit is where organic matters are processed for energy and nutrient recovery (Lazarova, et al., 2012). Johansson, et al., 2008 has shown that it is important to minimize direct gas emissions and recover nutrients for the environmental outcome of sludge treatment. The best sludge management options for each technology are analyzed in this study, and system boundaries are defined as a result. Sludge treatment methods (anaerobic digestion, lime stabilization) are included in the study, as well as emissions generated during the operation and maintenance of the sludge treatment plant. In several areas, wastewater treatment results in increased sludge production, and sludge removal approaches are becoming increasingly restricted (Renou, et al., 2008). Organic waste produced will be subject to more stringent landfill regulations, and incineration will face considerable social opposition in many countries. In comparison to the main wastewater treatment processes, sewage sludge disposal was perceived as a minor concern. However, in the last couple of years, owing to the strong growth in sludge production and the strengthening of its disposal regulations, it became one of the relevant pollution control systems (Suh & Rousseaux, 2002). Combining these unit processes is the solution to a variety of scientific and environmental concerns about wastewater and sludge treatment. However, this situation requires a systematic

This article is protected by copyright. All rights reserved.

process of evaluating and comparing wastewater treatment plant's environmental efficiency. The functional unit used for the analysis was 1 m^3 of treated wastewater. The study shows that the consumption of electricity to carry out treatment has significant contribution in all impact categories. The use of secondary treatment has positive impact in reduction of eco-toxicity, eutrophication and human toxicity.

In water politics throughout the world, a paradigm shift is necessary to prevent future damage to sensitive ecosystems and the aquatic environment and emphasize that wastewater is an essential resource to efficiently manage water security. The research goal is to conduct LCA on different wastewater and sludge treatment methods to evaluate the resources consumption, pollutant emissions and their consequent environmental impacts of alternative wastewater treatment methods during their operation period. The analysis outcomes will give insights into the environmental hotspots of the treatment methods and suggest areas for improvement. The 'gate-to-grave' approach is applied in the present work; it means that the system boundary is considered when the wastewater is received to the plant site and ends when sludge is treated and finally disposed. The LCA is formulated using the principles and guidelines from the ISO 14040/14044:2006 standards.

MATERIAL AND METHODS

In LCA, the comparable performance systems or products or procedures are selected by considering the impacts of their life cycle caused by alternatives. LCA enables decision-makers to compare two products of procedures and choose the one with the least environmental effect. LCA is a modelling method that evaluates a product's environmental implications across its whole life cycle, from raw material extraction to disposal or recycling. It may also be used to identify the most critical environmental impacts in the life phase and to establish baselines for

This article is protected by copyright. All rights reserved.

potential research. The environmental effects are assessed by determining and quantifying the resources and products used and the atmosphere's emissions throughout the life cycle (Garfí, et al., 2017). ISO 14040 and ISO 14044 LCA guidelines characterize four distinct but theoretically iterated phases in the LCA – (i) goal and scope, which involves a clear and consistent definition of the study; (ii) perform a life cycle inventory (LCI), based on energy and mass quantification; (iii) perform a life cycle impact assessment (LCIA) utilizing the information collected at the LCI phase in conjunction with the state-of-the-art impact assessment method; (iv) interpretation of the results and make some general conclusions and recommendations. LCA may be used to analyze and evaluate the contributions of various activities or systems to global environmental impacts. The assumptions made in this study are: (i) The BOD inlet is assumed to be the average in India (200 mg/l) for all treatment plants (CPCB, 2009); (ii) and sludge transportation distance for all WWTPs is assumed to be 20 km; (iii) according to previously published literature (Tillman, et al., 1998; Machado, et al., 2007; Kalbar, et al., 2012), the construction phase of a WWTP in India accounts for around 1% of the entire life cycle impacts of the WWTP. As a result, the construction phase is excluded from this research and focuses on the operation phase of the WWTPs; (iv) LCAs need the usage of software such as SimaPro or openLCA. These software's often incorporate a number of inventory databases. Since the databases were created mainly in the European context, they must typically be adapted when used in other regions as there are no national databases available in India to carry out the LCA, and the production of material and emission inventories has become difficult (Kalbar, et al., 2012). In this study, Simapro 9.1.1.1 software and Ecoinvent v3.6 database are used.

Accepted Article

This article is protected by copyright. All rights reserved.

Goal and Scope

The study aims to compare the environmental impacts of the four wastewater treatment methods: (i) activated sludge process (ASP), (ii) constructed wetland (CW), (iii) sequential batch reactor (SBR), and (iv) Up-flow anaerobic sludge blanket (UASB). The sludge from these wastewater treatments goes to considered sludge treatment methods (anaerobic digestion, anaerobic digestion with pre-treatment, lime stabilization, and lime stabilization with energy recovery), respectively. These methods are most common treatment alternatives in developing countries (Kamble, et al., 2019). According to (CPCB, 2013) the WWT methods used in India were ASP, Moving Bed Biological Reactor (MBBR), Membrane Bio Reactor (MBR), UASB, Waste Stabilization Pond (WSP), and SBR. As per the availability of inventory data, four WWT methods were chosen for the study. Figure-1 shows the boundary of the considered system of the analysis. The precise selection of system boundaries has a significant impact on LCA. The system boundary starts when raw wastewater goes for primary treatment and terminates at the point of application of treated sludge. First-order environmental effects, such as direct atmospheric pollution and effluent discharge, are included under the system boundary. It also considers second-order impacts such as electrical emissions and the material resources required for the treatment processes. The Functional Unit (FU) specifies how the product's functions are quantified. A FU's primary purpose is to serve as a point of reference for relating inputs and outputs. This point of reference is required to ensure LCA outcomes are comparable; this is especially important when comparing various systems and ensuring that comparisons are performed on a common basis**.** Several analyses consider population equivalent (p.e.) as FU, but they calculate it in very different ways; thus, they were not directly comparable. One p.e. represents 50g of BOD5 load per day in India (Arceivala & Asolekar, 2007). Other functional units used in research are m^3 (Kamble, et al., 2017; Singh, et al., 2018); kg of sludge

This article is protected by copyright. All rights reserved.

(dry basis) (Suh & Rousseaux, 2002; Hospido, et al., 2005; Cartes, et al., 2018). According to (Zang, et al., 2015), WWT related LCA should use more than one FU to define the system under study and avoid ambiguities in the conclusions.

Two separate FU have been used in this study since the LCA covers both wastewater and sludge treatment. For WWT, FU is 1 m^3 of treated wastewater, and for sludge treatment, FU is 1kg of sludge (dry basis) is considered.

Life Cycle Inventory

This is the LCA stage, which involves gathering system inputs and outputs for the end-of-life phase of a process. The system inputs are raw materials and energy, and the system outputs are materials such as compost, pollutants to air, water, and soil. Data of input/output, including raw materials, products, and emissions, are collected and recorded for each system boundary phase. The life cycle inventory (LCI) is critical since it directly influences the impact assessment phase and is important for the interpretation of results and conclusions. The inventory used in this study was compiled using the Ecoinvent v3.6 database and previously published research on the considered WWT methods (Tangsubkul, et al., 2005; Kalbar, et al., 2012; Cartes, et al., 2018). Since there is no central database on all types of inventories in India, creating a life cycle inventory of all units is challenging. More information on each WWTP and its inventories can be found in the subsections below.

Activated Sludge Process

ASP is the most commonly used WWT system. Although there are several different designs, all ASPs have three main components: an aeration tank (coarse bubble aeration) where microorganisms metabolize suspended and soluble organic matter; a settling tank for the

This article is protected by copyright. All rights reserved.

separation of activated sludge (AS) and effluent; and returned AS equipment to move settled AS from the sedimentation tank to aeration tank. All of these elements are included in the study's scope. WWT by ASP has high removal efficiencies up to 85-90% for BOD, 80-95% for COD, 85-90% for Total Suspended Solids (TSS), and 60-75% phosphorus (Kumar, et al., 2010; Sastry, et al., 2013; Chandrakant, et al., 2015; Daud, et al., 2018). Material and emission inventory of this analysis is generated for a 50 KLD sewage treatment plant based on (Kalbar, et al., 2012). Waste-activated-sludge from the main treatment process is further treated using an anaerobic digestion (AD) process. Sludge from WWT is thickened prior to AD to reduce its volume. The ASP flow diagram is shown in Figure -2(a). In terms of solids recovery, a 90 percent efficiency was assumed (Hobson, 2003). The biogas produced is used to generate electricity. A loss of 5 percent of total biogas generated, with 50 percent directly emitted and the other 50 percent released and emitted during storage, was taken into account (Cakir & Stenstrom, 2005). It was assumed that dewatering would take place in a centrifuge using polyacrylamide, and the sludge would then be transported 20 km to the application site in a 21-metric-ton lorry. Table-1 shows the inventory data of ASP.

Constructed Wetland

CWs are integrated structures designed with natural processes, including wetland plants and microbial assemblies, assisting with wastewater management (Vymazal, 2010). According to Arceivala & Asolekar, 2007, the CWs are designed for a BOD_5 of 200 mg/l. With a sufficient detention time, BOD, TSS, COD, metals, and persistent organic compounds may be effectively removed from municipal wastewater (USEPA, 2000). With a much longer detention time, Nitrogen is predominantly removed by nitrification, subsequent denitrification, and ammonia volatilization at higher pH levels caused by algal photosynthesis (Vymazal, 2010). Constructed

This article is protected by copyright. All rights reserved.

Wetlands have high removal efficiencies for all parameters up to 85 % - 95% for COD, 80 - 95% for BOD, 85 - 95% for TSS and 90 - 98% for ammonia (Šereš, et al., 2017; Stefanakis, 2020; Alayu & Leta, 2021). When evaluating the environmental effect of CWs, it is important to account that when CO_2 is sequestered in biomass, CW will serve as a carbon sink. CWs, but on the other hand, release significant quantities of greenhouse gases (de Klein & van der Werf, 2014) especially CH_4 and N_2O. In the study, Phragmites karka (P. karka) is assumed to be used to treat sewage. P. karka is a kind of aquatic weed that grows on the surface of the water. Because of their fast development, emergent macrophytes appear to be promising for long-term wastewater treatment. The Phragmites plant provides oxygen to the heterotrophic microbes in the constructed wetland, which is an essential function for WWT (Chavan & Dhulap, 2012). The CWs pollution inventory is shown in Table-1. After the primary treatment, wastewater enters the horizontal flow wetland, causing solids to settle. After being treated in CW, sludge is thickened, then chemically stabilized with lime, dewatered, and transported to a landfill site for disposal. A 21-metric-ton lorry was assumed to have been used to transport the sludge over a 20-km distance. The Figure – 2 (b) depicts the phase flow diagram of CWs.

Sequential Batch Reactor

SBR is a modification of ASP. Wastewater is applied to a single batch reactor in this method and then processed and discharged to eliminate undesirable components. Equalizing, aerating, and clarification will all be done with a single batch reactor. They are generally used for WWT with low or intermittent-flow conditions (Gupta, et al., 2012). The process diagram of SBR is shown in Figure – 2 (c). SBR is a sewage treatment technology that is still under development, and the appropriate management platform and management mode must be updated and enhanced regularly. The disadvantage of conventional SBR technology includes an extended

This article is protected by copyright. All rights reserved.

treatment cycle, frequent switching of inlet and outlet water, high needs for equipment automation, high amount of energy, and considerable water level changes (Jing, et al., 2021). SBR has very high removal efficiencies up to 90 - 95% for BOD, 90 – 96% for COD, 92- 97% for TSS, 80 – 90 % for ammonia, and 80- 90% for nitrogen and phosphorus (Khalil, et al., 2008; Singh & Kazmi, 2018). The inventory data is generated for 33 MLD capacity SBR plant (Gupta & Singh, 2012) based on the design procedure given in (Metcalf & Eddy, 2003). The sludge from the SBR process is assumed to be subjected to anaerobic digestion before applying to the land. The inventory data of SBR is given in Table - 1.

Up-flow Anaerobic Sludge Blanket

The UASB reactor is the widely used anaerobic treatment system for wastewater. UASB is a suspended growth reactor that promotes granulation to sustain a high microbial biomass concentration (Khanal, et al., 2017). Anaerobic treatment systems such as UASB are being promoted due to several advantages: low operation cost, low excess sludge production, high COD removal efficiency, ability to cope with pH, influent concentration fluctuations, and energy generation in the form of biogas. Anaerobic WWT was formerly not employed because they could not manage large organic loading rates and might be unstable in treatment effectiveness. The UASB overcomes these problems since it can withstand high organic loading rates, and adding a recycle promotes biological population development, resulting in shorter start-up periods. It also demonstrates the traditional benefits of most anaerobic wastewater treatment techniques, such as biogas production, reduced sludge formation, and low land requirements compared to aerobic systems (Dutta, et al., 2018). The removal efficiencies of UASB are up to 80-90% for BOD, 70-80% for COD, and 85-85% for TSS (Khalil, et al., 2008; Saliba & von Sperling, 2017; Daud, et al., 2018). The removal efficiency

This article is protected by copyright. All rights reserved.

for nitrogen and phosphorus is about 60-65% (CPCB, 2013). Table 1 shows the emission inventory. A75 MLD capacity sewage treatment plant based on (Kalbar, et al., 2012) is used to generate LCI. UASB sludge is assumed to be treated with lime stabilization before being disposed off, having an energy recovery facility. The UASB phase flow diagram is shown in Figure – 2 (d).

Life Cycle Impact Assessment (LCIA)

An LCIA aims to identify and quantify the potential environmental consequences of a product, operation, or service. LCIA comprises several steps: choosing the appropriate impact categories, assigning the elementary flows, characterization, normalization, grouping, and weighting (Nieuwlaar, 2013). LCIA methods take two different approaches: midpoint and endpoint. The midpoint method embraces the categories in the middle of the cause-effect chain without computing their final damages, such as human health or environmental damages. Global warming (kg CO_2 eq.), Acidification (kg SO_2 eq.), ionizing radiation (Bq C-14 eq.), and eutrophication (kg $PO4^{3-}$ eq.) are some examples of these midpoint categories. The endpoint method takes it a step further by converting the midpoint effect categories into more specific human and ecosystem damage categories.

LCIA methods attempt to link each LCI to the respective environmental impacts. The LCI results are categorized according to ISO 14044 into impact categories with an indicator. There are many approaches for calculating environmental effects. The most often used LCIA approaches are CML, IMPACT 2002+, Eco-indicator 99, and Eco-points 97. Damage assessment and weighting are not supported by the CML method. Eco-point 97 can calculate impacts in a single score; however, it does so by using 30 separate impact categories and does not reflect the ultimate damages to human health or resource depletion. The IMPACT 2002+

This article is protected by copyright. All rights reserved.

approach was used in this analysis. This method consists of four other methods: IMPACT 2002, Eco-indicator 99, CML (Guinée, et al., 2002) and Intergovernmental Panel on Climate Change (IPCC). The method consistently combines midpoint and endpoint categories. This method includes 15 midpoint level categories grouped into four damage impact scores or damage categories: Human Health, Ecosystem Quality, Climate Change, and Resources. Table - 2 shows the midpoint and damage categories of IMPACT 2002+ (Jolliet, et al., 2003). The following impact categories are chosen based on the studied processes and emissions: Global Warming, Toxicity, Eutrophication, Acidification, Respiratory Effects, and Fossil Depletion. These impact categories were selected because they are the most relevant to WWT and WWTP is known to release emissions to the environment and freshwater resources, which are currently some environmental challenges worldwide (Tangsubkul, et al., 2005; Mamathoni & Harding, 2021). Midpoint characterization factors are based on equivalency rules, which means that midpoint characterization scores are calculated in kg-equivalents of a substance relative to a reference substance. This study does not consider model uncertainty or uncertainty from the LCIA methods.

RESULTS AND DISCUSSIONS

A comparative LCA of wastewater treatment methods was performed in this study. The overall energy consumption of the plants through their life cycle has been found to be SBR > ASP > UASB > CW, which is consistent with previous studies conducted in India (0 - 1 kWh/m^3) (Singh & Kazmi, 2018; Kamble, et al., 2019). The analysis shows that electrical consumption contributes the most to global impact categories such as global warming, abiotic depletion, and acidification. The WWTP does not have a direct influence on the surrounding environment. Its environmental effect is attributed to the production of electricity (Ioannou-Ttofa, et al., 2016).

This article is protected by copyright. All rights reserved.

This is due to the extraction and combustion of fossil fuels, which emit pollutants and CO_2 into the atmosphere. The data obtained from the analysis of impact categories are as follows: -

Global Warming Potential (GWP)

Figure – 3 (a) shows that CWs have negative CO_2 emissions (-6.63E-03 kg CO_2 eq.) due to sequestration of CO_2 during the treatment. The most important contributor to CO_2 emissions is electricity usage. Due to its high electricity consumption, the SBR process has the highest impact on climate change (0.77 kg CO_2 eq.) compared to other treatment processes. The higher energy requirements of SBRs compared to ASP is because SBRs are often built for both organic and nutrient removal, which takes more oxygen and hence more energy (Kalbar, et al., 2012). However, the sludge treatment method used for SBR has the least impact due to pre-treatment and energy recovery from the AD process. The negative global warming potential of CWs suggests that natural treatment systems can help to reduce global warming. The results of this analysis are in line with the findings of (Kamble, et al., 2019).

Toxicity Potentials

Toxicity potential depends mainly on the heavy metals released in the air, water, and soil environment.

Human Toxicity (HT) is due to carcinogenic and non-carcinogenic effects. HT calculates the cumulative toxicological risk and probable consequences of releasing a certain amount of a chemical into the environment. As shown in Figure - 3 (b), SBR has the highest HT potential due to non-carcinogens 8.58E-03 kg C_2H_3Cl eq. (mainly due to emission of Arsenic 7.13E-03 kg C_2H_3Cl eq. from the electrical process) followed by ASP 4.15E-03 kg C_2H_3Cl eq. The

This article is protected by copyright. All rights reserved.

Characterization Factors (CF) of aquatic ecotoxicity are given for emissions into the air, water, and soil and quantify the ecotoxicity effects on (surface) freshwater. Ecotoxicity is measured in kg of Triethylene glycol (TEG) / kg and is calculated by dividing the damage CF of the substance under consideration by the damage CF of the reference substance (Triethylene glycol) (Jolliet, et al., 2003). Terrestrial ecotoxicity is calculated similarly. SBR has the highest Aquatic and terrestrial ecotoxicity, 35.9 kg TEG water and 8.54 kg TEG soil, respectively, shown in Figure - 3 (c). However, the sludge treatment process AD with pre-treatment has negative toxicity potential due to avoided products (fertilizers as ammonium sulfate and diammonium phosphate) and energy generation.

Eutrophication Potential

Eutrophication in WWTP is mainly due to the release of phosphorus, nitrogen, and ammonia and, to some extent, degradable organics in wastewater effluent. Figure – 3 (d) shows that UASB system had the highest EP (2.89E-02 kg $PO4^{3-}$Eq) because there is negligible nutrient removal in the UASB system followed by ASP because of its nutrient removal capacity of 40-50%. SBR has the lowest eutrophication potential (6.67E-03 kg $PO4^{3-}$Eq) because of its intrinsic nutrient removal capacity of 70-80%. This result is in line with the findings of (Gallego, et al., 2008).

Acidification Potential (AP)

Acidification is caused due to the release of NOx and SOx associated with energy production. SBR is found to have the highest Aquatic and Terrestrial AP, 5.56E-03 kg SO_2 eq. and 1.63E-02 kg SO_2 eq., respectively, followed by ASP as shown in Figure – 3 (e). The impact is majorly

This article is protected by copyright. All rights reserved.

due to the emission of Nitrogen oxide (0.012 kg SO_2 eq. for Terrestrial AP) and Sulfur dioxide (3.87E-03 kg SO_2 eq. for Aquatic AP).

Respiratory Effects

This category refers to impacts that are caused by inorganic substances. The CFs are given for air emissions only. SBR has the highest respiratory effects (1.37E-03 kg $PM_{2.5}$ eq.) due to the release of particulates from the electrical process, followed by ASP (7.17E-04 kg $PM_{2.5}$ eq.). Figure – 3 (f) shows the comparison of respiratory effects caused by WWT methods.

Fossil Depletion

Coal consumption has a major contribution to Fossil depletion. The study shows that due to the high energy consumption of the SBR process, it has the highest fossil depletion potential (5.35 MJ), followed by ASP (2.72 MJ) shown in Figure - 4, which matches with the study by (Kalbar, et al., 2012).

Damage Categories

Endpoint indicators are considered to represent changes in stresses at a point in a cause-effect chain where the overall outcome might significantly impact society's perception of it. In IMPACT 2002+ method, the midpoint impacts are grouped into four damage or endpoint categories: Human Health represented in DALY ("Disability Adjusted Life Years") i.e. years lost to early mortality and reduced quality of life due to illness; Ecosystem determined in PDF*m^2*yr ("potentially disappeared fraction of species over a certain area in m^2 over a certain year") (Jolliet, et al., 2003); Climate Change similar as global warming and estimated in kg

This article is protected by copyright. All rights reserved.

CO_2 eq.; and Resources expressed in MJ. Table -3 shows the damage category values of considered WWT methods.

Normalization

Normalization is a technique for analysing each impact's contribution to total damage by adding normalization variables to impact groups to make analysis simpler. Comparing various groups on the same graph with the same units makes it simple to analyze the data. As per IMPACT2002+ impact guide, "Normalization is accomplished by dividing the impact by the appropriate normalization variables." Figure - 5 indicate the normalized values of the midpoint groups of the considered WWT methods.

CONCLUSION

The importance of LCA in Indian WWTPs is highlighted in this analysis. These techniques are capable of addressing environmental challenges and maximizing waste economic and technical aspects.

1. The result implies that SBR has the most potential for global warming and resource degradation because of high electricity consumption.
2. Due to the less capacity to remove nutrients, UASB has the highest aquatic eutrophication potential. CWs, in comparison to all other systems, have the fewest environmental impacts and have been shown to help mitigate global warming due to their carbon sequestering potential.
3. These studies suggest that any given technology will work well in one specific category of impact and poorly in another, making it hard to compare depending on the impact

This article is protected by copyright. All rights reserved.

categories. The CWs, accompanied by UASB, are the required wastewater management strategies in accordance with the system boundary of the study, with which the environmental effects of wastewater treatment can be reduced.

4. The analysis also shows that the sludge treatment method, anaerobic digestion with a pre-treatment facility, has the least impact on the environment, followed by anaerobic digestion (without pre-treatment) and Lime stabilization with an energy recovery facility. The impact of pre-treatment on digestive efficiency was related to its impact on energy recovery, transport demands, and loading of nutrients, emphasizing the need for a life cycle assessment.
5. Precise information on environmental or human health threats cannot be provided directly by LCAs for wastewater systems. However, LCA input and output data can be used in additional methods, such as risk evaluation. Local and regional priorities are inconsistent with present LCIA approaches, necessitating modifications to the CFs to account for regional variances.
6. A full-scale environmental and economic review of technologies should also be carried out using the long-term collected information in relation to potential study requirements. The comprehensive framework developed in the study will aid in the formulation of a suitable decision-making approach to choose the right treatment technology for wastewater.

Declaration of Competing Interest

The authors declare no conflicts of interest concerning the research, authorship, and publication of this article.

Accepted Article

This article is protected by copyright. All rights reserved.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

REFERENCES

Aenab, A. & Singh, S., 2015. Critical Assessment of River Water Quality and Wastewater Treatment Plant (WWTP). International Journal of Advanced Research, Volume 3, pp. 405-411.

Alayu, E. & Leta, S., 2021. Performance efficiency and water quality index of a two-stage horizontal subsurface flow constructed wetland system polishing anaerobically treated brewery effluent. Journal of Water Process Engineering, Volume 42.

Alyaseri, I. & Zhou, J., 2017. Towards better environmental performance of wastewater sludge treatment using endpoint approach in LCA methodology. Heliyon, 3(3).

Arceivala, S. J. & Asolekar, S. R., 2007. Wastewater Treatment for Pollution Control and Reuse. 3 ed. s.l.:McGraw-Hill Education / India.

Awad, H., Alalm, M. & El-Etriby, H., 2019. Environmental and cost life cycle assessment of different alternatives for improvement of wastewater treatment plants in developing countries. Science of the Total Environment, Volume 660, pp. 57-68.

Beery, M. & Repke, J., 2010. Sustainability analysis of different SWRO pre-treatment alternatives. Desalination and Water Treatment, 16(1-3), pp. 218-228.

This article is protected by copyright. All rights reserved.

Accepted Article

Bhardwaj, R., 2005. Status of Wastewater Generation and Treatment in India, Vienna: IWG-Env, International Work Session on Water Statistics.

Cakir, F. & Stenstrom, M., 2005. Greenhouse gas production: A comparison between aerobic and anaerobic wastewater treatment technology. Water Research, 39(17), pp. 4197-4203.

Cartes, J., Neumann, P., Hospido, A. & Vidal, G., 2018. Life cycle assessment of management alternatives for sludge from sewage treatment plants in Chile: does advanced anaerobic digestion improve environmental performance compared to current practices?. Journal of Material Cycles and Waste Management, 20(3), pp. 1530-1540.

Chandrakant, G., Jaswanth, P., Reddy, S. & Kiranmai, G., 2015. Design & Performance Evaluation of Wastewater Treatment Plant-D at Tirumala. International Journal of Scientific & Engineering Research, 6(7), pp. 1672-1688.

Chavan, B. & Dhulap, V., 2012. DESIGNING AND TESTING OF WASTEWATER IN CONSTRUCTED WETLAND USING PHRAGMITES KARKA. International Journal of Physical and Social Sciences, 2(12), pp. 205-211.

Consoli, F. et al., 1993. Guidelines for Life-Cycle Assessment:, s.l.: Society of Environmental Toxicology and Chemistry (SETAC).

CPCB, 2009. STATUS OF WATER SUPPLY, WASTEWATER GENERATION AND TREATMENT IN CLASS-I CITIES & CLASS-II TOWNS OF INDIA, s.l.: CONTROL OF URBAN POLLUTION, SERIES: CUPS/ 70 / 2009 - 10.

CPCB, 2013. Performance Evaluation of Sewage Treatment Plants under NRCD, New Delhi: Ministry of Environment and Forests, Govt. of India.

This article is protected by copyright. All rights reserved.

CPCB, 2021. National Inventory of Sewage Treatment Plants, Delhi: Central Pollution Control Board.

Daud, M. et al., 2018. Review of Upflow Anaerobic Sludge Blanket Reactor Technology: Effect of Different Parameters and Developments for Domestic Wastewater Treatment. Journal of Chemistry.

de Klein, J. J. & van der Werf, A. K., 2014. Balancing carbon sequestration and GHG emissions in a constructed wetland. Ecological Engineering, Volume 66, pp. 36-42.

Dutta, A., Davies, C. & Ikumi, D., 2018. Performance of upflow anaerobic sludge blanket (UASB) reactor and other anaerobic reactor configurations for wastewater treatment: a comparative review and critical updates. Journal of Water Supply: Research and Technology—AQUA, 67(8), pp. 858-884.

Gallego, A., Hospido, A., Moreira, M. & Feijoo, G., 2008. Environmental performance of wastewater treatment plants for small populations. Resources, Conservation and Recycling, 52(6), pp. 931-940.

Garfí, M., Flores, L. & Ferrer, I., 2017. Life Cycle Assessment of wastewater treatment systems for small communities: Activated sludge, constructed wetlands and high rate algal ponds. Journal of Cleaner Production, Volume 161, pp. 211-219.

Guinée, J. et al., 2002. Handbook on Life Cycle Assessment: Operational Guide to the ISO Standards. 1 ed. s.l.:Springer Netherlands.

Gupta, D. & Singh, S., 2012. Greenhouse Gas Emissions From Wastewater Treatment Plants: A Case Study of Noida. Journal of Water Sustainability, Volume 2.

Accepted Article

This article is protected by copyright. All rights reserved.

Gupta, S., Ramakrishnan, A. & Hung, Y., 2012. SEQUENCING BATCH REACTORS. In: Handbook of Environment and Waste Management. s.l.:s.n., pp. 511-562.

Gupta, S., Singh, S. & Gandhi, V., 2018. A study on Sewage Treatment and Disposal in Delhi. International Journal of Advance Research and Innovation, 6(2), pp. 88-91.

Hao, X. et al., 2019. Environmental impacts of resource recovery from wastewater treatment plants. Water Research, Volume 160, pp. 268-277.

Hobson, J., 2003. CH4 AND N2O EMISSIONS FROM WASTE WATER HANDLING. Good Practice Guidance and Uncertainty Management in National Greenhouse Gas Inventories.

Hospido, A. et al., 2005. Environmental Evaluation of Different Treatment Processes for Sludge from Urban Wastewater Treatments: Anaerobic Digestion versus Thermal Processes (10 pp). The International Journal of Life Cycle Assessment, 10(5), pp. 336-345.

Ioannou-Ttofa, L., Foteinis, S., Chatzisymeon, E. & Fatta-Kassinos, D., 2016. The environmental footprint of a membrane bioreactor treatment process through Life Cycle Analysis. Science of The Total Environment, Volume 568, pp. 306-18.

Jing, Y., Wei-li, L. & Fu-chuan, H., 2021. Review of SBR process in effluent treatment. s.l., E3S Web of Conferences, IAECST 2020.

Johansson, K. et al., 2008. Sewage sludge handling with phosphorus utilization – life cycle assessment of four

alternatives. Journal of Cleaner Production, 16(1), pp. 135-151.

Accepted Article

This article is protected by copyright. All rights reserved.

Jolliet, O. et al., 2003. IMPACT 2002+: A new life cycle impact assessment methodology. The International Journal of Life Cycle Assessment, 8(6).

Kalbar, P., Karmakar, S. & Asolekar, S., 2012. Assessment of wastewater treatment technologies: Life cycle approach. Water and Environment Journal, Volume 27.

Kamble, S. et al., 2017. A soil biotechnology system for wastewater treatment: technical, hygiene, environmental LCA and economic aspects. Environmental Science and Pollution Research, 24(15), pp. 13315-13334.

Kamble, S., Singh, A., Kazmi, A. & Starkl, M., 2019. Environmental and economic performance evaluation of municipal wastewater treatment plants in India: a life cycle approach. Water Science and Technology, 79(6), pp. 1102-1112.

Kaur, R., Wani, S., Singh, A. & Lal, K., 2012. Wastewater production, treatment and use in India.

Khalil, N., Sinha, R., Raghav, A. & Mittal, A., 2008. UASB TECHNOLOGY FOR SEWAGE TREATMENT IN INDIA: EXPERIENCE, ECONOMIC EVALUATION AND ITS POTENTIAL IN OTHER DEVELOPING COUNTRIES. Alexandria, Egypt, Twelfth International Water Technology Conference, IWTC12 2008.

Khanal, S., Giri, B., Nitayavardhana, S. & Gadhamshetty, V., 2017. 10 - Anaerobic Bioreactors/Digesters: Design and Development. Current Developments in Biotechnology and Bioengineering, pp. 261-279.

Kumar, K., Kumar, P. & Babu, M., 2010. PERFORMANCE EVALUATION OFWASTE WATER TREATMENT PLANT. International Journal of Engineering Science and Technology, 2(12), pp. 7785-7796.

This article is protected by copyright. All rights reserved.

Lazarova, V., Choo, K. & Cornel, P., 2012. Water - Energy Interactions in Water Reuse. s.l.:IWA Publishing.

Machado, A. et al., 2007. Life cycle assessment of wastewater treatment options for small and decentralized communities. Water Science and Technology, 56(3), pp. 15-22.

Mamathoni, P. & Harding, K., 2021. Environmental performance of extended activated sludge and sequential batch reactor using life cycle assessment. Cleaner Environmental Systems, Volume 2.

Mateo-Sagasta, J. et al., 2015. Global Wastewater and Sludge Production, Treatment and Use. In: Wastewater: Economic Asset in an Urbanizing World. Dordrecht: Springer, pp. 15-38.

Metcalf & Eddy, 2003. Wastewater Engineering: Treatment, Disposal, and Reuse. 4th edition revised by Tchobanoglous, G., Burton, F.L. and Stensel, D.H. ed. New York, USA: McGraw Hill Inc..

Morsy, K., Mostafa, M., Abdalla, K. & Galal, M., 2020. Life Cycle Assessment of Upgrading Primary Wastewater Treatment Plants to Secondary Treatment Including a Circular Economy Approach. Air, Soil and Water Research, 13(1), pp. 1-13.

Nieuwlaar, E., 2013. Life Cycle Assessment and Energy Systems. Reference Module in Earth Systems and Environmental Sciences.

Pasqualino, J., Meneses, M., Abella, M. & Castells, F., 2009. LCA as a Decision Support Tool for the Environmental Improvement of the Operation of a Municipal Wastewater Treatment Plant. Environmental Science & Technology, 43(9), pp. 3300-3307.

This article is protected by copyright. All rights reserved.

Renou, S., Thomas, J., Aoustin, E. & Pons, M., 2008. Influence of impact assessment methods in wastewater treatment LCA. Journal of Cleaner Production, 16(10), pp. 1098-1105.

Saliba, P. & von Sperling, M., 2017. Performance evaluation of a large sewage treatment plant in Brazil, consisting of an upflow anaerobic sludge blanket reactor followed by activated sludge. Water Science and Technology, 76(8), pp. 2003-2014.

Sastry, S., Rao, B. & Nahata, K., 2013. Study of Parameters Before And After Treatment of Municipal Waste Water From An Urban Town. Global Journal of Applied Environmental Sciences, 3(1), pp. 41-48.

Šereš, M., Hnátková, T., Vymazal, J. & Vaněk, T., 2017. REMOVAL EFFICIENCY OF CONSTRUCTED WETLAND FOR TREATMENT OF AGRICULTURAL WASTEWATERS. Chemistry Journal of Moldova, 12(1), pp. 45-52.

Singh, A., Kazmi, A.; Starkl, M.; Sayanekar, S.; Herlekar, M., 2018. Sewage management challenges in mega cities in India: a case study of Mumbai. DESALINATION AND WATER TREATMENT, Volume 116, pp. 329-341.

Singh, N., Banyal, P. & Kazmi, A., 2016. Techno-Economic Assessment of Full Scale MBBRs Treating Municipal Wastewater Followed by Different Tertiary Treatment Strategies: A Case Study from India. Nature Environment and Pollution Technology, 15(4), pp. 1311-1316.

Singh, N. & Kazmi, A., 2018. Performance and Cost Analysis of Decentralized Wastewater Treatment Plants in Northern India: Case Study. Journal of Water Resources Planning and Management, 144(3).

Accepted Article

This article is protected by copyright. All rights reserved.

Stefanakis, A., 2020. Constructed Wetlands for Sustainable Wastewater Treatment in Hot and Arid Climates: Opportunities, Challenges and Case Studies in the Middle East. Water, 12(6).

Suh, Y. & Rousseaux, P., 2002. An LCA of alternative wastewater sludge treatment scenarios. Resources, Conservation and Recycling, 35(3), pp. 191-200.

Tabesh, M., Feizee Masooleh, M., Roghani, B. & Motevallian, S., 2018. Life-Cycle Assessment (LCA) of Wastewater Treatment Plants: A Case Study of Tehran, Iran. International Journal of Civil Engineering, 17(7), pp. 1155-1169.

Tangsubkul, N. et al., 2005. Life Cycle Assessment of Water Recycling Technology. Water Resources Management, 19(5), pp. 521-537.

Tillman, A., Svingby, M. & Lundström, H., 1998. Life cycle assessment of municipal waste water systems. The International Journal of Life Cycle Assessment, 3(3), pp. 145-157.

UNESCO, 2012. Managing water under uncertainty and risk. The United Nations World Water Development Report 4, Paris: United Nations Educational, Scientific and Cultural Organization.

USEPA, 2000. Wastewater Technology Fact Sheet, s.l.: EPA.

Vymazal, J., 2010. Constructed Wetlands for Wastewater Treatment. Water, Volume 2, pp. 530-549.

Wang, X., Liu, J., Ren, N. & Duan, Z., 2012. Environmental profile of typical anaerobic/anoxic/oxic wastewater treatment systems meeting increasingly stringent treatment standards from a life cycle perspective. Bioresource Technology, Volume 126, pp. 31-40.

This article is protected by copyright. All rights reserved.

WHO & UN-Water, 2014. UN-water global analysis and assessment of sanitation and drinking-water (GLAAS) 2014 report: investing in water and sanitation: increasing access, reducing inequalities, s.l.: World Health Organization.

Zang, Y. et al., 2015. Towards more accurate life cycle assessment of biological wastewater treatment plants: a review. Journal of Cleaner Production, Volume 107, pp. 676-692.

Accepted Article

This article is protected by copyright. All rights reserved.

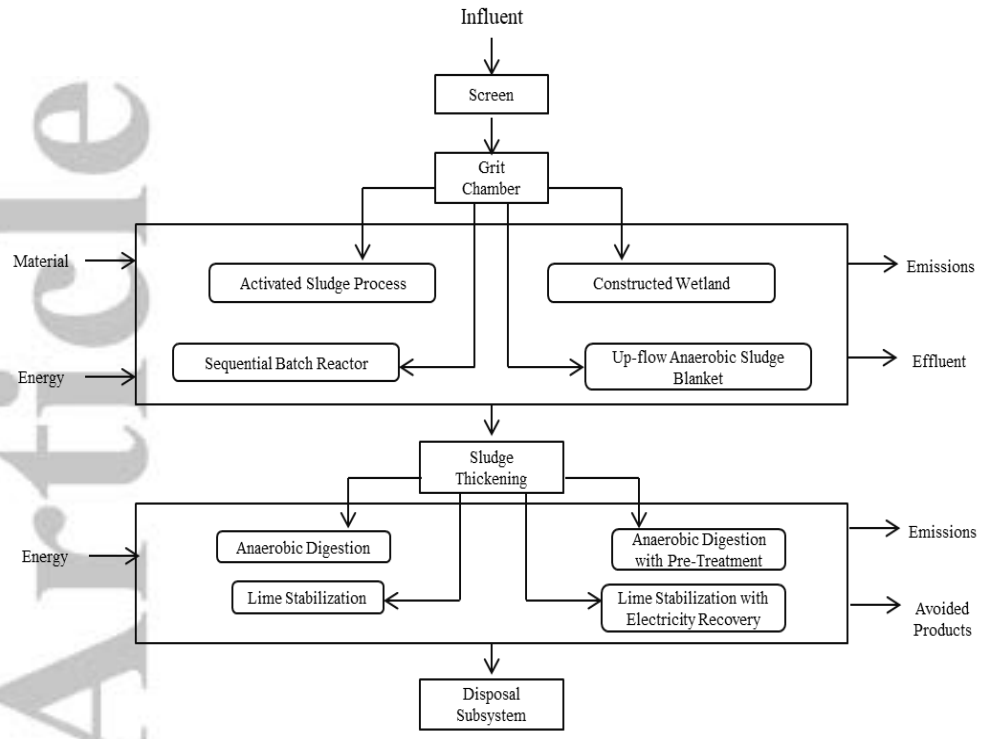


Figure - 1 System Boundary of the WWT Process

Accepted Article

This article is protected by copyright. All rights reserved.

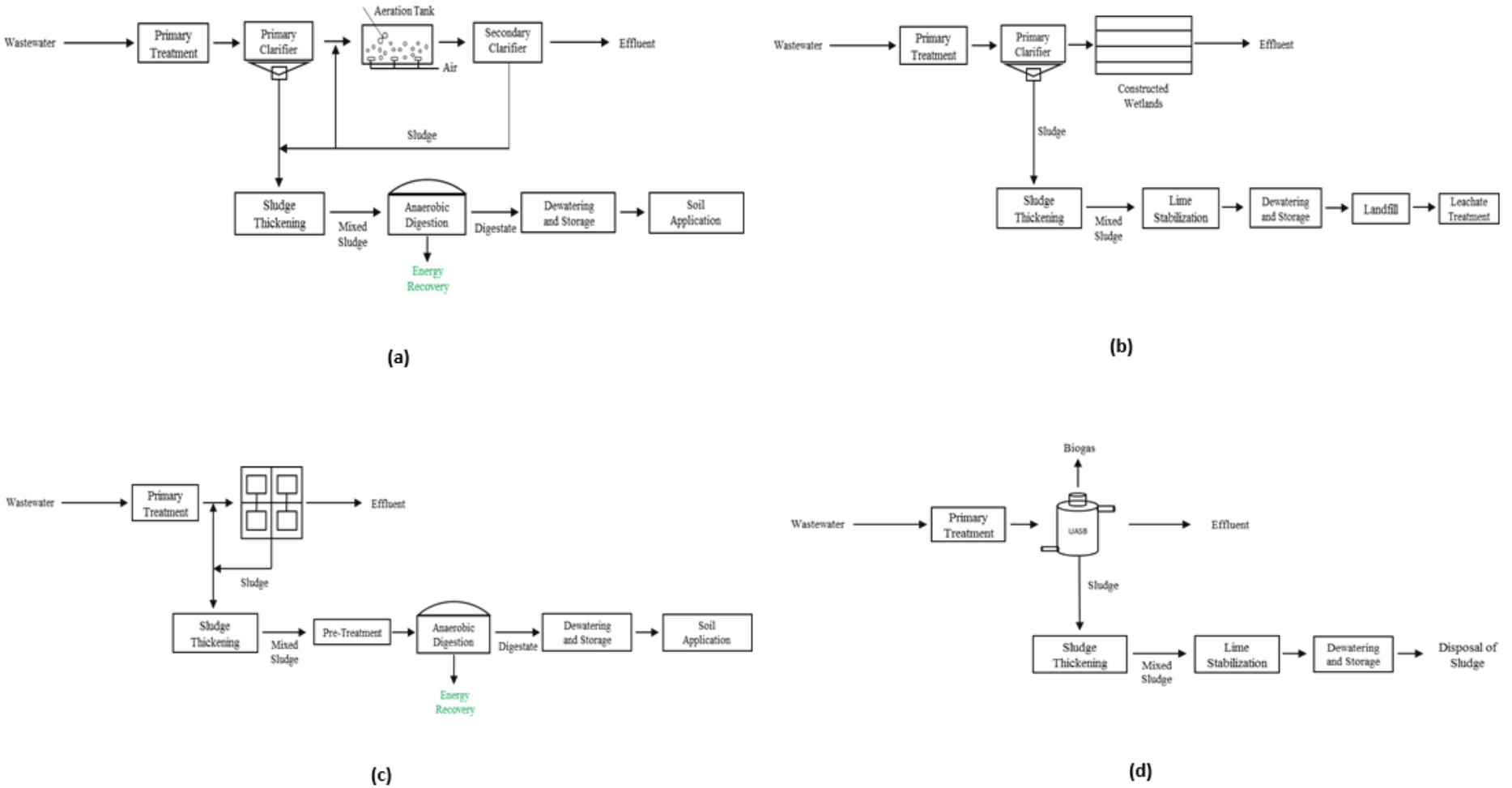
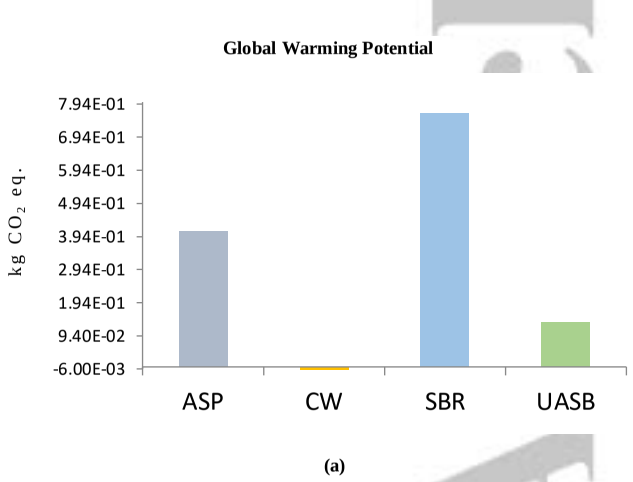
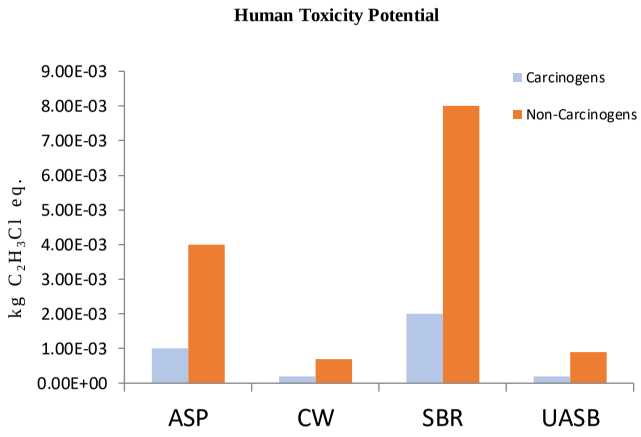


Figure – 2 Flow Diagram of (a) Activated Sludge Process (b) Constructed Wetlands (c) Sequential Batch Reactor (d) Up-flow Anaerobic Sludge Blanket

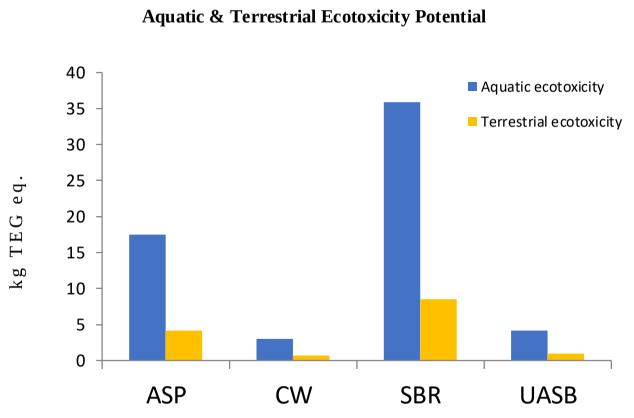
This article is protected by copyright. All rights reserved.



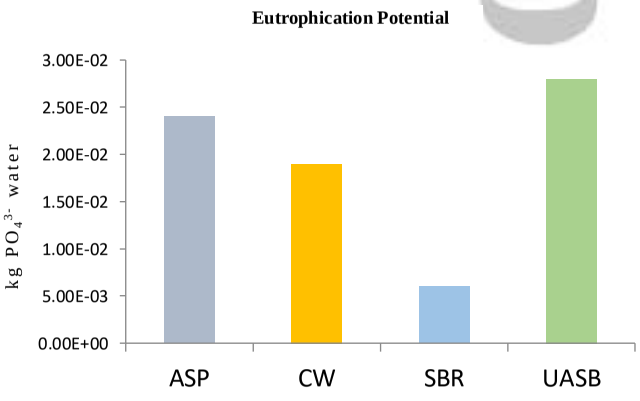
(a)



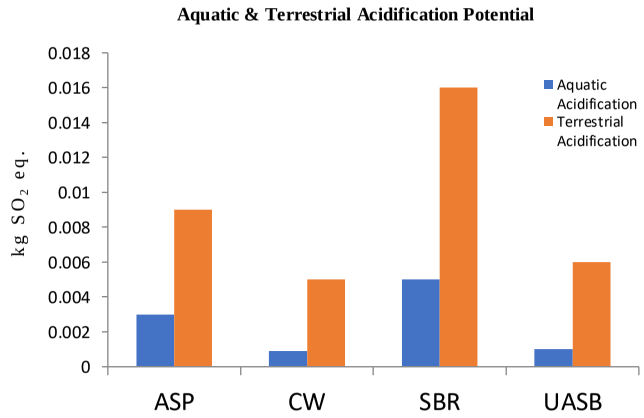
(b)



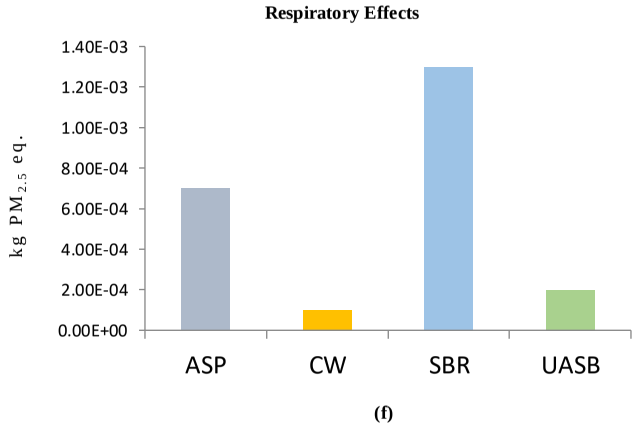
(c)



(d)



(e)



(f)

Figure - 3 Comparison of Environmental impacts of various Wastewater Treatment methods. ASP -Activated Sludge Process; CW- Constructed Wetlands; SBR- Sequential Batch Reactor; UASB-Up-flow Anaerobic Sludge Blanket

This article is protected by copyright. All rights reserved.

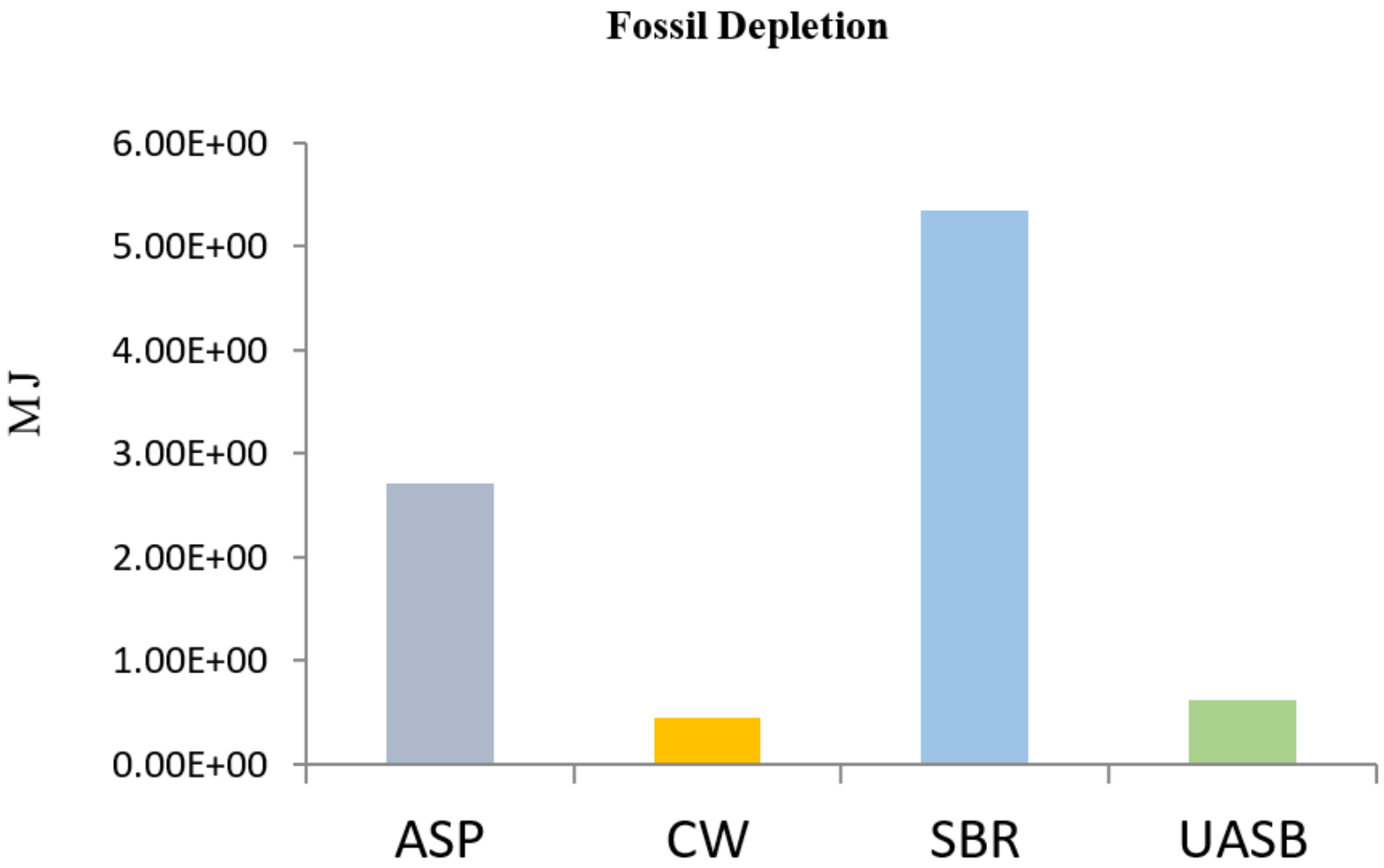


Figure - 4 Comparison of Fossil depletion potential of WWT method

This article is protected by copyright. All rights reserved.

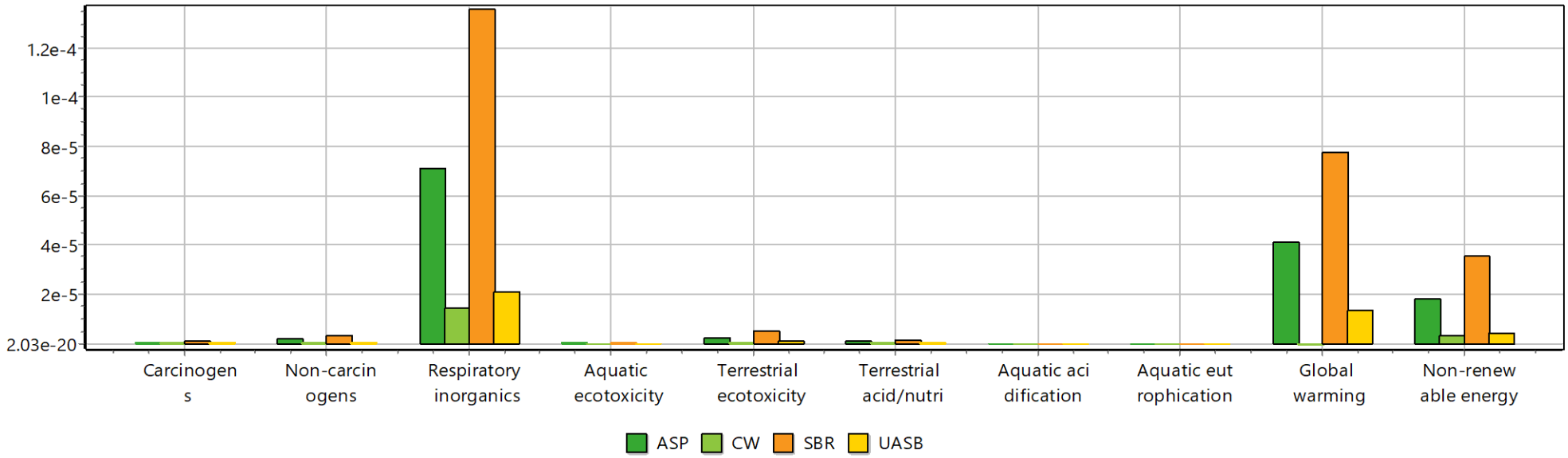


Figure - 5 Normalized values of midpoint impact categories of WWT methods

This article is protected by copyright. All rights reserved.

Table - 1 **Inventory Data of WWTPs (Tangsubkul, et al., 2005; Kalbar, et al., 2012; Cartes, et al., 2018)**

Parameters	Unit	ASP	CW	SBR	UASB
Wastewater Treatment*					
Input from Technosphere					
Electricity	kWh	0.173	0.0219	0.309	0.0329
Emissions to Air					
Carbon Dioxide	kg	0.197	-0.0438	0.351	0.0822
Sulfur dioxide	g	1.62	0.219	2.87	0.438
Nitrogen oxides	g	0.701	0.11	1.24	0.241
Carbon monoxide	g	0.932	0.219	1.6	0.625
Mercury	mg	0.0252	0.00329	0.0438	0.00471
Particulates	g	0.416	0.0515	0.723	0.0964
Emissions to Water					
COD	kg	0.070	0.0548	0.05	0.0986
Nitrogen, total	kg	0.0329	0.0219	0.0145	0.0986
Phosphorus, total	kg	0.006	0.005	0.000997	0.00767
Heavy metals	g	2.29	1.96	0.964	2.49
Emission to Soil					
Nitrogen, total	g	1.81	0.997	2.82	3.5
Phosphorus, total	g	0.449	0.252	0.707	0.877
Output to Technosphere					
Sludge	kg	0.05	0.02	0.01	0.02
Sludge Treatment**					
Input from Technosphere					
Electricity	kWh	0.166	0.0975	0.235	0.0975
Heat	kWh	0.64	-	0.65	-
Truck	tkm	0.074	0.06	0.1	0.06
Polyacrylamide	kg	0.0091	0.0091	0.0091	0.0091
Lime	kg	-	0.165	-	0.165
Avoided Products					
Ammonium sulfate, as N	kg	0.012	-	0.011	-
Phosphate fertiliser, as P2O5	kg	0.0016	-	0.0014	-
Electricity	kWh	0.7	-	0.84	0.0649

This article is protected by copyright. All rights reserved.

Heat	kWh	0.75	-	0.895	-
Emissions to Air					
Methane, biogenic	kg	0.0089	0.024	0.0106	0.024
Carbon dioxide, biogenic	kg	0.6	0.308	0.712	0.308
Nitrogen oxides	kg	8E-5	-	0.0001	-
Dinitrogen monoxide	kg	0.00017	-	0.00016	-
Hydrogen sulfide	kg	1E-5	-	1E-5	-
Ammonia	kg	0.0007	0.014	0.00085	0.014
Emissions to Water					
Phosphate	kg	0.0272	0.029	0.028	0.029
Nitrate	kg	0.3	0.262	0.264	0.262
Ammonium, ion	kg	0.058	0.06	0.058	0.06
Nitrogen dioxide	kg	0.0034	0.0036	0.0034	0.0036

* Functional Unit is per m^3

** Functional Unit is per kg

Accepted Article

This article is protected by copyright. All rights reserved.

Table - 2 Impact categories of IMPACT 2002+

Midpoint Category	Unit	Damage Category	Unit
Carcinogens	kg C_2H_3Cl eq.	Human Health	DALY
Non-Carcinogens	kg C_2H_3Cl eq.		
Respiratory inorganics	kg $PM_{2.5}$ eq.		
Respiratory organics	kg C_2H_4 eq.		
Ionizing radiation	Bq C-14 eq.		
Ozone Layer Depletion	kg CFC-11 eq.		
Aquatic ecotoxicity	kg TEG water	Ecosystem Quality	PDF*M^2*yr
Aquatic eutrophication	kg PO_4^{3-} limited		
Terrestrial ecotoxicity	kg TEG soil		
Aquatic acidification	kg SO_2 eq.		
Terrestrial acid/nutrient	kg SO_2 eq.		
Land occupation	m^2 organic arable land		
Global warming	kg CO_2 eq.	Climate Change	kg CO_2 eq.
Non-renewable	MJ	Resources	MJ
Mineral extraction	MJ		

Accepted Article

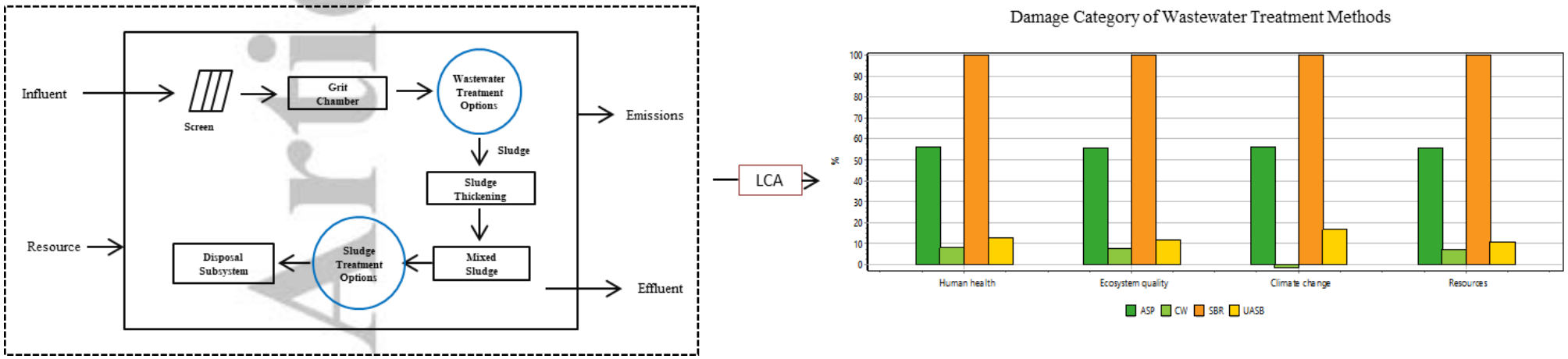
This article is protected by copyright. All rights reserved.

Table - 3 Damage category values of WWT methods

Damage Category	Unit	ASP	CW	SBR	UASB
Human Health	DALY	5.17E-07	1.04E-07	9.93E-07	1.54E-07
Ecosystem quality	PDF*m^2*yr	0.10	0.021	0.204	0.028
Climate change	kg CO_2 eq.	0.41	-0.006	0.77	0.13
Resources	MJ	2.72	0.45	5.35	0.62

Accepted Article

This article is protected by copyright. All rights reserved.



Graphical Abstract

This article is protected by copyright. All rights reserved.

View publication stats

A Novel method to Design FIR Digital Filter Using Whale Optimization

Suman Yadav
Department of Electronics and Communication Engineering
Indira Gandhi Technical University for Women
New Delhi, India
suman.yadav@bharatividyapeeth.edu

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

Richa Yadav
Department of Electronics and Communication Engineering
Indira Gandhi Technical University for Women
New Delhi, India
yadav.richa1@gmail.com

Ashwni Kumar
Department of Electronics and Communication Engineering
Indira Gandhi Technical University for Women
New Delhi, India
drashwanikumar@gmail.com

2021 IEEE Bombay Section Signature Conference (IBSSC) | 978-1-6654-1758-7/21/$31.00 ©2021 IEEE | DOI: 10.1109/IBSSC53889.2021.9673406

***Abstract*—This research work is about designing 20th order digital Finite impulse response filters of types low pass, high pass, band pass and band stop. Whale optimization algorithm (WOA) is utilized in designing the said filter by reducing the deviation of the frequency response of the designed filter from the ideal filter frequency response. The strategy is to get those optimal filter coefficients for which the difference in frequency responses is minimum. Some statistics have been used to compare the performance of WOA based digital filter with Cat Swarm Optimization (CSO) and Particle Swarm Optimization (PSO) based FIR filters in terms of pass band and stop band attenuation (dB) and ripples value in both the band of filter. It is observed that WOA based digital FIR filter have reduced ripples from 25 to 40 percent in both bands and have maximum attenuation in stop band.**

Keywords— filter design, cat swarm optimization algorithm, particle Swarm optimization, whale optimization

I. Introduction

Filters have been significant part of signal processing, in various areas as for processing of speech signals, satellite images, biomedical images, audio signals and in Communication System. Digital filters are mainly categorised in two parts, FIR (Finite Impulse response) and IIR (Infinite impulse response). This research work is based on the designing of digital FIR filter using optimization techniques.

There are some challenges while designing FIR filter to reduce ripples in like reduction of fluctuation in bands of filter, with attenuation in stop. The order of filter is also a factor in designing as it is related to hardware implementation and power consumption. Researchers are trying to establish a better trade-off among various specifications required in designing a filter. Earlier there were some traditional approaches for designing of a digital filter like window method, frequency sampling and Park-McClellan [1]. These techniques have drawbacks like to be struck in local minima, convergence speed is too slow and unable to achieve global optimum outcome. To overcome these above specified problems, different meta-heuristic optimization algorithm are used. These algorithms proved to be efficient in designing digital FIR filter. Optimization algorithms that have been used in designing digital FIR filter are Genetic Algorithm (GA) [2], Differential Evolution (DE) [3], Particle Swarm optimization (PSO) [4], Cuckoo Search Algorithm (CSA) [5], Artificial Bee Colony Algorithm (ABC)[6], Grey Wolf optimization (GWO) [7], Cat Swarm Optimization (CSO) [8], Bat Algorithm (BA) [9], Flower pollination Algorithm (BFA)[10-11], Grass hopper optimization algorithm [12-13], Vortex search algorithms [14-15], Biogeography optimization algorithm [16], Modified differential evolution ant colony optimization algorithm [17] etc.

Proposed work presents whale optimization algorithm (WOA) that provides an efficient exploration and exploitation strategy to attain global optimum output. It aids in designing FIR digital filter with reduced fluctuation in bands, also in approaching frequency response of ideal filter [18]. The unique feature of WOA is that it utilizes a spiral to simulate bubble net attacking methodology which makes it competitive with respect to other algorithms.

The research paper is organised as follows; FIR filter design model is detailed in Section 2. Segment 3 explains employed algorithm. Simulations outputs are given in part 4 and 5 of paper. Section 6 concludes the proposed work.

II. Design Model of Fir

The frequency response of FIR filter:

$$A(z) = \sum_{t=0}^{t-1} a(t) z^{-t} \qquad (1)$$

$$= a(0) + a(1)z^{-1} + \cdots \ldots + a(t-1)z^{t-1} \qquad (2)$$

$a(t)$ in (1) denotes the filter coefficients for filter to be designed whereas order of the filter is $t-1$ given in (2). The objective is to minimize the difference between the frequency response of the designed and ideal filter to acquire optimal filter coefficients. The idea is to approach frequency response of ideal filter to the extent possible.

Designed filter impulse response adhere to symmetrical requirement as mentioned in (3) below for linear phase:

$$a(t) = a(T-t) \quad for\ 0 \leq t \leq T \qquad (3)$$

A linear phase filter will avoid any loss in useful data at time of filtering. Following linear phase symmetry, half of the filter coefficients are determined, and leftover are flipped to obtain full impulse response.

978-1-6654-1758-7/21/$31.00 ©2021 IEEE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:52:11 UTC from IEEE Xplore. Restrictions apply.

Digital FIR filter magnitude response:

$$A(\omega) = \sum_{t=0}^{t-1} a(t) e^{-j\omega t} \quad (4)$$

$A(\omega)$ in (4) is the magnitude response for optimized coefficients. Filter can be LP, HP, BP, and BS based on $a(t)$ that is the response of filter.

Ideal magnitude response of LPF:

$$A_{id}(\omega) = \begin{cases} 1 & \forall \omega \in [0, \omega_p] \\ 0 & otherwise \end{cases} \quad (5)$$

Ideal magnitude response of HPF:

$$A_{id}(\omega) = \begin{cases} 0 & \forall \omega \in [0, \omega_p] \\ 1 & otherwise \end{cases} \quad (6)$$

Ideal magnitude response of BPF:

$$A_{id}(\omega) = \begin{cases} 1 & \forall \omega \in [\omega_l, \omega_u] \\ 0 & otherwise \end{cases} \quad (7)$$

Ideal magnitude response of BSF:

$$A_{id}(\omega) = \begin{cases} 0 & \forall \omega \in [\omega_l, \omega_u] \\ 1 & otherwise \end{cases} \quad (8)$$

$A_{id}(w)$ in (5-8) are magnitude response of ideal FIR digital filter. Ideal low pass filter (LPF) and high pass filter (HPF) have cut off frequency ω_p whereas ideal band pass filter (BPF) and band stop filter (BSF) have ω_l and ω_u as cut off frequencies of pass band and stop band.

Equation (9) represents error fitness function, proposed work uses mean absolute difference as error fitness function:

$$B(\omega) = \sum_{\forall \omega} (|A_{id}(\omega)| - |A(\omega)|) \quad (9)$$

$B(\omega)$ represents the deviation of designed filter frequency response from ideal filter frequency response. The objective is to reduce $B(\omega)$ as much as possible to achieve optimal filter coefficients.

III. Whale Optimization Algorithm

WOA is based hunting mechanism of humpback whales. This algorithm includes two main phases; first one is to random search for the prey that comes under exploration phase and second one use bubble net attacking method that comes under exploitation phase. Each phase is explained with detailed mathematical model in the following subsections.

A. Exploration Phase

In the given phase, humpback whales searches for the prey randomly in accordance with their position. The next position of the search agent is determined in keeping with arbitrary selected search agent instead of best search agent.

$$\vec{D} = \left| \vec{P} . \overrightarrow{Y_{rand}} - \vec{Y} \right| \quad (10)$$

$$\vec{Y}(t+1) = \overrightarrow{Y_{rand}} - \vec{B}.\vec{D} \quad (11)$$

$\vec{D}$ is a variable in (10) that defines the distance between position of a particular search agent $\vec{Y}$ and a arbitrarily selected search agent $\overrightarrow{Y_{rand}}$ from the population. It is done by choosing $|\vec{B}| > 1$, in (11) to focus on global search that is exploration.

$\vec{B}$ *and* $\vec{P}$ represents coefficient vectors, are evaluated as shown in eq. (12-13).

$$\vec{B} = 2 \cdot \vec{b} \cdot \vec{r} + \vec{b} \quad (12)$$

$$\vec{P} = 2 \cdot \vec{r} \quad (13)$$

$\vec{b}$ is a variable in (12) which decreased from 2 to 0 with increment in iteration. $\vec{r}$ in (13) is a random variable having value between [0 1]. After every iteration search agents modify their position to find out best search agent.

B. Encircling Prey

The locality of prey is identified by humpback whales. WOA presumes that victim prey is current foremost candidate solution and is very near to prime solution. Following the identification of location of best search agent, all agents in search space will attempt to modify their positions according to best search agent. This conduct is constituted mathematically through equations under shown:

$$\vec{D} = \left| \vec{P} \cdot \overrightarrow{Y^*}(t) - \vec{Y}(t) \right| \quad (14)$$

$$\vec{Y}(t+1) = \overrightarrow{Y^*}(t) - \vec{B} \cdot \vec{D} \quad (15)$$

Here t is present iteration in (14). $\overrightarrow{Y^*}(t)$ is the location of the best search agent (nearest to prey) till present iteration. $\vec{Y}(t)$ denotes position vector, whereas | | denoted absolute value. $\cdot$ indicates the multiplication done elementwise. $\vec{Y}(t+1)$ in (15) is the present position of the search agent.

The position of a particular search agent is modified depending on the location of current best search agent in search space. The value of $\vec{B}$ and $\vec{P}$ are varied to reach the locations near to best search agent according to current position.

C. Exploitation Phase (Bubble net attacking)

There are mainly two approaches under bubble net attaching method that is shrinking encircling mechanism and spiral updating location in exploitation phase.

- Shrinking encircling mechanism, humpback whales makes circles for capturing prey, shrinks slowly. It is accomplished with decreasing "b" from 2 to 0. B is also reducing and limited to [-b, b] band, so that circles can shrink. Updated position of the search agent can be located among actual location of search agent and location of best current search agent.
- Spiral Updating mechanism: A helix or a spiral shape is made through bubbles by humpback whales around the prey. The distance between the prey and

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:52:11 UTC from IEEE Xplore. Restrictions apply.

humpback whale during the creation of spiral shape is designed by the formula stated below in (16):

$$\vec{Y}(t+1) = \vec{D}.e^{bl}.\cos(2\pi l) + \overrightarrow{Y^*}(t) \qquad (16)$$

$\overrightarrow{Y^*}(t)$ gives the locality of prey and $\vec{Y}(t+1)$ is position of whale. There is 50 % probability of the mechanism chosen by humpback whales to enclose prey, that is shrinking encircling and spiral updating.

$$\vec{D} = \left|\overrightarrow{Y^*}(t) - \vec{Y}(t)\right|$$

$$\vec{Y}(t+1) = \overrightarrow{Y^*}(t) - \vec{B}\cdot\vec{D} \quad \text{for } p < 0.5$$

$$= \vec{D}.e^{bl}.\cos(2\pi l) + \vec{Y}(t) \quad \text{for } p \geq 0.5$$

p value lies between [0 1]

IV. Simulation Outcomes

A. Low Pass Filter

Simulations have been performed on MATLAB software. To design digital FIR LPF filter using WOA, cut off frequency has been set to $0.5 * \pi$. The number of iterations is taken as 500 to get the desired output. The size of populations is set to 30, whereas for PSO and CSO it is set to be 50, leads to lower computations in designing filter with proposed algorithms. The scope of frequency is between $[0 \text{ to } \pi]$. The range of frequency coefficients should be between -1 to 1. Error fitness function is minimized using WOA to get the optimized filter coefficients as given in Table 1. Magnitude response is plotted using these optimized these coefficients as shown in Fig.1.

TABLE 1. FILTER COEFFICIENTS OF FIR LPF AND HPF

$a(t)$	WOA (LPF)	WOA (HPF)
$a(0) = a(20)$	-0.0011	-0.0006
$a(1) = a(19)$	0.0353	0.0351
$a(2) = a(18)$	0.0006	0.0020
$a(3) = a(17)$	-0.0456	-0.0456
$a(4) = a(16)$	-0.0013	-0.0017
$a(5) = a(15)$	0.0636	0.0636
$a(6) = a(14)$	0.0016	0.0015
$a(7) = a(13)$	-0.1062	0.1062
$a(8) = a(12)$	-0.0006	-4.0e-05
$a(9) = a(11)$	0.3183	0.3183
$a(10)$	0.5013	0.4987

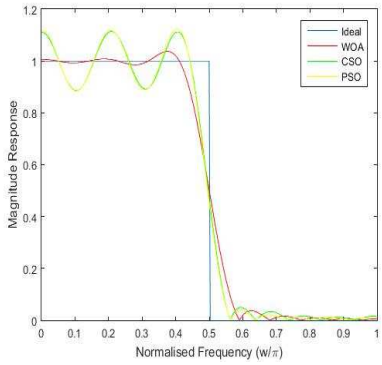


Fig. 1. Magnitude response of LPF of 20th order

B. High Pass Filter

Certain specifications has been set to design FIR high pass filter using proposed algorithm, cut off frequency for the pass band is 0.5π with 20th order. Filter coefficients should not extent beyond the range to -1 to +1. Based on optimal filter coefficients as mentioned in Table 1, the said filter is designed and its magnitude response is plotted in Fig 2.

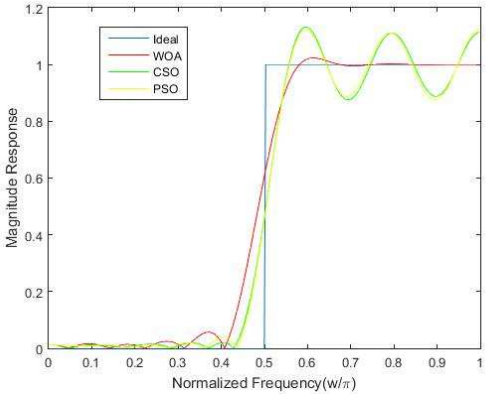


Fig. 2. Magnitude response of HPF of 20th order

C. Band Pass Filter

The specification for the BPF, are two cut off frequencies for the pass band and stop band which are fixed as 0.30π and $0.60\,\pi$ .WOA minimizes the difference among desired and ideal magnitude response of the filter, and obtain the optimal filter coefficient The range for the coefficients of filter is set between +1 and -1 for 20th order filter as mentioned in Table 2. Magnitude response for the designed filter is plotted in Fig. 4.

TABLE 2. FILTER COEFFICIENTS OF FIR BPF AND BSF

$a(t)$	WOA (BPF)	WOA (BSF)
$a(0) = a(20)$	0.0078	-0.0005
$a(1) = a(19)$	0.0006	0.0075
$a(2) = a(18)$	-0.0753	0.0761
$a(3) = a(17)$	-0.0030	-0.0274
$a(4) = a(16)$	0.0617	-0.0010
$a(5) = a(15)$	-0.0037	-0.0634
$a(6) = a(14)$	0.0951	-0.0950

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:52:11 UTC from IEEE Xplore. Restrictions apply.

$a(7) = a(13)$	-0.0014	0.1333
$a(8) = a(12)$	-0.3022	0.0012
$a(9) = a(11)$	-0.0002	0.4447
$a(10)$	0.3977	0.1007

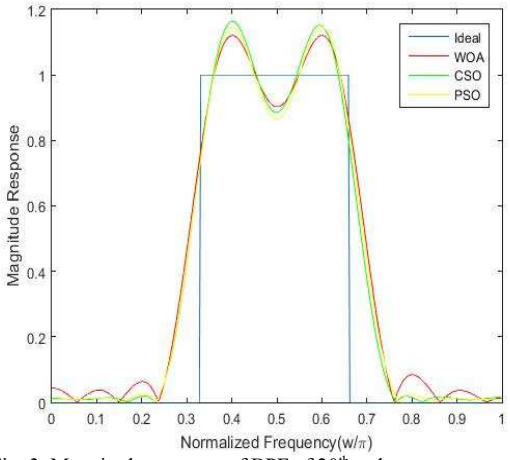


Fig. 3. Magnitude response of BPF of 20th order

D. Band Stop Filter

For BSF cut off frequencies are fixed to 0.33π and 0.66π for 20th order filter. Filter coefficients values are taken between -1 to +1. WOA technique is used to decrease the difference of magnitude responses between designed and ideal filter. The optimal filter coefficients for which the difference is minimum are selected to design BSF, specified in Table 2. Magnitude response using these coefficients are plotted in Fig 4.

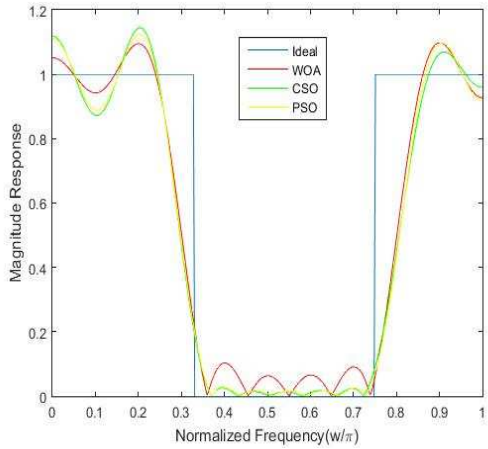


Fig. 4. Magnitude response of BPF of 20th order

V. Analysis of Simulation Results

TABLE 3. COMPARISON OF VARIOUS ALGORITHM IN MAXIMUM PASS BAND RIPPLE (MPR) FOR FILTER ORDER 20

TYPE	MPR (WOA)	MPR (CSO) [8]	MPR (PSO) [8]
LPF	0.0940	0.1320	0.1220
HPF	0.1210	0.1630	0.1460
BPF	0.0980	0.1440	0.1250
BSF	0.0900	0.1640	0.1230

TABLE 4. COMPARISON OF VARIOUS ALGORITHM IN TERMS PASS BAND ATTENUATION FOR FILTER ORDER 20

FILTER TYPE	Pass band attenuation (dB) WOA	Pass band attenuation (dB) CSO [8]	Pass band attenuation (dB) PSO [8]
LPF	20.91	15.70	18.20
HPF	20.53	17.58	18.27
BPF	18.34	15.75	16.71
BSF	20.17	16.83	18.06

It is predicted from Table 3 and 4, that ripple fluctuation is least in pass band in case of FIR LPF, HPF, BPS and BSF designed using proposed algorithm. Pass band attenuation is also higher for the designed filter using proposed algorithm. It can be said that WOA based filter is effective in stopping the unwanted signal in stop band. WOA based filter approaches ideal frequency response of FIR digital in an effective way.

TABLE 5. COMPARISON OF VARIOUS ALGORITHM IN MAXIMUM STOP BAND RIPPLE (MSR) FOR FILTER ORDER 20

TYPE	MPR (WOA)	MPR (CSO) [8]	MPR (PSO) [8]
LPF	0.0110	0.0200	0.0396
HPF	0.0179	0.0208	0.0393
BPF	0.0143	0.0189	0.0250
BSF	0.0131	0.0248	0.0296

TABLE 6. COMPARISON OF VARIOUS ALGORITHM IN TERMS STOP BAND ATTENUATION FOR FILTER ORDER 20

FILTER TYPE	Pass band attenuation (dB) WOA	Pass band attenuation (dB) CSO [8]	Pass band attenuation (dB) PSO [8]
LPF	34.81	33.99	28.03
HPF	34.12	33.62	28.10
BPF	34.48	34.47	32.03
BSF	33.73	32.11	30.56

It is depicted from Table 5 and 6 that magnitude of ripples in stop band is lowest with higher attenuation for FIR LP, HP, BP and BS filters designed with proposed optimization technique as compare to the filters designed using other

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:52:11 UTC from IEEE Xplore. Restrictions apply.

optimization algorithms. This characteristic is very much desirable to approach ideal frequency response.

VI. Conclusion

The proposed work aims to design 20th order digital FIR LPF, HPF, BPF and BSF with an objective to achieve minimum ripples in pass band and higher attenuation in stop band. The said filter is constructed using WOA while reducing error fitness function as much as possible to come closer to ideal frequency response of the digital filter. It is concluded that both objectives are met for the filter designed using WOA as compared to the other filter based on PSO and CSO algorithms.

References

[1] J.H. McClellan, and T.W. Parks, "A personal history of the Parks-McClellan algorithm," *IEEE signal processing magazine*, vol. 22, no. 2, pp.82-86, 2005.

[2] K. Boudjelaba, F. Ros, and D. Chikouche, "An efficient hybrid genetic algorithm to design finite impulse response filters," *Expert Systems with Applications*, vol. 41, no. 13, pp. 5917-5937, 2014.

[3] K. S. Reddy, and S. K. Sahoo, "An approach for FIR filter coefficient optimization using differential evolution algorithm," *AEU-International Journal of Electronics and Communications*, vol. 69, no. 1, pp. 101-108, 2015.

[4] M. Kumar, "Fractional Order FIR Differentiator Design Using Particle Swarm Optimization Algorithm," *International Journal of Numerical Modelling: Electronic Networks, Devices and Fields*, vol. 32, no. 2, 2019.

[5] I. Sharma, B. Kuldeep, A. Kumar, and V. K. Singh, "Performance of swarm-based optimization techniques for designing digital FIR filter: A comparative study," *Engineering science and technology an international journal*, vol. 19, no. 3, pp. 1564-1572, 2016.

[6] N. Karaboga, "A new design method based on artificial bee colony algorithm for digital IIR filters" *Journal of the Franklin Institute*, vol. 346, no. 4, pp. 328-348, 2009.

[7] S. Yadav, M. Kumar, R. Yadav, and A. Kumar, "A Novel Approach for Optimal Digital FIRFilter Design Using Hybrid Grey Wolf and Cuckoo Search Optimization," *Springer Scienceand Business Media LLC*, 2020.

[8] S. K. Saha, S. P. Ghoshal, R. Kar, and D. Mandal, "Cat swarm optimization algorithm for optimal linear phase FIR filter design," *ISA Transactions*, vol. *52, no.* 6, pp. 781-794, 2013.

[9] A. Goyanka, B. Rachuri, T. K. Rawat and R. Barsainya, "Bat algorithm for the design of fractional order FIR differentiators and fractional FIR Hilbert transformers: A comparative study," *In2017 8th international conference on computing, communication and networking technologies (ICCCNT)*, pp. 1-7, IEEE, 2017.

[10] S. Singh, A. Ashok, T. R. Rawat, and M. Kumar, "Optimal IIR System Identification Using Flower Pollination Algorithm," *In: IEEE First International Conference on Power Electronics, Intelligent Control and Energy Systems,* pp. 1-6, 2016.

[11] A. Kumar, I. Sharma, and L. K. Balyan. "Design of low power multiplierless FIR filter with enhanced adder efficiency using flower pollination optimization," *Applied Acoustics* vol. 174:107792, 2021.

[12] S. Yadav, M. Kumar, R. Yadav and A. Kumar, "A Novel approach for Optimal Design of Digital FIR Filter Using Grasshopper Optimization Algorithm," *Intl. J. of Wavelets, Multiresolution and Information Processing (IJWMIP)*, 2 Jul, 2021.

[13] D. Tanay, R. M. Aich, S. Dhabal, and P. Venkateswaran, "Finite impulse response filter design using grasshopper optimization algorithm and implementation on FPGA," *In 2020 IEEE Applied Signal Processing Conference (ASPCON)*, pp. 313-317, 2020.

[14] S. Yadav, M. Kumar, R. Yadav and A. Kumar, "Design of Optimal two-dimensional FIR Filters with Quadrantally Symmetric Properties Using Vortex Search Algorithm," Journal of Circuits, Systems, and Computers Vol. 29, no. 8, 20202050155 (23 pages).

[15] S. Yadav, M. Kumar, R. Yadav and A. Kumar,"A Novel Approach to Design Optimal 2-DDigital Differentiator Using Vortex Search Optimization Algorithm," *Journal of Multidimensional Systems and Signal Processing*, Feb 2020.

[16] S. Srivastava, A. K. Dwivedi, and D. Nagaria, "Low Complexity FIR Filter Design using Biogeography Optimization Algorithm and its Improved Version," *In 2020 IEEE Students Conference on Engineering & Systems (SCES)*, pp. 1-5, 2020.

[17] John, Tintu Mary, and Shanty Chacko, "Efficient VLSI architecture for FIR filter design using modified differential evolution ant colony optimization algorithm," *Circuit World*, 2020.

[18] S. Mirjalili, "Whale Optimization Algorithm", *Advance in Software Engineering*, vol. 95, pp. 51- 67, May 2016.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:52:11 UTC from IEEE Xplore. Restrictions apply.

2021 IEEE 8th Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON) | 978-1-6654-0962-9/21/$31.00 ©2021 IEEE | DOI: 10.1109/UPCON52273.2021.9667572

A review on Finger vein based Recognition

Ajai Kumar Gautam
Delhi Technological University
India
ajai.gautam@gmail.com

Rajiv Kapoor
Delhi Technological University
India
rajivkapoor@dce.ac.in

***Abstract* -- One of the most well-researched methods in biometric is finger vein-based identification for authenticating an individual. As this biometrics involves inbuilt security, such as the liveliness of the individual concerned, it cannot be replicated or altered. Further, it can be combined easily with fingerprint, palm print, palm & dorsal hand vein, and other biometrics to improve recognition accuracy as some extremely secure applications require. This paper focuses on the research works and their analysis in recognizing individuals using palm and finger veins.**

Keywords -- Biometrics, Finger Vein Recognition, Palm Vein Recognition.

I. INTRODUCTION

Finger vein or Palm Vein based recognition techniques are used widely over many applications. A Finger Vein-based authentication system has an IR camera. It captures the images of the finger vein of a person. From the captured image, either the whole picture is used for extracting features or vein structures are taken out, and then features are extracted from the vein structure or vein points

These extracted features are then compared to the characteristics features of the individual that have already been saved in the database. These features include both global as well as local. Then, with the help of the classifiers algorithm, the probe or the test vein image is classified to either one of the existing images in the database or declared as not found in the current database.

Therefore, an automatic vein recognition system includes an IR camera and processing elements, i.e., CPU and memory, to store the vein database. Fig.1 depicts various steps generally followed in recognizing a person using finger veins.

II. LITERATURE SURVEY

Vein-based Recognition has been the recent and most worked out research area relating to biometric Recognition. Today, it has many challenges ranging from designing appropriate sensors, external environmental factors such as temperature, humidity, misplacement or misalignment of the finger, physiological changes, occlusions, illumination variance, shading, and many more. Still, it also has many inherent advantages, such as it cannot be forged, only a living person's vein pattern can be extracted, which enhances its security. Its capturing environment is amiable.

The following are typical steps in a person identification using finger veins are shown in fig 1.

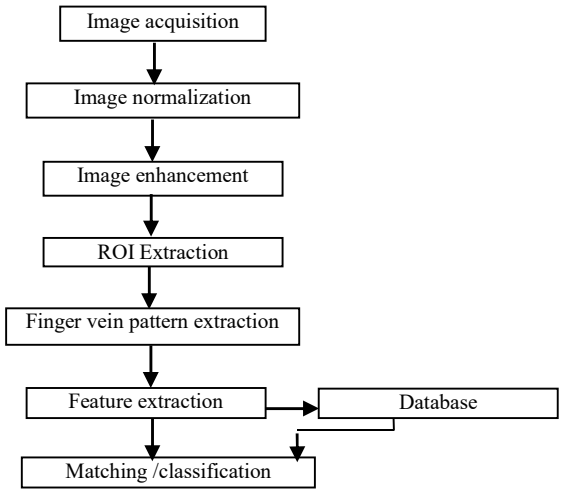


Fig. 1. It shows the typical steps generally followed in a finger vein recognition system.

Person recognition using finger veins can be grouped into different categories, namely-

Whole finger vein pattern-based, also referred to as non-vein region-based.

Finger vein pattern-based- A vein point is identified using these techniques by detecting an association between the present location and its surroundings pixels –These methods can be categorized further into – point pattern-based (including both finger vein region and non- vein region

In [1]author has used distance and spatial distribution for improving recognition accuracy. The Author uses two-step processes. In the first step k-nearest neighbors method, the initial probe is taken using the nearest centroid neighbors. The nearest centroid is scattered over the investigation and as close to the probe as possible.

In the next step, the probe is grouped using the sparse representation technique on selected k-nearest centroid neighbors. The Author reported an improvement in terms of recognition accuracy and processing time. These results are obtained as the training sample space for SRC-based grouping is minimized by using the nearest centroid neighbors,

However, the Author uses some manual cropping for the finger vein database. The author uses the "Hong Kong

978-1-6654-0962-9/21/$31.00 ©2021 IEEE.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:48:06 UTC from IEEE Xplore. Restrictions apply.

Polytechnic University finger image" (HKPU) database. One can replace manual cropping with some automated software algorithm.

In [2], the Author proposes how to generate a template with the least intra-class variation (distance). The Author presents the weighted LSR (least Square Regression) based model. This model will not only generate but also improve the template to be enrolled. This process of generation of the template is converted into an optimization problem. Once the template generates an improvement model is used to improve the template vein features. Authors claim that it's the first kind of work

The Author merges the multiple templates of the same class to create a template termed the super template. This super template is updated with the vein features from the vein images acquired during user verification. The main idea behind the improvement of the vein template is that the vein pattern changes when captured at different intervals of time.

In[3], the author proposes that finger tissue intensity distribution can be used along with core biometric and termed it soft biometric. The finger image is usually separated into two foreground & background layers. The foreground layer holds texture details. The other layer has intensity details, and this detail is used as a secondary trait for identification.

The author uses two different algorithms for separating foreground and background layers; one is the existing image layer separation algorithm based on relative smoothness as in [4], while the other is the Gaussian blur. The author also states that Gaussian blur is more efficient in terms of layer separation, noise suppression, computing time and highlights intensity distribution in the background layer.

The author extracted various soft features and found that the array of means & variance gives the best performance when combined with the texture as the primary trait of finger vein. Here a question arises can we use vein structures as a primary trait rather than using texture

The author also proposes a rotational correction by a gradient map. The authors also mention the following soft biometric traits that are already reported –

a. finger shape,
b. width of phalangeal joint
c. distance between finger edge and image centroid.

In[5], the researchers use vein and non-vein points in the finger vein image to identify objectives. Here, the author categorized image points into two groups: vein points (VP) and the 2nd group consists of points other than the vein, termed as non-vein points (NVP). These VP's or NVP's are arranged in the order of their feature value in their respective groups. The grouping threshold is calculated and is used to make groups of NP or NVP

Matching between the test image and the database is done by classifying the test image's points (VP and NVP). Then with the help of point location and point orientation, these points are arranged as matched points and non-matched points with the database.

Authors use this point grouping method in two existing vein-based recognition methods. One vein extraction using the Anatomical Structure Analysis based Vein Extraction (ASAVE) [6], other is a multi-orientation Gabor filter [7].

Overlapping points are used to calculate similarity scores, and non-overlapping points are used for dissimilarity scores. How to use this method of point grouping in other approaches that uses image points can be explored?

The feature values can extract venous and non-venous points and group all the image points in a vein-point approach. However, grouping points via attribute values by all image point approaches like Local Binary Pattern (LBP), etc., still needs more analysis.

In [8], the authors use the Pulse Coupled Neural Network (PCNN) for enhancing venous images. The improved images of the finger vein maximize the accuracy of identification.

This PCNN model has reduced computational complexity and higher efficiency. Further, the model parameters are calculated automatically without any training or empirical setting. One can also explore whether this model can be applied for segmenting finger-vein images.

In [9], the Authors proposed multimodal finger biometrics and used the graph-based method for the feature extraction from three separate modalities: finger vein, fingerprint, and finger knuckle print. The author also proposed an imaging system that captures these three modalities. The author uses a model for graph structure with weights that extract features. The basic assumption over the graph is that the local region has a high correlation with adjacent or surrounding areas than the far distant regions.

The graph is written like G (V, E, W), in which they have labeled V: Virtex, E: edge, and W: weight associated with the edge. For developing the graph, the image is divided into small blocks. These blocks are used in feature extraction. These features are represented as vertices in the graph. Once the features are extracted, a feature-level fusion is performed.

The authors use two types of fusion – one is serial fusion in this fusion weighted graph of each modality which can be represented as the compressed sparse row is combined to form an overall feature vector. Then Softmax, Support Vector Machine (SVM), and Extreme Learning Machine (ELM) classifiers are selected. A performance comparison is made for all three classifiers.

The other fusion method is coding fusion which combines the earlier generated three feature matrices (also called Adjacency matrices) using competitive coding. Then either a classifier or a matrix matching algorithm can be applied for recognition purposes. Obtaining graphical features of the finger tri-modal images addresses scale discrepancy, which is crucial for efficient and robust fusion. Further scope for improving Graph feature accuracy by using an image repainting model is left for future work.

Graph Convolutional Neural Network (GCNNs) have better outcomes in processing graph structure data, and as a result, the Author uses GCNNs to recognize finger tri-modal images based on graph features.

A deep fusion model for finger tri modalities based on GCNNs is to be developed for further work.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:48:06 UTC from IEEE Xplore. Restrictions apply.

In [10]authors proposed a pose correction method for reducing the finger pose variation in tri-modal finger images. The authors also proposed a method for reducing the effects of illumination variation on the finger images.
The pose angle can be measured using the finger's edge line , and hence it is corrected by rotation of the finger image. The image finger is rotated about the center of the finger direction line.
After the pose is corrected, the side background is eliminated using the Region of Interest (ROI) extraction technique. Three different ROI extraction techniques are used for three different modalities, namely – the core detection method for fingerprint [11], convex direction coding method for Finger Knuckle Point [12], the interphalangeal joint prior procedure applied on finger vein image [13].
Following the extraction of the ROI, the image is enhanced with k-channel, even symmetrically oriented Gabor filters. Output image of Gabor filter contains local positions and gradient features. A Gabor generalized method with symmetric local graph structure (G-SLGS) [10]is proposed to consider the target pixel's relationship with its flexible surrounding. Finally, these GSLGS are coded. Coding is done in two steps 1.-The center pixel, and its neighborhood pixel is compared in succession. If the value is greater than 1, it is assigned binary 1; otherwise, 0.
2.- Maximum value of all the coded values is selected as the target pixel value.
A block-level gray histogram is plotted for each of the modalities. These block-level histograms for each modality are fused separately in series. Finally, all three gray-level histograms are merged to generate one histogram for each finger tri-modal image. These histograms can be classified using SVM, ELM, and k-NN. The author uses a similarity score for matching the two histogram features. The similarity score is calculated by intersecting the two feature vectors.
The author indicated that the dimensionality of the feature vector could be reduced and proposed a local coding algorithm that can be applied to other biometric modalities. Other fusion techniques can also be explored.
In [14], the authors use multimodal biometric finger vein and palm vein and uses score level fusion. Fusion is completed using fuzzy logic. In this article, enhanced palm vein image is obtained after applying Contrast Limited Adaptive Histogram Equalization (CLAHE), while the enhancement of image containing finger vein pattern is performed by adaptive histogram equalization(AHE), CLAHE is generally suited where overall improvement of contrast is required.
For feature extraction, a modified 2-D Gabor filter is applied in hand vein images. Gabor filters are unaffected by changes in light, rotation, translation, and scale.
A smoothness parameter η is introduced in the Gabor filter definition. The extraction of features from palm veins is done using a gradient-based feature detection method.
Firstly the gradient magnitude and gradient orientation are calculated at each pixel level. If both exceed the threshold, then it is considered as part of the feature vector.
For matching, individual finger and palm images are compared using the Euclidean distance metric. This metric is then passed as input to the fuzzifier. The fuzzifier converts it into a matching score –as low, medium, and high. If the matching score from the fuzzifier is high for both modalities, then only a person is said to be recognized.
Future work consists of combining the attributes of the finger and palm veins at various levels such as feature level stage another level could be at the decision level and then analyzing the results; the proposed work may be extended. Other distance measurements, such as city block and Hamming distance, and Hausdorff can also be used to compare vein features.

In [15], these authors define a local learning method ~~that~~ preserves manifold structures of pixel difference value (PDV). Specifically, an asymmetric graph is constructed in such a way so that every point is made up of Its K-nearest neighbor anchors that are combined in a linear way. Finally, using unsupervised learning techniques, a feature map is obtained that maps these PDV's into binary codes. These binary codes are of low dimensionality. Here Author also combines the two existing methods Discriminative Binary Descriptor and Anchor Based Manifold Binary Pattern (AMBP) [15] at feature level (image representation level) that improves the recognition accuracy. In this paper, future work will be the application of this algorithm other biometrics like palm vein-based recognition.

In [16], this Author proposes a bilayer restoration method that tries to sort out skin scattering in finger images, including dorsal images, thus enhances the overall images that have vein patterns, but it has noisy data background.

In [17]Author uses two-finger vein images which they called it as composite image and use it for authentication. A 3-channel composite image is formed. This composite image is used on the input side of the Convolutional Neural Network (CNN). Authors claim that this method does not take the difference image as the input, therefore less prone to noise. Also, this method uses all the layers of CNN that improve accuracy. The authors use a deep, densely connected convolution network which is termed as DenseNet. One of the reasons that CNN is used in Recognition is that the preprocessing technique does not affect the CNN results. The author also uses a shift matching technique to overcome the effects of misalignment of finger vein images during enrollment and test probe images.
Here authors use the whole finger image after extracting from ROI as the input for the CNN. Also, they use the practical scenario, i.e., classes used in training are different from testing or probe image. However, the authors do not use the thick part (leftmost) and the nail part (rightmost) of the finger vein image owing to poor illumination.
Further, the author also does not use the bright light areas of the image as obtaining a vein pattern in that area is difficult.

Further, it is possible to increase processing speed while maintaining accuracy. DenseNet can be made more efficient By minimizing the number of layers and transition layers; also, other modalities related to veins may be analyzed from shift matching.
In [18] author uses block uniform LBP and block 2-directional, 2-D Principal Component Analysis (PCA) method for vein pattern recognition. In this method, the author is able to reduce data redundancy with improved performance.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:48:06 UTC from IEEE Xplore. Restrictions apply.

TABLE I: EXISTING DIFFERENT APPROACHES (THAT USES THE WHOLE FINGER, ALSO REFERRED TO AS NON-VEIN PATTERN-BASED METHODS)

LBP Based	**Minutiae & Key point of SIFT based**	**PCA based**	**High-level features based**
Poor discrimination ability Low recognition rate	Finger vein images have a small number of minutiae. Low recognition rate	In these methods, a larger number of images are used to obtain a transformation matrix through training, and this is normally not possible in real-world applications. .	recognition process is time-consuming It gives a high recognition rate with less consistency
Various Local feature extraction methods are:- LBP LLBP ELBP DBC BMULBP Global Feature extraction methods PCA Super pixel-based feature(Superpixel histogram	finger displacement and vein deformation are major issues		

In vein pattern-based method approaches through observing the interaction between the present point and its neighboring points, these methods were able to locate the vein point. The selection of the nearby points in the vein point identification will be used to classify these approaches.

In [19] author proposes a method that tries to overcome the effects of breaks in the vein pattern. A Direction-Variance-Boundary Constraint Search model is used for restoring gaps in vein pattern drawback of this method is that it consumes a lot of time.

Using the intrinsic finger vein recognition performance is not satisfactory because of defective vein network extraction and weak matching. A thorough examination and analysis of the vein's anatomical structure are required.

TABLE II: VEIN PATTERN-BASED METHODS APPROACHES

Cross-sectional profile	**Based on Neighborhood area:-**	**Based on full image:-**
The vein of the cross-sectional profile is typically less grey than the others. The profile can be represented as a valley in shape. <u>earlier approaches</u> **1. repeated line tracking** The depth of the valley is used to trace the vein point after it has been found .Finally, Tracking times are used to label the vein pattern. (in binary). **2. region growth** In vein point identification, the grey value is used to consider the valley's symmetry. **3. maximum curvature point** To find the vein pattern's centerline, the cross-sectional profile's local maximum curvature is determined. This process is carried out in four different directions. The resultant central lines are then merged by optimizing each pixel's curvature value. **4. Wide line detector(WLD):** The grey value is taken from a circular neighborhood area in this case.	Wide line detectors (WLD) along with Gabor are used To characterize the current point, and the neighborhood has a predefined radius. The current pixel is classified as a vein point if the number of neighbors with higher pixel values exceeds the predefined threshold. One of the papers constructs a series of self-similar Gabor templates. These templates are used to obtain the vein pattern by performing convolution of templates with finger vein images.	The valley-like vein pattern was observed using this approach, which represented the finger vein image as a 3-D geometrical structure and used the negative mean curvature to detect it.
All of these approaches use template matching to match the number of matched vein points.		

TABLE III: GABOR FILTER IS GENERALLY USED IN VARIOUS BIOMETRIC MODALITIES AND IS GENERALLY USED FOR OBTAINING TEXTURE AND ORIENTATION INFORMATION CONTAINED IN THE IMAGE.

Methods	**Biometric modalities**	**Pros/cons/works**
Gabor wavelets [20]	Palm vein features, finger Vein Features	Insensitive to rotation and shift variations, variation of intensity can be handled effectively
homomorphism [21] filter and a Gabor filter	Finger Vein	It increases both the illumination and the reflectance of an image at the same time. Therefore homomorphic filter is applied for enhancing image contrast.
Gabor filtering[22]	Finger Vein	Image enhancement is performed, and thus the issue with the low contrast vein image is rectified with this algorithm. The author has solved the issue normally associated with histogram(HE) equalization that is losing the details of HE algorithm,

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:48:06 UTC from IEEE Xplore. Restrictions apply.

TABLE IV: METHODS WHICH APPLIES CNN / DEEP NEURAL NETWORK IN RECOGNITION USING FINGER VEIN PATTERN

Methods	Pros/cons/applications
Radon Transform is used for extracting features while An artificial neural network is used for classification [23]	Used in driver recognition. Radon transform used to capture information available in limited high-valued coefficients
To isolate features, a CNN with four layers is used, and Euclidean distance is used to find the nearest match. [24]	Advantage - simultaneously extraction of features is possible, It reduces data dimensionality, and classification is performed using one network structure. It is also robust to noise and small misalignment
light convolutional neural networks [25]	It uses the center loss function. It also employs dynamic regularization in CNN
Lightweight CNN [26]	Addresses Variations in finger thickness and vein pattern induced by axial finger rotation that degrades image quality. However, accuracy needs further improvement
Merged CNN[27]	Uses many similar CNN with shorter paths. Similar CNN is inputted with different quality images. However, lower-quality images affect the accuracy
Two-stream convolutional neural networks[28]	There are only three convolution layers in its two-channel network. It focuses on the displacement problem. Mini ROI is identified and finally merged with the original image, which yields better results
CNN with Supervised Discrete Hashing[29]	It enhances performance also it reduces the template size requirement, but still, accuracy remains the foremost issue, and it requires a two-step training phase
Uses CNN & recurrent neural network (RNN)[30]	Constraint-free finger vein recognition. It uses an array of cameras to capture the image. Image restoration is used. RNN is used to handle multiple frames. a long Short term memory (LSTM) network is used to process Hand movement in image frames. Exposure time affects recognition accuracy

In this [31] author has discussed various loss functions that are used in the CNN and how these loss functions improve recognition accuracy by improving the variation between inter-class subjects and decreasing in the case of intraclass. This has been explored on the hand, palm, and finger vein modalities. The author also emphasizes that if we initialize CNN with the weights that have been obtained by training, using a large dataset will improve accuracy.

III. CONCLUSION

In this review on finger vein /palm vein-based Recognition, it has been found that the current biometric Recognition is focused on finger/palm vein biometrics. Combining more than one modalities of an individual around the finger, like palm & finger vein, dorsal vein, and fingerprint, will definitely increase recognition performance—also, the computation cost increases. Further, these modalities can be captured at the same time or using some of the soft biometrics along with core biometrics. Computation cost can be optimized without compromising on accuracy. In fact, we have seen in some cases authors try to cut down the processing time by reducing the number of layers in CNN. In some cases, the author uses more than one finger image to improve the accuracy. Also, a new direction has been observed where users can be recognized on the fly with no constraint.

REFERENCES

[1] S. Shazeeda and B. A. Rosdi, "Nearest Centroid Neighbor Based Sparse Representation Classification for Finger Vein Recognition," *IEEE Access*, vol. 7, pp. 5874–5885, 2019, doi: 10.1109/ACCESS.2018.2889506.

[2] H. Qin and P. Wang, "A template generation and improvement approach for finger-vein recognition," *Inf.*, vol. 10, no. 4, pp. 1–19, 2019, doi: 10.3390/info10040145.

[3] W. Kang, Y. Lu, D. Li, and W. Jia, "From noise to feature: Exploiting intensity distribution as a novel soft biometric trait for finger vein recognition," *IEEE Trans. Inf. Forensics Secur.*, vol. 14, no. 4, pp. 858–869, 2019, doi: 10.1109/TIFS.2018.2866330.

[4] Y. Li and M. S. Brown, "Single image layer separation using relative smoothness," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 2752–2759, 2014, doi: 10.1109/CVPR.2014.346.

[5] L. Yang, G. Yang, K. Wang, H. Liu, X. Xi, and Y. Yin, "Point grouping method for finger vein recognition," *IEEE Access*, vol. 7, pp. 28185–28195, 2019, doi: 10.1109/ACCESS.2019.2901017.

[6] L. Yang, G. Yang, Y. Yin, and X. Xi, "Finger Vein Recognition with Anatomy Structure Analysis," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 28, no. 8, pp. 1892–1905, 2018, doi: 10.1109/TCSVT.2017.2684833.

[7] A. Kumar and Y. Zhou, "Human identification using finger images," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 2228–2244, 2012, doi: 10.1109/TIP.2011.2171697.

[8] L. Lei, F. Xi, and S. Chen, "Finger-Vein Image Enhancement Based on Pulse Coupled Neural Network," *IEEE Access*, vol. 7, pp. 57226–57237, 2019, doi: 10.1109/ACCESS.2019.2914229.

[9] H. Zhang, S. Li, Y. Shi, and J. Yang, "Graph fusion for finger multimodal biometrics," *IEEE Access*, vol. 7, pp. 28607–28615, 2019, doi: 10.1109/ACCESS.2019.2902133.

[10] S. Li, H. Zhang, Y. Shi, and J. Yang, "Novel local coding algorithm for finger multimodal feature description and recognition," *Sensors (Switzerland)*, vol. 19, no. 9, 2019, doi: 10.3390/s19092213.

[11] H. B. Kekre and V. A. Bharadi, "Fingerprint's core point detection using orientation field," *ACT 2009 - Int. Conf. Adv. Comput. Control Telecommun. Technol.*, pp. 150–152, 2009, doi: 10.1109/ACT.2009.45.

[12] L. Zhang, L. Zhang, D. Zhang, and H. Zhu, "Online finger-knuckle-print verification for personal authentication," *Pattern Recognit.*, vol. 43, no. 7, pp. 2560–2571, 2010, doi: 10.1016/j.patcog.2010.01.020.

[13] J. Yang and Y. Shi, "Finger-vein ROI localization and vein ridge enhancement," *Pattern Recognit. Lett.*, vol. 33, no. 12, pp. 1569–1579, 2012, doi: 10.1016/j.patrec.2012.04.018.

[14] S. Bharathi and R. Sudhakar, "Biometric recognition using finger and palm vein images," *Soft Comput.*, vol. 23, no. 6, pp. 1843–1855, 2019, doi: 10.1007/s00500-018-3295-6.

[15] H. Liu, G. Yang, L. Yang, K. Su, and Y. Yin, "Anchor-based manifold binary pattern for finger vein recognition," *Sci. China Inf.*

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:48:06 UTC from IEEE Xplore. Restrictions apply.

Sci., vol. 62, no. 5, pp. 1–16, 2019, doi: 10.1007/s11432-018-9651-8.

[16] W. You, W. Zhou, J. Huang, F. Yang, Y. Liu, and Z. Chen, "A bilayer image restoration for finger vein recognition," *Neurocomputing*, vol. 348, pp. 54–65, 2019, doi: 10.1016/j.neucom.2018.06.085.

[17] J. M. Song, W. Kim, and K. R. Park, "Finger-Vein Recognition Based on Deep DenseNet Using Composite Image," *IEEE Access*, vol. 7, pp. 66845–66863, 2019, doi: 10.1109/ACCESS.2019.2918503.

[18] N. Hu, H. Ma, and T. Zhan, "Finger vein biometric verification using block multi-scale uniform local binary pattern features and block two-directional two-dimension principal component analysis," *Optik (Stuttg).*, vol. 208, no. September 2019, 2020, doi: 10.1016/j.ijleo.2019.163664.

[19] T. Liu, J. Xie, W. Yan, P. Li, and H. Lu, "Finger-vein pattern restoration with Direction-Variance-Boundary Constraint Search," *Eng. Appl. Artif. Intell.*, vol. 46, pp. 131–139, 2015, doi: 10.1016/j.engappai.2015.09.004.

[20] R. Wang, G. Wang, Z. Chen, Z. Zeng, and Y. Wang, "A palm vein identification system based on Gabor wavelet features," *Neural Comput. Appl.*, vol. 24, no. 1, pp. 161–168, 2014, doi: 10.1007/s00521-013-1514-8.

[21] J. Yang, J. Yang, and Y. Shi, "Finger-vein segmentation based on multi-channel even-symmetric Gabor filters," *Proc. - 2009 IEEE Int. Conf. Intell. Comput. Intell. Syst. ICIS 2009*, vol. 4, pp. 500–503, 2009, doi: 10.1109/ICICISYS.2009.5357631.

[22] X. Li, J. Ji, Y. Xu, S. He, and Q. Zhou, "Research on Low Contrast Image Enhancement Algorithm," *IEEE Adv. Inf. Technol. Electron. Autom. Control Conf.*, vol. 2021, pp. 1974–1977, 2021, doi: 10.1109/IAEAC50856.2021.9391127.

[23] J.-D. Wu and S.-H. Ye, "Driver identification using finger–vein patterns with radon transform and neural network.," *Expert Syst. Appl.*, vol. 36, no. 3, pp. 5793–5799, 2009.

[24] S. Ahmad Radzi, M. Khalil-Hani, and R. Bakhteri, "Finger-vein biometric identification using convolutional neural network," *Turkish J. Electr. Eng. Comput. Sci.*, vol. 24, no. 3, pp. 1863–1878, 2016, doi: 10.3906/elk-1311-43.

[25] W. T. Dongdong Zhao, Hui Ma, Zedong Yang, Jianian Li, "Finger vein recognition based on lightweight CNN combining center loss and dynamic regularization," *Infrared Phys. Technol.*, vol. 105, 2020.

[26] W. Yang, W. Luo, W. Kang, Z. Huang, and Q. Wu, "FVRAS-Net: An Embedded Finger-Vein Recognition and AntiSpoofing System Using a Unified CNN," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 11, pp. 8690–8701, 2020, doi: 10.1109/TIM.2020.3001410.

[27] I. Boucherit, M. O. Zmirli, H. Hentabli, and B. A. Rosdi, "Finger vein identification using deeply-fused Convolutional Neural Network," *J. King Saud Univ. - Comput. Inf. Sci.*, no. xxxx, 2020, doi: 10.1016/j.jksuci.2020.04.002.

[28] W. K. Yuxun Fang, Qiuxia Wu, "A novel finger vein verification system based on two-stream convolutional network learning," *Neurocomputing,* vol. 290, pp. 100–107, 2018, doi: https://doi.org/10.1016/j.neucom.2018.02.042.

[29] C. Xie and A. Kumar, "Finger vein identification using Convolutional Neural Network and supervised discrete hashing," *Pattern Recognit. Lett.*, vol. 119, pp. 148–156, 2019, doi: 10.1016/j.patrec.2017.12.001.

[30] R. S. Kuzu *et al.*, "On-the-Fly Finger-Vein-Based Biometric," vol. 15, pp. 2641–2654, 2020.

[31] R. S. Kuzu, E. Maiorana, and P. Campisi, "Loss functions for CNN-based biometric vein recognition," *Eur. Signal Process. Conf.*, vol. 2021-Janua, pp. 750–754, 2021, doi: 10.23919/Eusipco47968.2020.9287517.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:48:06 UTC from IEEE Xplore. Restrictions apply.

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.08.475476; this version posted January 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

A visual atlas of genes' tissue-specific pathological roles

Priyadarshini Rai[1†], Atishay Jain[2†], Neha Jha[1†], Divya Sharma[1†], Shivani Kumar[2†], Abhijit Raj[1†], Apoorva Gupta[3], Sarita Poonia[1], Smriti Chawla[1], Angshul Majumdar[2,4,5], Tanmoy Chakraborty[2,5*], Gaurav Ahuja[1*] and Debarka Sengupta[1,2,5,6*]

1. Department of Computational Biology, Indraprastha Institute of Information Technology-Delhi (IIIT-Delhi), Okhla, Phase III, New Delhi-110020, India.
2. Department of Computer Science and Engineering, Indraprastha Institute of Information Technology-Delhi (IIIT-Delhi), Okhla, Phase III, New Delhi-110020, India.
3. Department of Biotechnology, Delhi Technological University, Shahbad Daulatpur, Main Bawana Road, Delhi-110042, India.
4. Department of Electronics & Communications Engineering, Indraprastha Institute of Information Technology-Delhi (IIIT-Delhi), Okhla, Phase III, New Delhi-110020, India.
5. Centre for Artificial Intelligence, Indraprastha Institute of Information Technology-Delhi (IIIT-Delhi), Okhla, Phase III, New Delhi-110020, India.
6. Institute of Health and Biomedical Innovation, Queensland University of Technology, Brisbane, Australia.

† These authors contributed equally
* For correspondence: debarka@iiitd.ac.in, gaurav.ahuja@iiitd.ac.in, tanmoy@iitd.ac.in

Keywords: text mining, Word2vec, gene, disease, organ, tissue

Key points:

- We mined ~18 million PubMed abstracts to extract latent knowledge pertaining to tissue specific pathological roles of genes.
- We found well-defined gene modules implicated in disease pathogenesis in anatomically proximal tissues.
- We demonstrated an ahead of time discovery of the association between DNMT3A/BCOR with CYLD cutaneous syndrome (CCS), as a knowledge synthesis use-case.

Abstract

Dysregulation of a gene's function, either due to mutations or impairments in regulatory networks, often triggers pathological states in the affected tissue. Comprehensive mapping of these apparent gene-pathology relationships is an ever daunting task, primarily due to genetic pleiotropy and lack of suitable computational approaches. With the advent of high throughput genomics platforms and community scale initiatives such as the Human Cell Landscape (HCL) project [1], researchers have been able to create gene expression portraits of healthy tissues

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.08.475476; this version posted January 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

resolved at the level of single cells. However, a similar wealth of knowledge is currently not at our finger-tip when it comes to diseases. This is because the genetic manifestation of a disease is often quite heterogeneous and is confounded by several clinical and demographic covariates. To circumvent this, we mined ~18 million PubMed abstracts published till May 2019 and selected ~6.1 million of them that describe the pathological role of genes in different diseases. Further, we employed a word embedding technique from the domain of Natural Language Processing (NLP) to learn vector representation of entities such as genes, diseases, tissues, etc., in a way such that their relationship is preserved in a vector space. Notably, Pathomap, by the virtue of its underpinning theory, also learns transitive relationships. Pathomap provided a vector representation of words indicating a possible association between *DNMT3A*/*BCOR* with *CYLD* cutaneous syndrome (CCS). The first manuscript reporting this finding was not part of our training data.

Background

The number of published biomedical manuscripts has sharply increased since the beginning of the genomic era. Nearly 3000 manuscripts are published daily in peer-reviewed journals [2]. This has resulted in a voluminous corpus of scientific literature, outreaching by far, the limit of human comprehension. Given the complexity of human biological systems, it is nearly impossible to get a qualitative as well as a quantitative view of organ/tissue wide pathological roles of genes by an internet search of the literature. Today, a great deal of community level effort is channeled towards creating a molecular atlas of healthy human organs by charting out the functionally distinct cellular subtypes in tissues of interest. Such effort is scarce in disease biology due to the presence of innumerable covariates and the diversity of the disease states. A vast majority of knowledge-based archiving disease-gene association is painstakingly, manually curated. DisGeNET [3] and Online Mendelian Inheritance in Man (OMIM®) [4] are noteworthy in this regard. NLP based efforts in this regard have mainly concentrated on extraction of disease-gene association from biomedical literature which is an automated version of manual data curation [5]. While these repositories are imperative to take our disease biology investigations forward, none of these are equipped to give a comprehensive atlas-like view of genes' tissue specific pathological roles.

Here we present Pathomap, a visual atlas of genes' tissue-specific pathological roles. To achieve this, we took the help of word embedding technologies from the domain of NLP. Word embedding typically employs efficient neural networks to learn word-representations by ingesting millions of documents. Mikolov and colleagues first proposed the very idea of learning a continuous bag of representation of words at Google LLC [6]. The field has since evolved dramatically, and in recent times, transformer based models such as BERT (bidirectional encoder representations from transformers) [7] have emerged as state-of-the-art. In the past years, text mining and language modeling have gained prominence in the fields of biomedical research. Gideon et al. used word embedding techniques to identify associations among

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.08.475476; this version posted January 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

different brain regions [8]. Similar approaches have been used in predicting compound-protein interactions [9] and drug discovery [10]. BERT is highly resource intensive. In a recent effort by Lee and colleagues, embeddings have been produced for several billions of words by processing voluminous biomedical literature, using BERT [2]. We hypothesized that such generic word embeddings may not capture the essence of molecular pathology. To circumvent this, we first screened ~6.1 million abstracts that describe an association between certain genes' dysregulation and diseases from a database of ~18 million biomedical abstracts. Word embeddings were learned exclusively on the basis of these selected abstracts.

In contrast to the generic BioBERT embeddings, the Pathomap word vectors were found to preserve semantic relationships between words in a manner that is consistent with our knowledge of the pathogenesis of diverse diseases. Furthermore, these disease-focused word embeddings allowed us to extract latent knowledge beyond the ambit of the training literature corpus.

Materials and Methods

Data Description. We downloaded ~18 million abstracts [11] from PubMed that were published till May 2019. We postulated that abstracts capture the essence of the entire manuscript [28]. Also, due to paywalls, many articles are not accessible for mining.

Manual annotation of abstracts. We looked for articles in the field of molecular biology and genetics, and thereby manually curated two sets of abstracts - first, describing genes' direct roles in disease pathogenesis, and second, abstracts that do not cite apparent disease-gene relationships. The former consisted of 1,363 abstracts (**Supplementary File 1**) and the latter consisted of 768 abstracts (**Supplementary File 2**). The abstracts with information about gene-disease relationships were labeled as 'relevant' or 'positive' data and the remaining abstracts that did not mention gene and corresponding disease relationships were labeled as 'non-relevant' or 'negative' data. We ensured that the number of positively and negatively annotated abstracts is approximately the same in number. For the annotation of abstracts, we also referred to articles from GWAS (genome-wide association study), COSMIC(Catalogue Of Somatic Mutations In Cancer), and OMIM (Online Mendelian Inheritance in Man) databases.

Stratifying pathological abstracts from non-pathological abstracts. An abstract stratification strategy was developed to shortlist the abstracts that depict the pathological role of genes. We used positive and negative sets of abstracts as seed data. These abstracts were further divided into train and test sets in a 75:25 split. We used BioSentVec [12] to transform these abstracts into vectors of dimension 700 that are suitable for machine learning. We evaluated the performance of three classifiers, namely, Support Vector Machine (SVM) [13], Extreme Gradient Boosting (XGBoost) [14], and Logistic Regression (LR) [15] on the associated binary classification task. SVM outperformed the other two methods in terms of prediction accuracy (**Supplementary Table S1, Supplementary Table S2, Supplementary Figure S1**). We classified all ~18 million

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.08.475476; this version posted January 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

abstracts using the SVM model (https://github.com/Priyadarshini-Rai/Pathomap/tree/main/Pathomap/inst/extdata). As a stringent selection criterion, we qualified an abstract as *patho-abstract* if the SVM model returned prediction probability was greater than 0.7. After classification, we obtained ~6.1 million abstracts as *patho-abstract.*

Word2vec methodology. The ~6.1 million abstracts obtained after classification were subjected to train the Word2vec skip-gram model. We performed pre-processing by first performing tokenization using python's nltk's word tokenizer. After that, we removed punctuations and stop-words from the list of tokens. We found that lower-casing the articles helped reduce the computational overhead and vocabulary size without significantly affecting the performance. We finally used the *patho-abstracts* to train our primary Word2vec model. Phrases provided by 'gensim' in python were used to handle bigrams. Window size and vector length were considered to be 5, and 300 respectively. Word embeddings trained on *patho-abstracts* are shared on Zenodo (https://zenodo.org/record/5237312). A tissue specific *Patho-score* is calculated simply as the cosine distance between the vectors associated with a tissue name and a certain gene symbol. **Figure 1A** outlines the method workflow.

Pathomap for visualizing the impact of a gene on different tissues. For wide dissemination and ease of use, we created a Shiny application (https://the-ahuja-lab.shinyapps.io/Pathomap), which is freely available to users. An R package is also compiled to give users programmatic access to the embeddings. Both these mediums allow users to visualize a gene's tissue specific pathogenic roles and compare them with tissue-wide expression levels, obtained from GTEx gene expression archives [16]. We will update the Shiny application and R package yearly to incorporate new discoveries. Of note, we used the gganatogram R package to visualize *Patho-scores* on a human body template [17]. For visualizing gene expression levels, we applied median normalization and log transformation to the GTEx expression matrix.

Insights and Applications

Semantic similarity among word-relationships. To evaluate the meaningfulness of the word vectors, we constructed a case-study to visualize the semantic similarity between disease-symptom-gene relationships. We considered three such triplets — schizophrenia-hallucination-ZNF804A [18]; epilepsy-seizure-TCF4 [19]; migraine-headache-ATP1A2 [20]. We visualized vectors associated with these words using Principal Component Analysis (PCA) (**Figure 1B**). The relative positioning of the words across the triplets exhibits consistent directionalities, such that there exist consistent vector operations between words that represent concepts such as 'symptom of ' and 'causative gene for'. Vectors associated with about 2 million words, trained on *patho-abstracts* are made available on Zenodo (https://doi.org/10.5281/zenodo.5237312). These can be used as an exploratory tool to visualize, discover, and confirm more such relationships.

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.08.475476; this version posted January 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Disease associated genes are shared among anatomically proximal tissues. Diseases are best understood in terms of the dysregulation of genes and other regulatory components. It's not convenient to explore if genes share tissue specific disease associations. To validate this, we plotted a heatmap depicting 1000 genes with the highest standard deviation of *Patho-score* across 34 main tissue types (**Figure 2**). We see the emergence of neat gene modules with reproducible intensity patterns of *Patho-scores* across diverse tissues. For example, all important brain regions (hypothalamus, cerebellum, frontal cortex, amygdala, and hippocampus) share the disease causing genes. Similarly, skeletal muscle and adipose tissue are found to share the disease associated genes. Notably, metabolic interrelationships and cross-talk of signals derived from both skeletal muscle and adipose tissue have been long investigated [21]. The prostate and the breast are two organs with several similarities: lobular glands, gonadal steroid sensitivity, and diseases treated by some comparable therapeutic approaches, primarily in the context of cancer. In fact, pathological evaluation of these two organs is similar in nature [22]. Both these organs shared similar patterns associated with *Patho-scores*. We failed to see such meaningful structures when we attempted to reproduce this heatmap using pre-trained BioBERT embeddings (**Supplementary Figure S2**). This highlights the power of context specific learning of word-vectors.

Disease linked genes are uniformly distributed across the human genome. We speculated that there could be genomic hotspots of disease linked genes. To test this, we divided the length of each of the human chromosomes into contiguous chunks spanning 1 million base pairs and counted the total number of genes located in each of these DNA stretches. Further, we counted the number of times gene(s) located in each of these stretches appeared in the *patho-abstracts*. We found these values to be highly correlated across the 1 million base-pair long DNA stretches (**Supplementary Figure S3**) with a Spearman's correlation coefficient of 0.7.

Visual atlas of disease linked genes. Since the beginning of the genomic era, several studies have been conducted to create molecular portraits of healthy tissues using transcriptomic, proteomic, and epigenomic assays. Single cell transcriptomics has fueled such efforts by providing unprecedented resolution into phenotypic heterogeneity in cells. With the advent of a high throughput single-cell transcriptomics platform, the past few years have seen a sharp spike in single cell atlasing effort, disentangling versatile molecular architecture of healthy tissues. It is difficult to undertake such studies for diseased tissues due to the non-deterministic nature of genetic aberrations that underpin their associated pathogenesis. Pathomap is the first unbiased, metainformation platform to map genes' pathogenic role at the tissue level. We developed an R package as well as a Shiny web server to facilitate gene queries. For example, **Figure 1C** and **Figure 1D** depict *PSEN2* expression and *Patho-score* respectively across human tissues. Mutations in the *PSEN2* gene are known to have a causal role in Alzheimer's disease [23]. When it comes to gene expression, *PSEN2* is minimally variable across tissues. Several, well-known disease linked genes were tested to ensure the integrity of the platform (**Supplementary Figures S4-19**)

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.08.475476; this version posted January 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Word analogy and knowledge synthesis. Word embeddings can be used to find domain specific analogies. We tested various well-known relationships. For instance, the most similar gene to diabetes associated with *TCF7L2* [24] was found to be *CDKAL1* with a cosine similarity of 0.69. Notably, *CDKAL1* is also implicated in type 2 diabetes [25]. Second, we asked, "diabetes is to the pancreas, as asthma is to?" We obtained trachea, bronchi, lung as top hits with a cosine similarity of 0.56, 0.56, and 0.52 respectively. To the question "diabetes is to diuresis as asthma is to?", we got the answer as bronchoconstriction, cough, bronchospasm with a cosine similarity of 0.58, 0.54, and 0.53 respectively. We also tried to predict the type of cancer based on the gene — "*BRCA* is to breast cancer as *KRAS* is to which cancer?", where the most similar response given by the embedding is colorectal cancer having a cosine similarity score of 0.62 [26]. *TCF7L2* variants have been associated with type 2 diabetes in multiple ethnic groups [24]. "Diabetes is caused by aberrations in *TCF7L2* as schizophrenia is caused by which gene?" Among all genes, *ZNF804A* was returned as the top gene [18]. Such analogies can be represented through algebraic equations involving concerned word-vectors as follows.

$$TCF7L2 - diabetes + schizophrenia \sim ZNF804A$$

Word embeddings are capable of deriving latent knowledge. The latest publication in our abstract repository was dated May 2019. A paper by Davies et al., published on Oct 17, 2019, reported molecular profiling data of the genomic landscape (42 benign and malignant tumors across 13 individuals from four multigenerational families) and discovered recurrent mutations in epigenetic modifiers *DNMT3A* and *BCOR* in 29% of benign tumors [27]. To the best of our knowledge, this was the first report linking these genes with *CYLD* cutaneous syndrome (CCS). We obtained high cosine similarity between these genes and *CYLD*: ('CYLD','DNMT3A') = 0.30 ('CYLD','BCOR') = 0.37.

Benchmarking of Pathomap's automated approach against manual curation. DisGeNET features a large number of disease-gene relationships, which can be considered as ground truth data [3]. We found 9727 such relationships for which disease and gene names were present among the tokens for which word embeddings were produced. These are considered positive samples. An equal number of randomly disease-gene pairs were populated and associated *Patho-scores* were obtained. These *Patho-scores* are considered as negative samples. We hypothesized that if Pathomap's automated approach is viable, we should typically expect clearly higher *Patho-scores* for experimentally validated disease-gene relationships, logged by DisGeNET. We compared the distribution of *Patho-scores* associated with positive and negative samples (**Figure 3A**). The *Patho-scores* associated with the positive samples showed significant right skew in distribution as compared to negative samples, depicting high fidelity of the *Patho-scores* above a cut-off of 0.3. We compared the distributions using the Kruskal-Wallis test (*p-value* < 2.2e-16). Pathomap has about half of the DisGeNET annotated diseases in the database (**Figure 3B**). Further, we downloaded the list of human genes from NCBI and gene-disease associations from the DisGeNET database. The human genes list downloaded from NCBI has ~25K genes. We found ~15K of these genes in *Pathomap*, whereas ~9K genes in the DisGeNET database. ~8.8K genes were common in both databases.

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.08.475476; this version posted January 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Discussion and Conclusion

Pathomap provides a strategy to obtain an unbiased continuous representation of disease causing genes and their tissue specificity. This can accelerate targeted investigations into diverse diseases. *Patho-scores* and associated color intensities indicate the extent of a gene's involvement in causing disease in a particular tissue/organ. Pathomap does not provide specifics of the gene's involvement in a particular disease. For that one can look into PubMed. It is well known that gene expression matrices are low-rank. Our analysis of the top 1000 tissue wide differentially disease-implicated genes suggests that their pathological roles are reproducibly patterned (**Figure 2**). This is a high-level, empirical view and by design, trustworthy.

A major caveat of the current study is that we did not perform species wise segregation of the articles. As such, the current scores do not discriminate between mice, humans, or other primates. We believe this won't introduce any nuisance factors since animal models are typically used to mimic human diseases only. Pathomap can be conveniently used for deciding genes for diagnostic gene-panels. We demonstrated Pathomap's ability to synthesize previously unknown associations between diseases and genes scopes. Pathomap may also serve as an orthogonal approach for Gene Ontologies for the narrow and precise investigation of outputs from high-throughput experiments such as RNA-Seq or MicroArray. It can also assist in prioritizing the gene selection for the large-scale loss of function studies.

We compared Pathomap's automated literature mining approach to the high fidelity manual curation approach taken by DisGeNET. We found *patho-scores* to be extremely reliable beyond a moderately conservative cutoff. While human cell atlasing is a reality, a pathological atlas might take significant effort and resources, which we predict to not arrive anytime soon. Till the time this vision becomes a reality, we trust that Pathomap will assist the community in narrowing the investigation scopes.

Availability: Pathomap is freely available as both a web server and an R software package. Web Server: https://the-ahuja-lab.shinyapps.io/Pathomap/; R package: https://github.com/Priyadarshini-Rai/Pathomap/blob/main/Pathomap/README.md.

Acknowledgment

DS conceived the study. PR wrote the code and performed all analyses with assistance from AJ, SK, AR, and AG. NJ, DSH (Divya Sharma), SC, SP performed manual annotations of abstracts. NJ and DSH performed analysis on word analogies and knowledge synthesis. TC and GA supervised NLP related developments, biological interpretation alongside DS. AM assisted in improving the machine learning workflow.

References

1. Han, Xiaoping, Ziming Zhou, Lijiang Fei, Huiyu Sun, Renying Wang, Yao Chen, Haide Chen, et

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.08.475476; this version posted January 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

al. 2020. “Construction of a Human Cell Landscape at Single-Cell Level.” *Nature* 581 (7808): 303–9.
2. Lee, Jinhyuk, Wonjin Yoon, Sungdong Kim, Donghyeon Kim, Sunkyu Kim, Chan Ho So, and Jaewoo Kang. 2020. “BioBERT: A Pre-Trained Biomedical Language Representation Model for Biomedical Text Mining.” *Bioinformatics* 36 (4): 1234–40.
3. Piñero, Janet, Àlex Bravo, Núria Queralt-Rosinach, Alba Gutiérrez-Sacristán, Jordi Deu-Pons, Emilio Centeno, Javier García-García, Ferran Sanz, and Laura I. Furlong. 2017. “DisGeNET: A Comprehensive Platform Integrating Information on Human Disease-Associated Genes and Variants.” *Nucleic Acids Research* 45 (D1): D833–39.
4. McKusick, Victor A. 2007. “Mendelian Inheritance in Man and Its Online Version, OMIM.” *American Journal of Human Genetics* 80 (4): 588–604.
5. Bravo, Àlex, Janet Piñero, Núria Queralt-Rosinach, Michael Rautschka, and Laura I. Furlong. 2015. “Extraction of Relations between Genes and Diseases from Text and Large-Scale Data Analysis: Implications for Translational Research.” *BMC Bioinformatics* 16 (February): 55.
6. Mikolov, Tomas, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. “Efficient Estimation of Word Representations in Vector Space.” *arXiv [cs.CL]*. arXiv. http://arxiv.org/abs/1301.3781.
7. Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. “BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding.” *arXiv [cs.CL]*. arXiv. http://arxiv.org/abs/1810.04805.
8. Rosenthal, Gideon, František Váša, Alessandra Griffa, Patric Hagmann, Enrico Amico, Joaquín Goñi, Galia Avidan, and Olaf Sporns. 2018. “Mapping Higher-Order Relations between Brain Structure and Function with Embedded Vector Representations of Connectomes.” *Nature Communications* 9 (1): 2178.
9. Wan, Fangping, and Jianyang (michael) Zeng. n.d. “Deep Learning with Feature Embedding for Compound-Protein Interaction Prediction.” https://doi.org/10.1101/086033.
10. Agarwal, Pankaj, and David B. Searls. 2008. “Literature Mining in Support of Drug Discovery.” *Briefings in Bioinformatics* 9 (6): 479–92.
11. Kim, Donghyeon, Jinhyuk Lee, Chan Ho So, Hwisang Jeon, Minbyul Jeong, Yonghwa Choi, Wonjin Yoon, Mujeen Sung, and Jaewoo Kang. 2019. “A Neural Named Entity Recognition and Multi-Type Normalization Tool for Biomedical Text Mining.” *IEEE Access* 7: 73729–40.
12. Tshitoyan, Vahe, et al. "Unsupervised word embeddings capture latent knowledge from materials science literature." Nature 571.7763 (2019): 95-98.
13. Chen, Qingyu, Yifan Peng, and Zhiyong Lu. 2019. “BioSentVec: Creating Sentence Embeddings for Biomedical Texts.” In *2019 IEEE International Conference on Healthcare Informatics (ICHI)*, 1–5.
14. Chang, Chih-Chung, and Chih-Jen Lin. 2011. “LIBSVM: A Library for Support Vector Machines.” *ACM Trans. Intell. Syst. Technol.*, 27, 2 (3): 1–27.
15. Chen, Tianqi, and Carlos Guestrin. 2016. “XGBoost: A Scalable Tree Boosting System.” In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–94. KDD ’16. New York, NY, USA: Association for Computing Machinery.
16. Hosmer, David W., Jr, Stanley Lemeshow, and Rodney X. Sturdivant. 2013. *Applied Logistic Regression*. John Wiley & Sons.
17. Lonsdale, John, Jeffrey Thomas, Mike Salvatore, Rebecca Phillips, Edmund Lo, Saboor Shad, Richard Hasz, et al. 2013. “The Genotype-Tissue Expression (GTEx) Project.” *Nature Genetics* 45 (6): 580–85.
18. Maag, Jesper L. V. 2018. “Gganatogram: An R Package for Modular Visualisation of Anatograms and Tissues Based on ggplot2.” *F1000Research* 7 (September): 1576.
19. Harrison, Paul J. 2017. “ZNF804A : Insights From the First Genome-Wide Significant Schizophrenia Gene.” *Biological Psychiatry*. https://doi.org/10.1016/j.biopsych.2017.03.024.
20. Matricardi, Sara, Paolo Bonanni, Giulia Iapadre, Maurizio Elia, Elisabetta Cesaroni, Alberto Danieli, Susanna Negrin, et al. 2021. “Epilepsy, Electroclinical Features, and Long-Term Outcomes in Pitt-Hopkins Syndrome due to Pathogenic Variants in the TCF4 Gene.” *European Journal of Neurology: The Official Journal of the European Federation of Neurological Societies*, September. https://doi.org/10.1111/ene.15104.
21. De Fusco, Maurizio, Roberto Marconi, Laura Silvestri, Luigia Atorino, Luca Rampoldi, Letterio

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.08.475476; this version posted January 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Morgante, Andrea Ballabio, Paolo Aridon, and Giorgio Casari. 2003. "Haploinsufficiency of ATP1A2 Encoding the Na+/K+ Pump α2 Subunit Associated with Familial Hemiplegic Migraine Type 2." *Nature Genetics* 33 (2): 192–96.

22. Argilés, Josep M., Joaquín López-Soriano, Vanessa Almendro, Sílvia Busquets, and Francisco J. López-Soriano. 2005. "Cross-Talk between Skeletal Muscle and Adipose Tissue: A Link with Obesity?" *Medicinal Research Reviews* 25 (1): 49–65.
23. Algaba, Ferran, Vicente Peg, Maria Rosaria Raspollini, and Antonio Lopez-Beltran. 2019. "♂♀ Prostate and Breast Pathology: Similarities and Differences." In *Gynecologic and Urologic Pathology: Similarities, Differences and Challenges*, 155–70. Cambridge University Press.
24. Lanoiselée, Hélène-Marie, Gaël Nicolas, David Wallon, Anne Rovelet-Lecrux, Morgane Lacour, Stéphane Rousseau, Anne-Claire Richard, et al. 2017. "APP, PSEN1, and PSEN2 Mutations in Early-Onset Alzheimer Disease: A Genetic Screening Study of Familial and Sporadic Cases." *PLoS Medicine* 14 (3): e1002270.
25. Florez, Jose C., Kathleen A. Jablonski, Nick Bayley, Toni I. Pollin, Paul I. W. de Bakker, Alan R. Shuldiner, William C. Knowler, David M. Nathan, David Altshuler, and Diabetes Prevention Program Research Group. 2006. "TCF7L2 Polymorphisms and Progression to Diabetes in the Diabetes Prevention Program." *The New England Journal of Medicine* 355 (3): 241–50.
26. Steinthorsdottir, Valgerdur, Gudmar Thorleifsson, Inga Reynisdottir, Rafn Benediktsson, Thorbjorg Jonsdottir, G. Bragi Walters, Unnur Styrkarsdottir, et al. 2007. "A Variant in CDKAL1 Influences Insulin Response and Risk of Type 2 Diabetes." *Nature Genetics* 39 (6): 770–75.
27. Amado, Rafael G., Michael Wolf, Marc Peeters, Eric Van Cutsem, Salvatore Siena, Daniel J. Freeman, Todd Juan, et al. 2008. "Wild-Type KRAS Is Required for Panitumumab Efficacy in Patients with Metastatic Colorectal Cancer." *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology* 26 (10): 1626–34.
28. Davies, Helen R., Kirsty Hodgson, Edward Schwalbe, Jonathan Coxhead, Naomi Sinclair, Xueqing Zou, Simon Cockell, Akhtar Husain, Serena Nik-Zainal, and Neil Rajan. 2019. "Epigenetic Modifiers DNMT3A and BCOR Are Recurrently Mutated in CYLD Cutaneous Syndrome." *Nature Communications* 10 (1): 4717.

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.08.475476; this version posted January 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

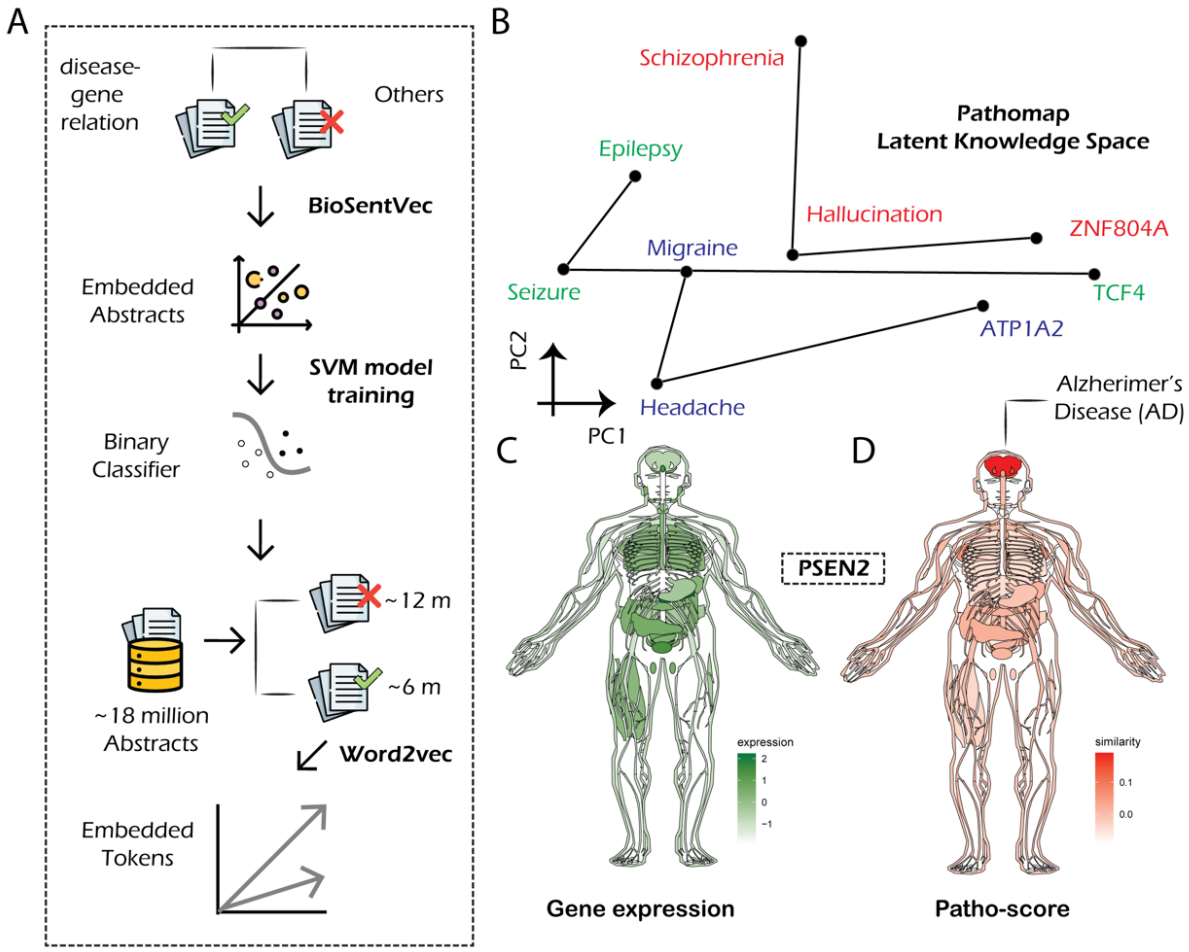
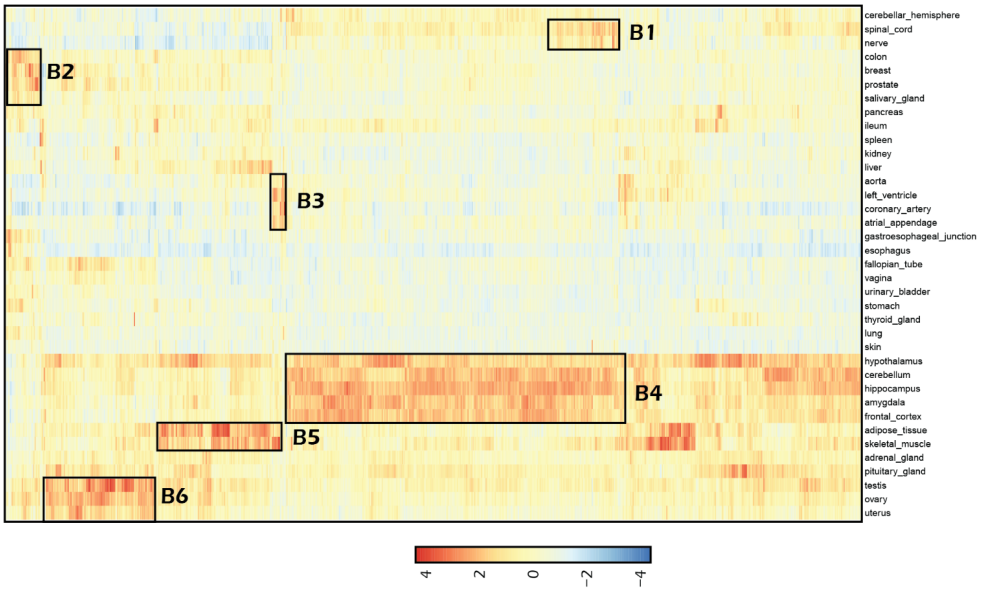


Figure 1: (A) Steps involved in representation learning from abstracts describing disease-gene associations. (B) Congruence of disease-symptom-causative gene triplets (schizophrenia-hallucination-ZNF804A; epilepsy-seizure-TCF4; migraine-headache-ATP1A2) in the word-vector space. (C) Expression and pathological role of *PSEN2* are shown as an example, on the human body template. While tissue wide expression levels of the gene are largely uniform, *Patho-score* shows remarkably higher intensity in the brain. Notably, mutations in *PSEN2* are the cause of familial Alzheimer's disease.

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.08.475476; this version posted January 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



B1 NRSN1, NGF, ASIC3

B2 EGFR, BIC, KRAS, MUC1, HOXB13

B3 MPI, MVP, CASQ1

B4 GABRG2, CA1, CA3, DBN1, APP, EGR3, NOS1, PDYN, DRD3

B5 SLC25A25, LIPE, PPARG, PLIN2, METRNL

B6 HOXA11, TAF4B, DDX4, FIGLA, FSH-R

Some representative genes for each block with well known disease implications

Figure 2: Heatmap depicting *Patho-scores* indicating tissue specific pathological involvement of genes. Each column of the heatmap represents one of the 1000 genes having the highest standard deviation of *Patho-scores* across tissues. For each block (manually boxed on the basis of coherent color intensities) some representative genes are shown with known pathological implications in concerned tissues.

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.08.475476; this version posted January 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

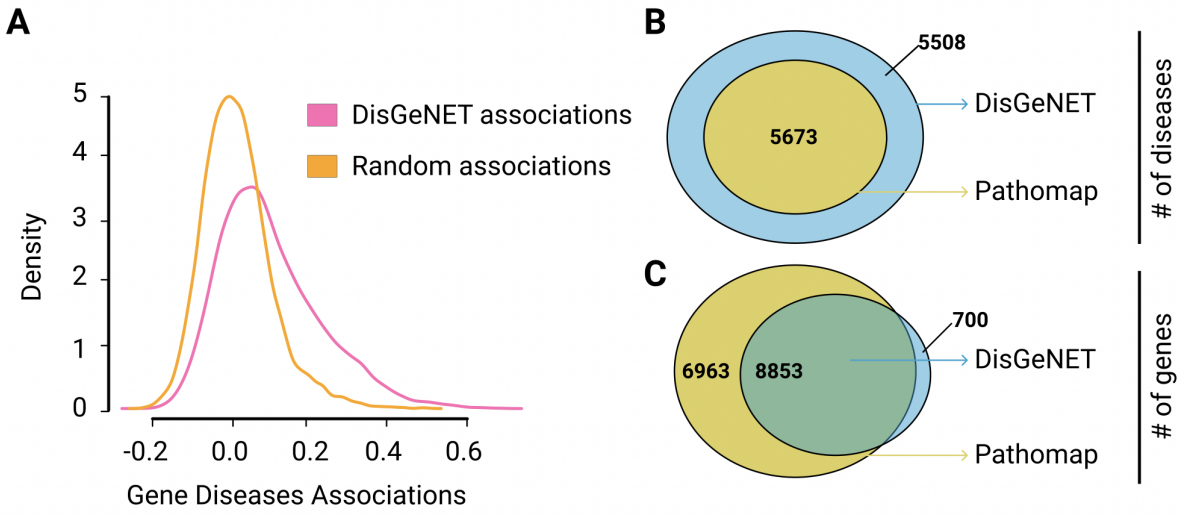


Figure 3: (A) The plot depicts the density distribution of the *Patho-score* of different gene-disease pairs from DisGeNET and compares it with the *Patho-score* distribution of random gene-disease pairs. The density distribution of random gene-disease pairs follows normal distribution whereas the gene-disease pairs from the DisGeNET database are right-skewed depicting the meaningful association between a gene and the corresponding disease. (B) The above figure represents that Pathomap covers more genes that are listed in the NCBI homo sapiens genes list in comparison to DisGeNET. (C) The gene-disease associations from the DisGeNET database has 11,181 disease names out of which 5673 disease names were also present in the Pathomap word embeddings.

An approach for automated descriptive answer verification.

Abesh Jha
Computer and Electronics Engineering Department
Kantipur Engineering College,
(Tribhuvan University)
Kathmandu, Nepal
magicianabesh123@gmail.com

Nikhil Mishra
Department of Computer Engineering
Delhi Technological University (DTU),
Delhi, India
nikhilmishra10@gmail.com

***Abstract*—Descriptive answer verification has been a modern-day requirement. It is undoubtedly a challenge for analysis and assessment of subjective answers along with their autonomous grading. Although objective answer evaluation has been in use for a long time, the analysis of subjective answers still possesses great complexity and there is no resolute or precise method for grading answers. Many methods have been implied and various metrics are currently being used for its evaluation with each one having its benefits and drawbacks. Thus, this research aims in tracking the feasibility of the metrics being used and providing an alternative approach for an automated system having relatively higher accuracy.**

***Index Terms*—Natural Language Processing(NLP), Text-rank algorithm, Cosine Similarity, Fuzzy Logic.**

I. Introduction

The concept of education and examination go hand in hand. It is essential to test the degree of knowledge absorbed by the learner to apply it practically. Since examinations are conducted periodically at a large scale, correcting examination papers can prove to be a soporific task. Analyzing hundreds of papers that contain monotonic answers has an element of declined enthusiasm attached to it, in addition to being time-consuming and unproductive. Hence, alternative approaches have been implemented with the purpose of computerization of evaluation but have been largely focused on objective answers which are relatively easy. However, objective examination solely cannot accommodate all types of questions and does not provide complete assurance in terms of the level of understanding of the subject. This research aims to automate the evaluation system for descriptive answers to make it fast processing and independent of manpower. It also guarantees the elimination of emotion from the whole checking procedure. Also, it intends to simplify the process of record-keeping and extraction. The proposed system will use a standard answer as a reference and utilize various metrics for comparison and subsequent allocation of marks through an unbiased checking procedure. Moreover, these kinds of automated systems not only lead to an efficient and unbiased checking system but are also time and cost-effective.

II. Literature Review

In March 2018, A. Shinde, C. Nirbhavane, S. Mahajan, V. Katkar, and S. Chaudhary presented a paper in which the proposed system constituted of three main steps: keywords and synonyms extraction, matching of keywords and synonyms, weighting the keywords, and generating score.[1] This evaluation system graded the answers depending upon the number of keywords matched. The major drawback of this system is that the keywords need to be provided in capital letters. For long sentences, this proves to be largely inconvenient. Also, the system only looks for those keywords that have been provided in the standard answer rather than extracting those from the students answer which may contain a few different words. Our system aims to overcome this by implementing the text rank algorithm to auto-extract keywords from both the answers separately. Similarly, another paper published in November 2014 by P. Sheeba provided another approach towards subjective answer verification. In this system, the evaluation process consists of four main steps: Pre-Processing, Part-of-Speech Tagging, Grading System, and Parse Tree. While the proposed system provided an element of semantic and syntactic analysis, it can only efficiently evaluate single sentences and an upgrade of the algorithm was required as explicitly mentioned in the paper. [2] Our system overcomes these hurdles as it efficiently evaluates multi-sentenced descriptive answers along with proper grammatical analysis. Another paper publishes in August 2013 by A. Kaur, M. Sasikumar, S. Nema, and S. Pawar proposes a graphical approach. The methodology used for the development of the algorithm is to match learners' single sentence answers with the standard single sentence answers by converting both answers into a graphical form and then intelligently matching their nodes and labels.[3] In the proposed approach, the graphical representation of descriptive answers to present complex information clearly and to represent knowledge in a machine-interpretable form has been considered. However, this approach fails when used for multiple sentences and consequently cannot be implemented in large-scale examinations. In our system, we

use various similarity matching metrics such as percent match, Levestanian distance, and cosine similarity to examine the degree of similarity between two answers. This method is better suited to evaluate long answers containing multiple sentences.

III. Scope Of Work

1) Implementation and analysis of text-rank algorithm for keyword extraction.
2) Automated grading of manual answers using five different metrics where syntactic analysis is not a requirement.

IV. Methodology

A. Proposed System

Our proposed system primarily consists of three major components: information extraction module, weighting module and score generation module as shown in Figure 1.

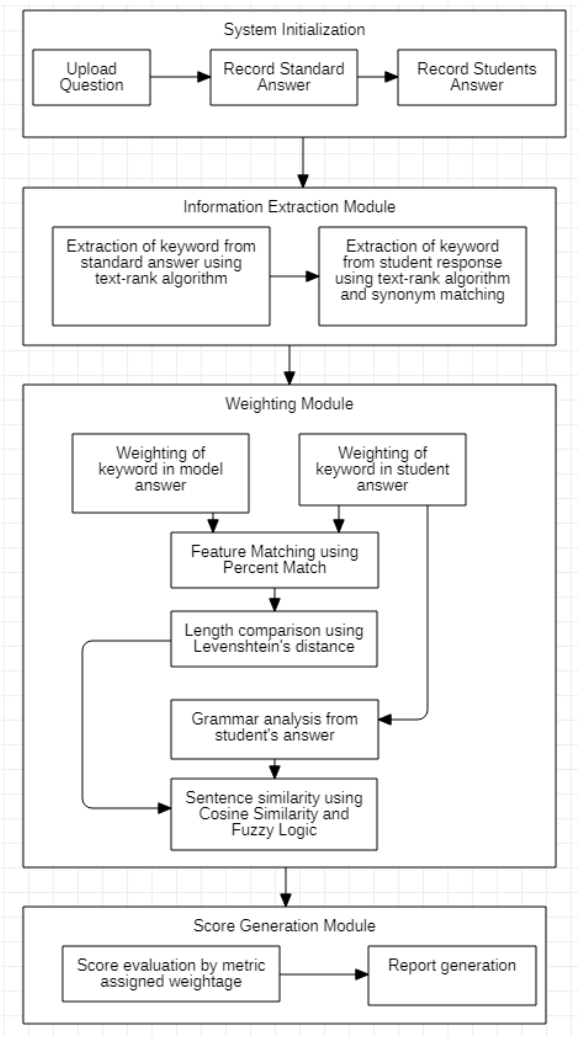


Fig. 1. System Diagram.

1) Initially, the system extracts keywords from the stored standard answer as well as the answer submitted by the user. Along with the keywords, the synonyms of the keywords submitted by the user will also be extracted and equated. The keywords that get repeated very often in a document are given less importance, while the keywords which occur rarely in a document will be assigned a higher value. The synonyms of the keywords will be extracted from the dictionary imported from a python library named PyDictonary.

2) The next step comprises of Information Extraction Module. The keywords are extracted and prioritized by the text-rank algorithm which updates the weights of every keyword according to the importance and frequency of words. Thus, only the most important keywords in a given paragraph are listed. Subsequently, keywords are extracted from the students answer as well and compared with the keywords from the standard answer. The weight of every matched keyword is added to obtain the final total.

3) After keywords are matched, we check if the length of the answer is close to the standard answer using Levenshtein distance. Successively, the percent match checks the accuracy of the percentage matched with the standard answer.

4) Further, cosine similarity is calculated to measure the similarity between two sentences by measuring the cosine of the angle between the two sentences by converting them into their respective vectors by implementing the word to vector process. Additionally, fuzzy string matching helps to find matching strings even if the string is not entered exactly.

5) Finally, the evaluation of the score is done based on marks predicted by the score generation module. This module provides the result based on the weight assigned to various metrics. Appropriate percentage is distributed according to the priority of the metrics.

B. Information Extraction Module

Text Rank Algorithm for Keywords Extraction

The task of the keyword extraction algorithm is to automatically identify a set of terms that best describe the document. Such keywords may constitute useful entries for building an automatic index for a document collection, can be used to classify a text or may serve as a concise summary for a given document. The Text Rank keyword extraction algorithm is fully unsupervised. The authors of the paper [4], introduce TextRank, a graph-based ranking model for text processing, and show how this model can be successfully used in natural language applications. First, the text is tokenized and annotated

with part of speech tags a preprocessing step required to enable the application of syntactic filters. To avoid excessive growth of the graph size by adding all possible combinations of sequences consisting of more than one lexical unit (ngrams), we consider only single words as candidates for addition to the graph, with multi-word keywords being eventually reconstructed in the post-processing phase. Next, all lexical units that pass the syntactic filter are added to the graph, and an edge is added between those lexical units that co-occur within a window of words. This is used to construct an undirected unweighted graph. After the graph is constructed, the score associated with each vertex is set to an initial value of 1 and the text-rank algorithm is run on the graph for several iterations until it converges usually for 20-30 iterations, at a threshold of 0.0001. Once a final score is obtained for each vertex in the graph, vertices are sorted in reversed order of their score and the top T vertices in the ranking are retained for post-processing.

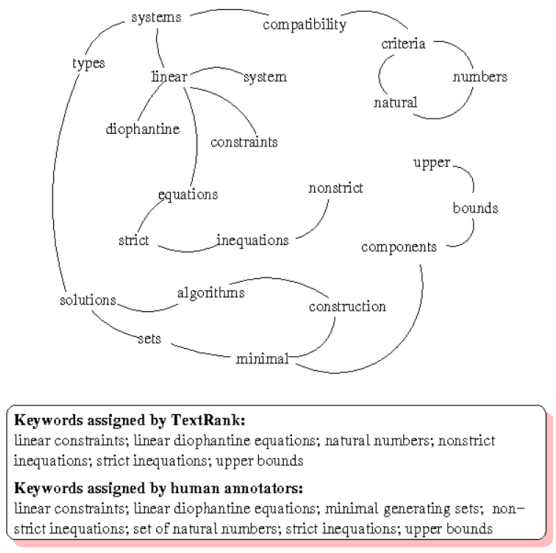


Fig. 2. Sample Graph for keyword extraction using Text Rank algorithm.[4]

Figure 2 shows a sample graph built for an abstract from our test collection. While the size of the abstracts ranges from 50 to 350 words, with an average size of 120 words, we have deliberately selected a very small abstract for illustration. For In this example, the lexical units found to have higher importance by the TextRank algorithm is (with the TextRank score indicated in parenthesis): numbers (1.46), inequations (1.45), linear (1.29), Diophantine (1.28), upper (0.99), bounds (0.99), strict (0.77). This ranking is different from the one rendered by simple word frequencies. For the same text, a frequency the approach provides the following top-ranked lexical units: systems (4), types (3), solutions (3), minimal (3), linear (2), inequations (2), algorithms (2). All other lexical units have a frequency of 1, and therefore cannot be ranked, but only listed.

It can be observed from Figure 2 that the text-rank algorithm prioritizes the words in terms of ranking which implies that if a student uses a keyword having a high rank, he receives a higher score as compared to the words which are given low priority by the algorithm.

C. Weighting Module (Evaluation of Metrics Used)

I)Levenshtein Distance

The difference in length between the two sentences is equally important. For example, a student might write an error-less sentence, however, be very limited in words. Thus marks must be deducted for the decreased length. Hence to calculate the length difference from the desired answer, Levenshteins distance is calculated from the standard answer. The Levenshtein distance is a string metric for measuring the difference between two sequences. Informally, the Levenshtein distance between two words is the minimum number of single-character edits (i.e. insertions, deletions, or substitutions) required to change one word into the other as shown in fig.3.

$$\mathrm{lev}_{a,b}(i,j) = \begin{cases} \max(i,j) & \text{if } \min(i,j) = 0, \\ \min \begin{cases} \mathrm{lev}_{a,b}(i-1,j)+1 \\ \mathrm{lev}_{a,b}(i,j-1)+1 \\ \mathrm{lev}_{a,b}(i-1,j-1)+1_{(a_i \neq b_j)} \end{cases} & \text{otherwise.} \end{cases}$$

Fig. 3. Mathematical representation for calculation of Levestian Distance. [5]

EXAMPLE : The Levenshtein distance between FLOMAX and VOLMAX is 3, since the following three edits change one into the other, and there is no way to do it with fewer than three edits.

II) Percentage match

Percentage match is used in the system to determine how closely two values match each other by calculating the Character Edit Distance between two String values, and also taking into account the length of the longer or shorter of the two values, by character count. In mathematical terms, the Character Match Percentage comparison uses the following formula to calculate its results as shown in equation 1 where:
In mathematical terms, the Character Match Percentage comparison uses the following formula to calculate its results as shown in equation 1 where:

CMP = Character Match Percentage
MCL = Maximum Character Length of the two values being compared, in characters
CED = the Character Edit Distance between two String values
CL = either the Maximum or Minimum Character

Length, depending on the setting of the Relate to shorter input option. If Relate to shorter input is set to No (as by default),the Maximum Character Length is used. If Relate to shorter input is set to Yes, the Minimum Character Length is used (that is, the number of characters in the shorter of the two values by character count).

$$CMP = \frac{MCL - CED}{CL} * 100 \quad (1)$$

III) Grammar

Grammar is one of the primary parameters used for a detailed evaluation of subjective answers since a sentence might contain just the important keywords but no correlation or grammatical equivalence between them which makes the answer highly equivocal. Hence, evaluation of grammar is inevitably a high priority to attain greater accuracy. The system uses grammar as one of the prominent evaluation metrics. We have used the Python library "Pydictionary" for analysis purposes. Our system uses three major processes for grammar.

Tokenization: It is used to create tokens by breaking up a sequence of strings into pieces such as words, keywords, phrases, symbols, and other elements for further processing.
Stemming: It is used to reduce the word to its word stem by removing the affixes or conversion to its respective root word.
Lemmatization: It is a normalization technique with benefits same as that of stemming. However, the major difference between them is that lemmatization is used to reduce the words to base form as opposed to stemming which just removes affixes.

In our system, tokenization occurs first which breaks the strings and sentences followed by stemming and lemmatization. A combination of stemming and lemmatization is required for better results. Since grammatical analysis tends to be very vague and complex, all these operations are performed using existing inbuilt libraries in our proposed system for faster and accurate evaluation.

IV) Fuzzy Logic

Approximate matching of strings is reviewed with the aim of surveying techniques suitable for finding an item in a database when there may be a spelling mistake or other error in the keyword. The methods found are classified as either equivalence or similarity problems.[6]
Fuzzy systems offer numerous methods that prove helpful in solving application problems, where the governing parameters bear fuzziness in nature. In particular, many variants of fuzzy decision-making models have been successfully explored as vocational guidance systems to assist decision-makers in recognition and guidance processes, especially in the Medical diagnostic realm.[8]
Fuzzy string search can be used in various applications such as a spell checker or a typos corrector. For example, when a user types "Missisaga" into Google, a list of hits is returned along with Showing results for "Mississauga". This implies that a search query returns results even if the user input contains additional or missing characters or other types of spelling errors. Fuzzy logic has been used in the proposed system in a similar context. It aims at providing some marks to the student's answers even if they are somewhat similar and doesnt require answers to be identical to the standard answer. Fuzzy logic is used in the system by importing the python library named fuzzywuzzy which takes two strings as input and returns a fuzzy value between them.

V) Cosine Similarity

Cosine similarity is used as a method for approximating how similar two words/sentence vectors are to each other. The intuition behind cosine similarity is relatively calculating the cosine of the angle between the two vectors to quantify how similar the two documents are.[7] The process of conversion of word to vector is done by a process called Bag Of Words. For example:
Sentence 1: AI is our friend and it has been friendly
Sentence 2: AI and humans have always been friendly

Step 1: We calculated Term Frequency using Bag of Words as shown in Fig.4:

Term Frequencies:

Sentence	AI	IS	FRIEND	HUMAN	ALWAYS	AND	BEEN	OUR	IT	HAS
1	1	1	2	0	0	1	1	1	1	1
2	1	0	1	1	1	1	1	0	0	1

Fig. 4. Calculation of Term Frequency by word count

Step 2: The main issue with term frequency counts shown above is that it favours the documents or sentences that are longer. One way to solve this issue is to normalize the term frequencies with the respective magnitudes or L2 norms. Summing up squares of each frequency and taking a square root, L2 norm of Sentence 1 is 3.3166 and Sentence 2 is 2.6458. Dividing above term frequencies with these norms, we get results as depicted in Fig 5.

Document	AI	IS	FRIEND	HUMAN	ALWAYS	AND	BEEN	OUR	IT	HAS
1	0.302	0.302	0.603	0	0	0.302	0.302	0.302	0.302	0.302
2	0.378	0	0.378	0.378	0.378	0.378	0.378	0	0	0.378

Fig. 5. Values after Normalization of Term Frequencies

Step 3: As we have already normalized the two

vectors to have a length of 1, we can calculate the cosine similarity with a dot product:
Cosine Similarity = (0.302*0.378) + (0.603*0.378) + (0.302*0.378) + (0.302*0.378) + (0.302*0.378) = 0.684
Therefore, cosine similarity of the given two sentences is 0.684.However, one major drawback of cosine similarity calculation using Bag of Words is semantics.While bag of words is intuitive and easy to implement, there is nothing in the document-term matrix which is capable of capturing the semantic similarity between words like hi and hey. Thus to overcome this, another feasible approach called word embeddings can be used which are basically vector representations of words which model semantic similarity through each words proximity to other words in the vector space.

D. Score Generation Module

The score is generated by the combination of all the metrics listed above. Appropriate weight is assigned to each of the metrics to obtain a final predicted score.
According to a published research paper entitled AI Answer Verifier, the score for marks is generated depending on the various factors and the total marks will be the sum of all the marks obtained from, individual, sections. Depending upon the priority, keywords are given 40 percent priority followed by grammar 30 and then synonyms 20 respectively and finally, the length of the answer is given 10 percent. The marks for length are dependent on the percentage of keywords matching and a making scheme is tabulated for giving marks to the length.[1] Thus, if the students write a completely different answer with no keywords matching, marks would not be allocated. Our proposed system also assigns similar distribution of percentage to the metrics, however, with some necessary changes. The problem with the distribution of weights mentioned above is that an answer may be grammatically correct but convey no actual meaning or be relevant to the question asked. Thus, to overcome this drawback, our proposed system only considers the evaluation of grammar only if at least 30 percent of the keyword is matched with the standard answer. Also, while grammar is an important factor to be taken into consideration, other parameters seem to be equally accountable for a proper evaluation. Hence, Grammar accounts for 20 percent of overall marks including synonym matching. Further, our proposed system takes cosine similarity and fuzzy logic into account equating to 20 percent of the total marks. Finally, the Levenshtein distance and percent match between the two answers: standard and students, accounts for ten percent each.

For example: A sample of standard answer and an answer written by student is taken as an input in the system as shown in fig.6. Further, fig.7 shows the

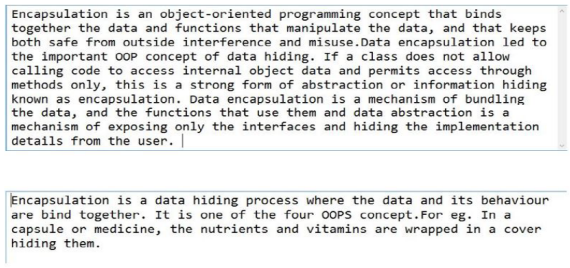


Fig. 6. Sample model of a standard answer and a manually written answer from the dataset.

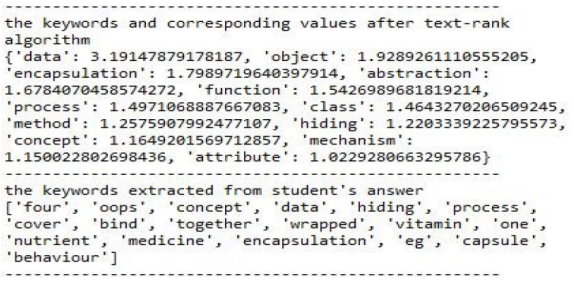


Fig. 7. Keyword extraction using Text rank algorithm for both standard and manual answers respectively.

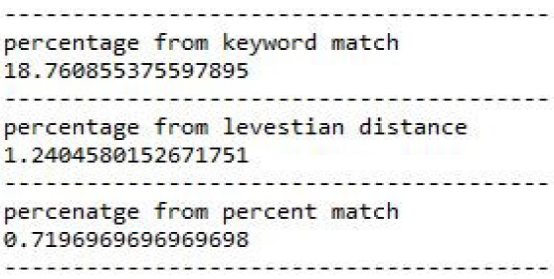


Fig. 8. Scores obtained for various metrics.

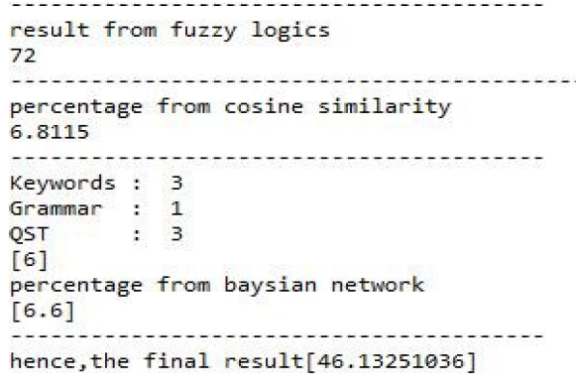


Fig. 9. Final score obtained after weight assignment and calculation.

important keywords extracted and their corresponding values after the implementation of Text Rank algorithm. Consequently, the marks obtained after the amount of keywords matched, levenshteins distnce , percent match and grammar is calculated individually as demonstrated in fig.8 . Finally,the respective weight assigned to each metrics is calculated to obtain the final score as shown in fig.9.

Result And Discussion

To check the reliability and the accuracy of the proposed system, we compared the results obtained from the system with the standard answers graded by some real evaluators. The dataset comprised of a standard answer and a readers commentary for a correct answer. Also numerous manually hand written answers by students for the same question was given and score was provided along with evaluation of mistakes.For experimental purpose, we took around 10 questions from the official GRE website where some prompt was given and students' answers were graded by real teachers. A score of 6 implied that it was a perfect response whereas a score of 1 was the lowest. We took the score 6 response as our standard answer and then inserted all the other responses having scores 5,4,3,2 and 1 respectively in our system. Respective scores obtained for each response were noted and then compared with the actual scores received for all the 10 responses. The scores were converted to a scale of 100. The results obtained can be demonstrated in a line chart as shown in Figure 10.

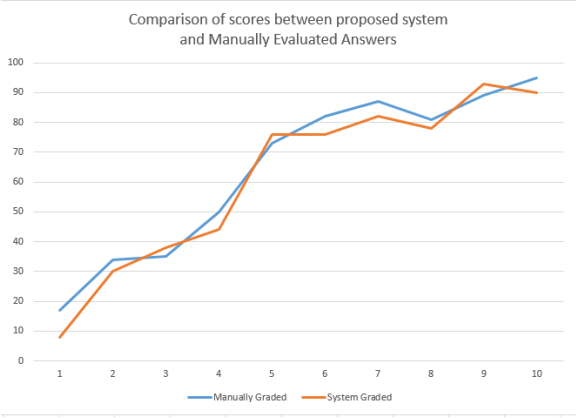


Fig. 10. Line Chart of manually evaluated scores and system evaluated scores for same answers.

It can be observed that the result is mostly similar in most of the answers apart from few outliers. This huge discrepancy in some cases can be credited to the fact that the standard answer may or may not include every possible answer. Since this system has a major limitation of syntactic analysis, an answer may be equally correct without possibly matching the standard answer. Therefore, in this case, the teachers grade it high but the system fails to do so. Such outliers were not taken into consideration or recorded in the graph shown. Also, the limited number of test results was due to the lack of availability of a dataset that mirrors our proposed system. To overcome these limitations, the system should be tested in a compromised environment where the data (students' answers) are more or less similar to that of the standard answer provided. Also, the question should be based on factual knowledge and keyword similarity rather than semantic analysis and cognitive reasoning.

Limitation

1) The students' answers should be more or less comparable to the standard answer stored. An entirely different variant of the answer, although correct, may not be graded properly by the system.
2) Lack of syntactic analysis. The system doesnt check for the meaning of words or how well an argument is supported or contradicted.

Future Enhancement

1) A Reinforcement learning approach will be used for better emulation of the manual checking process.
2) Currently the system is suited for only a single question. However, verification of multiple answer choices at the same time will be used.
3) The accuracy of the system will be checked on a large dataset comprising of manually graded answer sheets evaluated by real teachers.
4) Possible technologies such as deep learning will be incorporated in the system for higher accuracy.

Conclusion

Hence it can be seen that various metrics have their advantages and disadvantages. A trade-off between metrics is required to build a system having greater accuracy. Still, without the inclusion of syntactic analysis, the evaluation cannot be fully accurate. However, for grading descriptive answers by comparison with a standard answer, a combination of diverse metrics can be merged to obtain relatively higher precision. Although this inclusion accounts only for empirical and non-semantic answers, it nonetheless makes the evaluation process a lot easier for human accessors. Moreover, a Machine Learning approach is required for semantic analysis of answers and completely automated checking of descriptive answers.

References

[1] A. Shinde, C. Nirbhavane, S. Mahajan, V. Katkar, and S. Chaudhary, Ai answer verifier, International Journal of Engineering Research in Computer Science and Engineering, vol. 5, no. 3, pp. 143145, 2018.
[2] P. Sheeba, An approach to evaluate subjective questions for online examination system, International Journal of Innovative Research in Computer and Communication Engineering, vol. 2, no. 11, pp. 64106413, 2014.

[3] A. Kaur, M. Sasikumar, S. Nema, and S. Pawar, Algorithm for automatic evaluation of single sentence descriptive answer, International Journal of Inventive Engineering and Science, vol. 1, no. 9, pp. 112121, 2013.
[4] R. Mihalcea and P. Tarau, Textrank: Bringing order into text, in Proceedings of the 2004 conference on empirical methods in natural language processing, 2004,pp. 404411.
[5] R. Haldar and D. Mukhopadhyay, Levenshtein distance technique in dictionary lookup methods: An improved approach, arXiv preprint arXiv:1101.1232, 2011.
[6] P. A. Hall and G. R. Dowling, Approximate string matching, ACM computing surveys (CSUR), vol. 12, no. 4, pp. 381402, 1980.
[7] J. Ye, Cosine similarity measures for intuitionistic fuzzy sets and their applications, Mathematical and computer modelling, vol. 53, no. 1-2, pp. 9197, 2011.
[8] Thomas, A. et al. Intelligent Fuzzy Decision Making for Subjective Answer Evaluation Using Utility Functions. 2008 First International Conference on Emerging Trends in Engineering and Technology (2008): 587-591.

Microprocessors and Microsystems 89 (2022) 104412

Contents lists available at ScienceDirect

Microprocessors and Microsystems

journal homepage: www.elsevier.com/locate/micpro

An optimized weighted voting based ensemble model for DDoS attack detection and mitigation in SDN environment

Aastha Maheshwari *, Burhan Mehraj, Mohd Shaad Khan, Mohd Shaheem Idrisi

Department of Computer Science, Delhi Technological University, Delhi, India

ARTICLE INFO

Keywords:
Software defined networking
Distributed denial of service (DDoS)
Network security
Ensemble machine learning
Metaheuristic optimization

ABSTRACT

In recent years, software defined networking (SDN) has risen to prominence as a cutting-edge and promising networking approach. SDN is more secure and immune to DDoS attacks than traditional networks due to the bifurcation of the control and data planes, a global perspective of the whole network, centralized network control, and many other features. SDN still seems to be vulnerable to new DDoS attacks that can target the Application plane, control plane, or data plane, as well as the SDN architectural communication channels. In this paper, we propose a novel optimized weighted voting ensemble model to detect DDoS attack in an SDN environment. The proposed ensemble employs six base classifiers (two SVMs, two Random forests, and two Gradient Boosted Machines) that are differentiated by hyperparameter values. The optimal set of weights are identified by a novel hybrid metaheuristic optimization algorithm (BHO). Our approach eliminates false negatives by employing a novel dynamic fitness function. For training the ensemble framework we used CIC-DDoS2019 dataset. To examine the proposed model's performance, a testbed is developed using mininet and POX controller and loaded with multiple datasets. Our model outperforms all the recent approaches, achieving a low variance and high classification accuracies of 99.4163% and 99.3591% on CIC-DDoS2019 and CAIDA-2007 datasets respectively.

1. Introduction

The popularity of Web-based services and software has increased dramatically over the last two decades. Currently, the Internet is used by about 57 percent of the world's population [1]. As a result, there is a huge increase in Internet security concerns. Different security risks have often existed on the Internet. Worms, port scans, denial of service attacks, and Trojans, among other things, are all popular Internet outliers [2]. For such vast and complex networks, traditional network architectures and infrastructure do not have a flexible and obvious fix.

The three planes that make up network's work ability are the data plane, management plane and control plane. The vertical unification and ordination of the data and control planes into network devices using certain algorithms allow them to route, manage and alter network traffic in traditional IP-based networks. As a result, the network's entire structure becomes highly decentralized [3]. Devices are integrated with control logic in traditional networks. SDN is a dynamic programmable network replacing the fixed and complex traditional network. It is considered as one of the most significant networking breakthroughs [4].

SDN can be simply defined as the physical segregation of the control plane and the data forwarding plane of the network, and multiple devices in the network being controlled by control plane. SDN as stated by Kreutz et al. [5]:

1. Creating a separate unit called SDN controller by disengaging control functionality from network devices, and through this bifurcation of data and control plane network devices becomes fundamental forwarding components.
2. Flow-based, rather than destination-based, forwarding decisions are made using the packet's field values as a match criterion and collection of actions.
3. The network is programmable via APIs that communicate with underlying network devices in the data plane and run-on top of the SDN controller. APIs that run-on top of the SDN controller and interacts with fundamental network devices in the data plane.

The separation of data plane and control plane is at the heart of SDN architecture, which opens up a lot of possibilities for network design creativity [6]. SDN networks gain more programmability as a result

* Corresponding author.
E-mail addresses: aasthamaheshawari.am@gmail.com (A. Maheshwari), burhanmrj@gmail.com (B. Mehraj), shaadk7865@gmail.com (M.S. Khan), mohdshaheem91@gmail.com (M.S. Idrisi).

https://doi.org/10.1016/j.micpro.2021.104412
Received 2 August 2021; Received in revised form 4 October 2021; Accepted 10 December 2021
Available online 1 January 2022
0141-9331/© 2021 Published by Elsevier B.V.

of this bifurcation. As a result, network structure of SDN devices has become more manageable and organizable . In traditional networks, the control logic is embedded in the machines. If the policy has to be changed, vendor-specific attributes could be used to alter the logic on each system however that is a laborious task. Conversely SDN separates the control logic and puts it under unified control making it easy for system admins to modify control logic through APIs (southbound). Traditional switches are difficult to enforce new disruptive policies because they have minimal hardware potential and are vendor-specific. However, due to the unified controller in SDN, policy updation becomes very quick and convenient. In a conventional network, testing new policies is difficult due to limited area, but in an SDN network, there are far more possibilities for research.

A DoS attack is defined as a concerted effort to obstruct the use of specific network services by benign users.

To launch a Distributed Denial of Service (DDoS) attack, assailant start recruiting several devices which are geographically distributed. These infected computers are also referred to as bots and the collection of these bots is called a botnet [5]. The entire botnet is managed remotely by the bot-master [7–9] user. The attacker sends instructions to all the affected devices to initiate a DDoS attack, which would then be followed by unnecessary traffic on the target.

In the conventional network, the scale, frequency, intensity and complexity of DDoS attacks are increasing. This implies that the contemporary IDS for DDoS attacks are insufficient to tackle the obstacle. Attackers also strained techniques to circumvent current guarding shields [10,11]. However, SDN has features that have some benefits in defeating the DDoS attacks [11–15].

SDN can mitigate DDoS attacks by bifurcating the control plane and the data plane. But, because of its architecture, SDN itself is prone to DDoS attacks. SDN is divided vertically into three functional layers, as previously mentioned. DDoS attacks can be targeted to any of these layers. Layer wise vulnerabilities are summarized below:

- Data plane: In the data plane each switch has a limited sized flow table which makes it vulnerable to DDoS attack as an attack can cause flow table overflow [16]. Another potential vulnerability is buffer saturation [17]. Buffer memory is used in new flow rule generation process and its limited size makes data-plane vulnerable to DDoS attacks.
- Control-Data plane link: Sending numerous dummy packets to the switches, which transfers them to the controller through the control-data plane link. This flooding of packets through the channel can cause congestion in the channel.
- Control plane: Ultimately, as the attacker directs a flood of dummy packets to switches, all the requests are transferred by the switches to the controller, which causes the controller to become overburdened with fake requests, known as controller saturation.

In this paper, we propose an optimized weighted voting ensemble method to detect and mitigate DDoS attack in SDN environment that uses two SVMs, two random forests, and two Gradient boosted machines base classifiers to classify data. Two base learners of each type are distinguished in terms of the relevant hyperparameter values. The major contributions of our paper are:

1. A novel optimized weighted voting ensemble model.
2. A novel dynamic fitness function is proposed to reduce the false negatives while finding the optimal set of weights for the proposed ensemble.
3. BHO - A novel hybrid metaheuristic optimization algorithm is proposed which has improved exploration and exploitation capabilities and can utilize parallel processing to reduce the time complexity.

The remainder of this paper is arranged accordingly. A summary of the related works is given in Section 2. The main concepts used in our model are described in Section 3. The proposed model has been explained in Section 4. Experimental results are discussed in Section 5 using various performance metrics. Finally conclusion and future work is stated in Section 6.

2. Related work

DDoS identification and prevention have been a top priority for researchers in recent years. Many of the other researchers suggested DDoS attack prevention mechanisms, as well as strategies for minimizing the impact of DDoS attacks. This section provides a comprehensive discussion of current research in DDoS attack detection and prevention in the sense of SDN.

2.1. Information theory based approach

Giotis et al. [18] used a sFlow protocol to detect DDoS attack. This protocol limits the communication between the OpenFlow switch and controller thereby lowering the overload of controllers during the large network traffic conditions. These flow rules are also used by the mitigation module to block malicious flow i.e., DDoS attacks. However, the use of flow sampling using the sFlow protocol also affected the accuracy of detection.

Tsai et al. [19] employed a DDoS detection and prevention system based on entropy in which, an application on the controller monitors incoming traffic and computes the Entropy which is compared against a fixed entropy threshold. If it detects an intrusion, it will assist in the implementation of a prevention technique to block the relevant port, and this programme will then transfer this information to Network Intrusion Detection System (NIDS) for further inspection, and NIDS will be able to avoid this malicious traffic. One major limitation is choosing an appropriate entropy threshold value as it highly affects the detection rate.

Kalkan et al. [20] proposed a novel model JESS which is an acronym for Joint Entropy-based Security Scheme having three states, nominal, preparatory, and active mitigation stage. It is an entropy-dependent scoring system for detection and mitigation of DDoS attacks. It is based on the target IP address entropy, and dynamic combinations of TCP and IP layer attributes. This system resolves the problem of capable switch, which is a contentious problem in the literature, since it contradicts the fundamental properties of SDN such as distributed forwarding and centralized control. When the controller detects a threat, it notifies the switch. Then the attack packets are discharged by the Attack Mitigation module while protecting lawful packets.

Bawany et al. [21] introduced an flexible architecture known as SEcure and AgiLe (SEAL) that includes 3 different modules for solving the problem of DDoS attack detection and prevention : a-defence, c-defence, and d-defence. By using a customized category of Estimated weighted moving average filters, SEAL attains adaptability. Commonly used filter includes active, constructive and passive filters. For result analysis they generated traffic in controlled environment, which is the major shortcoming of proposed work.

Cui et al. [22] suggested using cognitive-inspired computation with dual address entropy to detect DDoS attacks. SVM learning was used for classification. Information of the switch from the flow table are obtained. DDoS attacks are observed where the entropy of the target is smaller and the entropy of the origin is exceeding the usual threshold of the traffic. To mitigate a DDoS threat, the DDoS attack defence dumps all table records. The proposed technique identifies attacks quickly and has a good detection rate and low FPR

Nada et al. [23] proposed a hybrid method which assigns a node trust value after detecting if any nodes are involved in DDoS attacks. Offending nodes are banned for certain period of time. But the flexibility of algorithm gives them an opportunity to reintegrate. But, if malevolent conduct is repeated, the nodes are forbid forever. The proposed approach is more robust as it takes into consideration the

dynamics of current network traffic and can be adapted in real time as the traffic complexity of a network varies. The major drawback is the choice of threshold value, which is selected experimentally in this work. The threshold value may be calculated using machine learning.

A common downside of these entropy-based techniques is the need to select the optimal detection threshold and poor accuracy at lower attack rates

2.2. Machine learning based approach

Li et al. [24] suggested a two-tier intrusion detection system (IDS) that identifies network anomalies by capturing global network flows. Two algorithms were used for extraction of features, i.e. swarm bat algorithm and Binary differential. Random forest algorithm is then used in order to characterize the movement of traffic to determine the system's possible interference by adapting the weights of samples using a voting ensemble. The downside is that this approach is not implemented and tested in real-world situations.

Phan et al. [25] proposed hybrid machine learning model to defend against a DDoS attack. They used modified History based IP Filtering known as enhanced HIPF (eHIPF) thereby empowering the model to detect and classify the network quickly. During the creation of eHIPF attack, to drop each individual packet at the border of the cloud, a flow mod message is sent with a command of drop down action. The proposed approach is inefficient to mitigate probe/reconnaissance attacks because the traffic volume generation of these attacks is very low

Myint et al. [26] used a tool for detecting DDoS attacks with minimal overhead using an advanced support vector machine. Volumetric and asymmetric characteristics are used to limit preparation time. Two forms of flooding: UDP flooding and ICMP flooding have been successfully identified. No information on the mitigation approach was provided by the authors.

Polat et al. [27] compared different ML algorithms based on feature selection. Authors used SVM, KNN, Naïve Bayes and ANN. In order to observe the accuracy they used multiple ML methods with varied feature selection. Since the emphasis was also on the best feature selection, a certain approach was used to pick many features.

Jia et al. [28] suggested a detection approach that would combine many multi-classifier de-compositions with a Single Value Decomposition (SVD). They believe that building different classifications is more accurate than static classifications. It would be insightful to measure the efficiency of the process when applied to a big dataset. The model produces downgrading results as compared to the K-NN algorithm.

Ahuja et al. [29] identified novel features for DDoS attack detection and logged them onto a CSV file to create a dataset specifically for DDoS attack. In the training phase, they used a hybrid consisting of a support vector classifier combined with an additional filtration by random forest (SVC-RF) and achieved a high accuracy while maintaining a low false alarm rate. The major drawback of the proposed hybrid is that it emphasizes more on the decision of random forest as compared to that of Support vector classifier.

2.3. Artificial neural network based approach

Cui et al. [30] developed a software-defined networking SD-Anti-DDoS defence mechanism. The proposed scheme includes 4 modules in sequence. Trigger module for a DDoS attack detection raise the reaction rate for the event trigger. BPNN technology is used to detect a traffic anomaly after activation by the trigger module to detect the attack. According to the path of detection module, the migration module obstruct traffic to minimize the impact of recent module. The other remaining modules are attack detection and traceback. The findings show that SD-Anti-DDoS can detect and identify the source of the attack easily within a second. However, the tool TFN2k they used to construct the traffic is outdated in current times.

Li et al. [31] used deep learning to solve the problem of DDoS attack detection in SDN environment. This software features a functionality analyser that analyses the unprocessed data units and generates detailed training information. The Flow Table Generator calculate flood entries and priorities of numerous Attack sets and sends them on the basis of statistics provided by the statistical module to the OpenFlow switch. Authors used a DDoS defence architecture tested by real-time attachments and used the ISCX dataset for training purposes. It was using hardware to design the experimental set-up. The major drawback of the proposed model was the problem of processing overhead, which can slow down the network significantly.

Novaes et al. [32] suggested a composite system for the detection and mitigation of port scanning and DDoS attacks in SDN network. This scheme is used in three stages to track attacks, including characterization, deviation and prevention. Authors use entropy metric for quantifying the network attribute and use LSTM (Last Short-Term Memory) for modelling the signature of each standard traffic attribute. The anomalies on the network were then detected using furious logic. They validate their approach by implementing floodlight control via Mininet simulation.

Gharvirian et al. [33] used Fast entropy rather than entropy for detecting attacks in SDN particularly DDoS, and therefore the detection time for the calculation is minimized and the threshold value is modified according to traffic conditions. Another good approach for improving the accuracy of the attack detection used in these experiments is the use of the neural perceptron network for the analysis of flux statistics. The neural network improves the accuracy of the detection and false warning rate in the course of this study and demonstrates the nature of an attack in examining the 3 travel functions of the network. Given the potency of the detection period, the proposed model used the neural network just for suspicious flows. Also, they have not mentioned the way to mitigate the attack.

Liu et al. [34] proposed a method for DDoS detection that combines the entropy method with the Particle Swarm Optimization based Back Propagation neural network (PSO-BP). The method uses the generic entropy method distributed on the switch to pre-diagnose the traffic. After identifying the switch at which the abnormal activity was first addressed, if the calculated entropy is greater than a certain threshold, the model uses PSO-BP neural network to extract the flow features of that switch which would be utilized for detection of DDoS attack. The major drawback of this approach is that the model might misclassify random entropy bursts in realistic networks as attacks. The change in network schema results into highly varied threshold value making this approach network scenario dependent.

Hannache et al. [35] introduced a Back Propagation Neural-network based Traffic Flow Classifier for real-time DDoS detection in SDN. The author trained the model using a custom-made dataset containing features extracted from both incoming packet messages received on SDN controller, and Openflow (OF) messages reciprocating between SDN controller and Virtual Open Flow Switches (VOFS). The model is capable of exploring the OF traffic by utilizing the fact that SDN allows deep packet analysis thereby controlling the comprehensive behaviour of the virtual environment. Along with detection, the proposed model also possesses the capability of mitigating the DDoS attacks. This paper lacked the extensive experimentation on multiple standard datasets which makes it susceptible to high variance problem.

2.4. Miscellaneous approach

Wang et al. [36] proposed DaMask which is a cloud-based DDoS detection approach. DaMask has is capable of detecting as well as mitigating the DDoS attack, for that it has two modules: DaMask-D and DaMask-M which are the detection module and mitigation module respectively. If the attack is detected by the DaMask-D module, along with the generation of alert, the information regarding the malicious packet is sent to DaMask-M to mitigate the attack. The author put

more emphasis on tackling DDoS detection in SDN for cloud-based environments rather than any arbitrary environment.

In order to shield SDN network from flow table overloading DDoS attacks, Bhushan et al. [37] proposed a novel flow table sharing based approach. The authors suggested using other space in the flow table for relatively little overhead contact. As a consequence, the rise in network resistance to flow table overflows DDoS attacks. The efficiency of the process in the attack is influenced to establish the holding time effectively. The efficacy of the procedure depends, however, on the time required.

Aleroud et al. [38] proposed a novel graph theory based approach. They used a catalogue of attack signatures from conventional network sets of data. This technique uses regularly labelled flows in order to analyse a new OpenFlow sample and to see how the existing signature is used for flow analysis. The drawback however is that the detection is based only on the signatures of the attack

Wang et al. [39] used a several path rerouting strategy to minimize the impact of LFA (Link Flooding Attacks), which at the time contains many target links. The method selects multiple optional routes and decides how much traffic to the optional connections should be redirected. The attacker has to locate a new target connection after rerouting. This will find the LFA bots and the malicious traffic from the LFA bots is then eliminated.

We learned from the literature review that ensemble learning employs weak base learners (with low accuracy) and that majority voting may result to high accuracy and low false alarm rate. Since, in an ensemble, base learners cannot perform equally well. Therefore, different weights should be assigned to the base learners in order to improve classification performance. To date, many different approaches have been proposed to find optimal weights of base learners in an ensemble. R. Burduk. [40] proposed a fuzzy set based approach in which interval valued fuzzy set is utilized to represent weights. There also exists several uses of probabilistic framework to find the optimal weights of base learners [41–43].

To find optimal weights, in addition to these analytical techniques, metaheuristic techniques can also be used. Zhang et al. [44] proposed differential evaluation based approach to find optimal weights and build an ensemble model with strong generalization ability. Onan et al. [45] defined a multi-objective function based on the classification accuracy of each base learners and used differential evaluation to optimize it to solve the problem of text sentiment classification and achieved a high classification accuracy. But, due to no data preprocessing their model underperformed.

3. Major concepts

3.1. DDoS attack on SDN

Software Defined Networking (SDN) network architecture provides many benefits thereby making it is resilient to DDoS attacks. However, SDN has some security issues, specifically its controller is vulnerable against the attacks of the Distributed Denial of Service (DDoS). When DDoS attacks occur against the SDN controller, the mechanism and connectivity capability of the controller are overwhelmed. Layer-by-layer overview of potential DDoS attacks on SDN are shown in Table 1.

Table 1
Layer-by-layer overview of potential DDoS attacks.

S.No.	SDN layer	DDoS attacks	Description
1	Application	Un-Authorized Application	Malicious applications are given access by instances of other applications and change network behaviour and are caused by network performance degradation.
2	Data plane	Flow table overflow	To store the flow tables, the switches have TCAM memory. The high cost and strength of memory makes this memory minimal.
		Spoofing of switch	IP address of the switch is spoofed and used to send control messages with changed address. Second switch is used to establish a connection with the controller [9].
		Buffer saturation	Packet header field must be buffered by switches. OpenFlow switches has restricted memory to buffer the data.
3	Control plane	Packet-in flooding	Overloading of the controller due to switch stimulated by spoofed IP's sending bulk packet_in messages [8].
		Saturation of controller	The central controller shall draw up a new flow rule for each switch in the network. The controller can be overloaded with several malicious requests [7,12].
4	Communicat--ions link	Congestion of south bound API	Switch requests new flow rules with Packet in messages from the controller for each new traffic flow and receive the responses with Packet out messages that trigger high traffic in the connection [6].

3.2. Support vector machine

One of the most effective binary classification techniques is the Support Vector Machine (SVM) [46]. SVM is widely used for its high generalization ability and strong resistance towards overfitting. SVMs tend to be resistant to overfitting because, rather than minimizing the squared error or an absolute value of an error they minimize structural risk.

Formulation of structural risk minimization (SRM) problem can be done in terms of data, given a dataset the objective $L(\theta)$ can be written as:

$$L(\theta) = \frac{1}{2n}\sum_{i=1}^{n}(h_\theta(x^i) - y^i)^2 + \frac{\lambda}{2}\sum_{j=1}^{d}\theta_j^2 \quad (1)$$

The first term in the Eq. (1) is mean squared error (MSE) term between learned model h_θ, and given labels y. The second term is the regularization term, which penalize larger weights and supports sparsity. λ is trade-off coefficient, a hyperparameter which affect the emphasis of regularization term.

By establishing a hyperplane or a collection of hyperplanes SVM classifies data into different classes. In SVM a hyperplane is described as a surface whereby the data points are completely sectioned in two classes to prevent the fall of data points of one class to the other, as shown in Fig. 1. The hyperplane with the greatest distance between the nearest points in the two groups is considered to be the best out of an excessive extent of hyperplanes [46]. A hyperplane is the collection of points $\vec{x}$ that satisfies the equation:

$$\vec{w}.\vec{x} + b = 0 \quad (2)$$

in Eq. (2) b and w are the bias and the weight vector respectively. The Lagrangian multipliers in the linear case can be used to solve these. The original maximum-margin hyperplane was introduced for linearly separable cases. But, SVM can also be useful in nonlinear classification problems, it does so by utilizing kernel trick with maximum margin hyperplanes. This enables the algorithm to transform feature space to a higher dimension and fit the maximum-margin hyperplane [47], as shown in Fig. 1. Commonly used kernel functions are listed in Table 2.

3.3. Random forest

Decision trees [48] are a powerful and widely utilized method for different machine learning tasks. However, they are rarely accurate due

Table 2
Commonly used kernel function.

Function	Equation
Linear kernel	$f(a_n, a_i) = (a_n, a_i)$
Polynomial kernel	$f(a_n, a_i) = (\gamma(a_n, a_i) + r)^d$
Sigmoid Kernel	$f(a_n, a_i) = \tanh(\gamma(a_n, a_i) + r)$
RBF kernel	$f(a_n, a_i) = exp(-\gamma \|\|a_n, a_i\|\|^2 + C)$

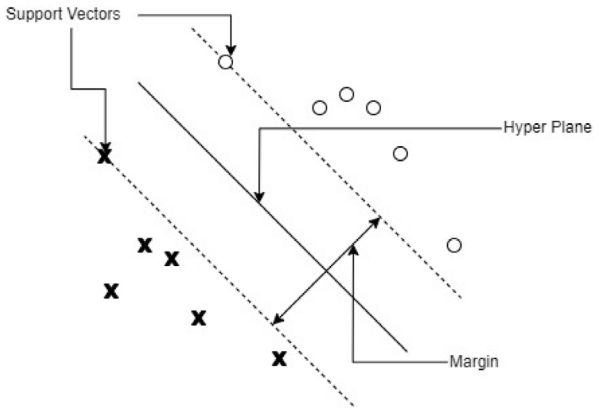


Fig. 1. Support vector machine.

to their low bias [49] yet high variance. Random forests [50] are a method of integrating a number of deep-seated decision-making trees trained in different sectors on the same training set with the aim of diminishing variance and enhancing model performance.

Brieman [50] integrates Bagging [51] with node-by-node random variable selection [52], given a training dataset $(x_i, y_i)_{i=1}^n$, repeatedly bagging (M times) chooses random subset samples of training set (using random replacement strategy) and fit trees to the chosen subset:

For m = 1, . . . , M:

1. Sample subset, with random replacement of n training data points from X, Y; say X_m, Y_m.
2. Using modified tree learning algorithm train a classification tree f_m on X_m, Y_m.

The modified tree learning algorithm picks a random sub-set of attributes at each division of the feasible candidate. "Feature bagging" is a term used to describe this procedure and it is shown in the algorithm 1.

Algorithm 1: Find spl^* (the best spilt) that divides f_t, among a random subset of D≤p input variables.

```
1  Function FindBestSplitRandom(f_t, D):
2      Δ = −∞
3      Draw D random indicies j_d from 1,...,p
4      for d = 1, ··· , d do
5          Find the best binary split spl*_{j_d} defined in X_{j_d}
6          if Δi(spl*_{j_d}) > Δ then
7              Δ = Δi(spl*_{j_d}, t)
8              spl* = spl*_{j_d}
9          end
10     end
11     return spl*
12 End
```

3.4. Gradient boosted machine

Gradient boosting [53–55] is a learning technique that combines the results of a large number of simple predictors to create an efficient committee with better performance than single members. Gradient boosting creates additive classification models by iteratively fitting a basic parameterized function (base learner) to current pseudo-residuals using least squares. As the iteration proceeds further, pseudo-residuals are being minimized as they are the gradient of loss function.

This machine learning algorithm is well-known for producing state-of-the-art results on a variety of complex datasets. GBM can be used for both regression and classification task.

Algorithm 2: Gradient boosted machine

```
   Input : (x_i, y_i)_{i=1}^n, L(y_i, F(x))
1  Begin:
2      F_o(x) = argmin_γ Σ_{i=1}^n L(y_i, γ)
3      for m = 1, ..., M do
4          for i=1 to n do
5              r_im = −[∂L(y_i, F(x_i)) / ∂F(x_i)]_{F(x)=F_{m−1}(x)}
6          end
7          for j=1,...,J_m do
8              Fit a regression tree to the r_im values and create
                terminal regions R_jm
9          end
10         for j=1,...,J_m do
11             γ_jm = argmin_γ Σ_{x_i∈R_ij} L(y_i, F_{m−1}(x_i) + γ)
12         end
13         F_m(x) = F_{m−1}(x) + ν Σ_{j=1}^{J_m} γ_m I(x ∈ R_jm)
14     end
15 End
   Output: F_M(x)
```

As inputs, the algorithm 2 takes a Dataset and a differentiable loss function. In step 2, the model is initialized with a constant value $F_0(x)$. M base learners are sequentially trained from lines 3 to 14. After calculating pseudo-residuals in steps 4–6, a base learner is trained to learn these pseudo residuals and create terminal regions in steps 7–9. In steps 10 to 12, the optimum value of γ_{jm} is calculated. In Step 13, using additive expansion technique Fm(x) is updated.

The procedure's learning rate is regulated by the "shrinkage" parameter ν, where $0 < \nu \leq 1$. Empirically, small values ($\nu \leq 0.1$) have been found to lead to a much better generalization error. I($x \in R_{jm}$) counts the number of data-points falling under the terminal region R_{jm}.

4. Proposed approach

This section describes our approach for detecting DDoS attacks in an SDN environment. The proposed approach consists of a learning phase followed by a detection and mitigation phase.

4.1. Learning phase

The architecture of the proposed optimized weighted voting ensemble model is given in Fig. 2. For training the model we used DDoS Evaluation Dataset (CIC-DDoS2019) [56] and for testing we used both CIC-DDoS2019 and CAIDA-2007 [57]. It is among the latest datasets for DDos Evaluation and is available in both CSV and PCAP format which makes it ideal for training the machine learning model as well as injecting it to SDN simulation (Mininet). The dataset consists of 12 different DDoS attacks using SYN PortScan, WebDDoS, SNMP, MSSQL, DNS, LDAP, SSDP, UDP, UDP-Lag, NetBIOS, NTP, and TFTP at different time. The availability of 80 features gives a huge amount of versatility in the precise selection of features and fine-tune it for our study.

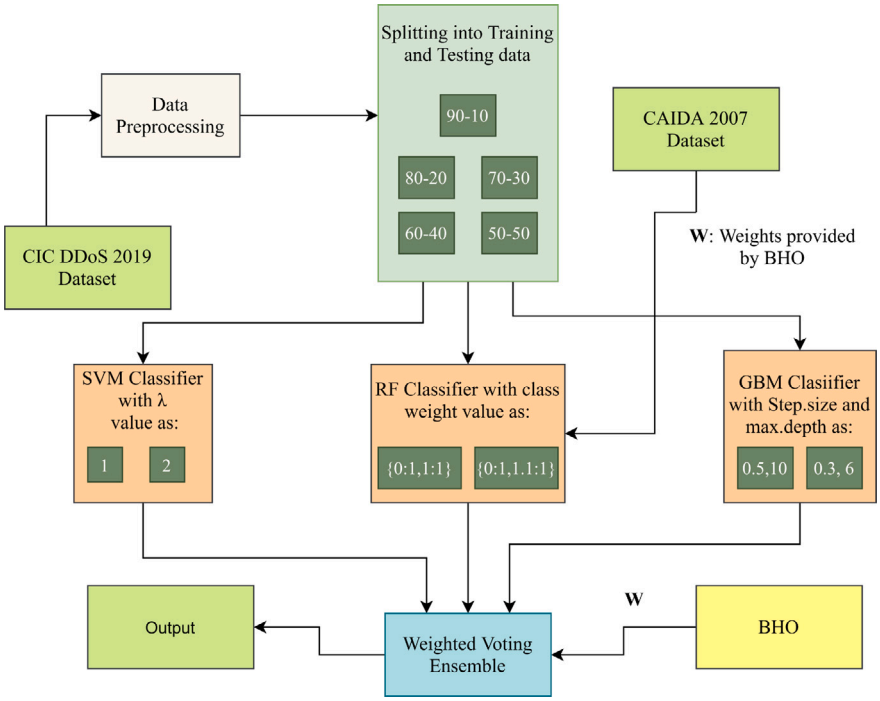


Fig. 2. Architecture of the proposed model.

In the first step, data is preprocessed to make it fit for training the ensemble. The dataset is highly unbalanced with more than 90% of packets related being DDoS which significantly increases the number of false positives. So, we balanced the dataset by undersampling it. We intended to balance the dataset without losing the versatile attack nature of the dataset. So, we subsampled the packets from each type of attack and used all the benign packets available in the dataset.

We chose a total of 20 features shown in Table 3. We wanted the model to be more generalized so we avoided the testbed related features like IP addresses and ports of both victim and attacker. From the definition, in a DDoS attack, the perpetrator tries to make the resources unavailable by flooding them with requests which make features related to the duration of each flow, packets per second and bytes per second most distinctive among all. During a DDoS attack, the machine becomes unresponsive which makes the Download versus Upload Ratio High for the host and makes the attack predictable. After analysing all such factors we procured the 20 most distinctive yet not machine-specific features for our study as shown in Table 3. After data preprocessing step dataset is divided into training and testing datasets.

Table 3
List of features used for training and testing the model.

Feature	Explanation
Flow duration	Duration of the flow (microsecs)
Tot Fwd Pkts	No. of total forward pkts
Fwd Pkt Len Max	Maximum size of forward pkts
Fwd Header Len	Total length for forward headers in bytes
Bwd Pkt Len Mean	Mean size of backward pkts
Tot Bwd Pkts	No. of total backward pkts
having Flow IAT Mean	Inter-Arrival Time of two pkts sent in the flow
Fwd IAT Tot	Total Inter-Arrival Time of two forward packets
Bwd IAT Tot	Total Inter-Arrival Time of two forward packets
Bwd PSH Flags	No. of times the PSH flag was set in backward packets
Fwd Pkts/s	pkts per second in forward direction
Bwd Pkts/s	pkts per second in backward direction
Pkt Len Std	Standard deviation length of a pkt in bytes
RST Flag Count	No. of pkts having RESET Flag
CWE Flag Count	No. of pkts having CWE Flag
Down/Up Ratio	Download and upload ratio of the flow
Fwd Byts/b Avg	Avg bytes bulk rate (ABBR) towards forward direction
Fwd Seg Size Min	Avg size of segment (ASoS) towards forward direction
Active Mean	Mean of the time a flow was active
PSH Flag Count	No. of pkts having PUSH Flag

In the second step, data is trained on each base learner. Since the performance of base learner depends on the choice of hyperparameter value. Therefore, to ensure that the base learners are used to their full potential, for each base learner we trained two different types with different hyperparameter values. We used linear SVM which uses linear kernel function for classification. Two linear SVM base learners are trained with hyperparameter λ (trade-off coefficient) value as 1 and 2. Similarly, two random forest (RF) base learners are trained with hyperparameter class_weight value as {0: 1, 1: 1} and {1: 1.2, 0: 1}. And the two gradient boosted machines (GBM) base learners are trained with hyperparameter step size value as 0.5 and 0.3, and max_depth as 10 and 6 respectively.

In the third step, the outcomes of SVM, Random forest and GBM base learners are ensembled using weighted majority voting (WMV). Littlestone and Warmuth [58] showed that providing weights to the base learners reduces the errors made by the ensemble system.

In this study, we use Optimized weighted voting ensemble which provide weights to all the base learners based on their classification accuracy. To find the optimal weights we propose a novel dynamic fitness function and BHO, which is a novel metaheuristic optimization algorithm. BHO is a hybrid of Brain Storm optimization algorithm (BSO) [59] and Henry's gas solubility optimization algorithm (HGSO) [60].

4.1.1. Proposed novel dynamic fitness function

Meta-heuristic optimization algorithms use fitness functions to guide simulations towards an optimal solution. The two main categories of fitness functions are static and dynamic fitness functions. The output of a static fitness function does not change with the same input over the course of iteration. A dynamic fitness function, on the other hand, may have different outputs for the same input as iteration changes.

Static fitness function for Optimized weighted voting ensemble:

Since the weights assigned to each base learner depends on their accuracy. Therefore, a fitness function can be defined as:

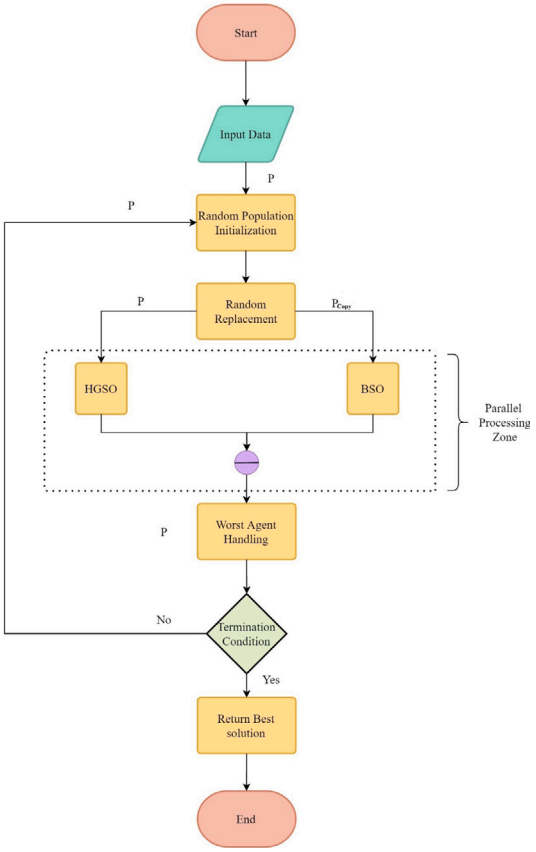


Fig. 3. Flowchart of the proposed model.

F (W, base learners, test dataset) =

$$1 - \text{Accuracy} = \frac{FN + FP}{TP + TN + FN + FP}$$

where, TN is true negative, TP is true positive, FP is false positive and FN is false negative.

Proposed dynamic fitness function for Optimized weighted voting ensemble:

Experimentally it was found that false negatives were more as compare to false positives, as show in experiment Section 5.2.3. Therefore, to reduce the number of false negatives we designed a dynamic fitness function as:

F (W, base learners, test dataset, σ) = $\frac{\alpha FN + \beta FP}{TP + TN + \alpha FN + \beta FP}$

The aim of this dynamic fitness function is to assign penalty weights α and β to FN and FP predictions respectively. Initially α will be more than β to reduce rate of FN, but as iteration proceeds both penalties approach each other to become equal to 1. σ is a hyper parameter which sets the maximum gap between α and β.

If T is the maximum number of iteration and t is the current iteration, then value of α and β are calculated as:

$\alpha = 1 + \gamma[1 - \sin(\frac{\pi t}{2T})]$; $\beta = 1 - \gamma[1 - \sin(\frac{\pi t}{2T})]$

$\gamma = \frac{\sigma(1-1/N)}{2}$; where γ takes into account the population size N and adjust σ accordingly.

4.1.2. BHO - novel hybrid meta-heuristic optimization algorithm

Metaheuristic algorithms are stochastic, population based computational intelligence paradigms that are specifically designed to solve complex optimization problems. Fig. 3 depicts flow chart of the proposed BHO algorithm.

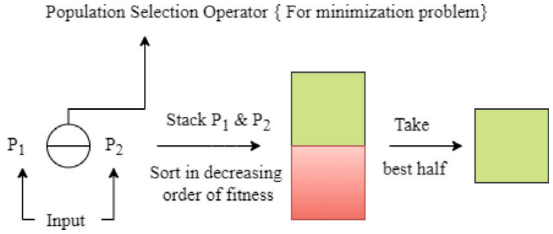


Fig. 4. Population selection operator.

BHO starts by taking in population size (N), number of groups (M), problem dimension (D), lower and upper bounds of solutions (LB and UB), maximum iterations (T) and other algorithm specific parameters as its input. Then it randomly initializes the population members in the limit of LB and UB and divides the population into M groups. Third step of BHO is Random replacement, where a group is chosen at random and best member of that group is replaced by a randomly generated solution using a greedy selection approach. This facilitates the exploration of new solutions.

Now in the fourth step of BHO, two copies of population are created (say P_1 and P_2), P_1 will undergo Henry's gas solubility optimization (HGSO) [60] while P_2 will undergo Brain storm optimization (BSO) [59]. In HGSO algorithm, α is a hyper parameter and it affects influence of other gases on ith gas of jth cluster [60] and in original implementation its value is taken as 1. We propose a constant called as amplification factor (AF) and it is used to increase α slowly in each iteration, which ultimately improves the convergence rate. α is initialized with a value of 1 and in each iteration, alpha is updated as $\alpha = \alpha * AF$, where $AF = 1 + \frac{0.1}{N}$

Fourth step can be parallelly processed to reduce time complexity. Now Population selection operator $\ominus$ is used to select N members out of P1 and P2. This step helps to improve the convergence rate of algorithm (see Fig. 4).

In the fifth step of BHO, population will be passed through worst agent handling module in which N_w worst solutions are selected and replaced with randomly generated solutions using greedy selection strategy. This step helps algorithm to escape local optima trap. Finally, after iterating for T times, best solution found so far is returned as an output of the algorithm.

4.2. Detection and mitigation phase

Firstly, results from each base learner are obtained. Then, using the weights obtained during the training phase, a weighted majority vote of the results obtained from each base learner is used as a final prediction to classify a new flow as either malicious or normal.

Final prediction = $majority[W_1.SVM_1, W_2.SVM_2, W_3.RF_1, W_4.RF_2, W_5.GBM_1, W_6.GBM_2]$

After deploying our model on POX controller, it monitors the traffic and as soon as the DDoS attack is detected, it blocks the victim port and prevent resources from further misuse.

5. Results and discussion

This section focuses on the experimental setup and performance evaluation of the proposed model for various dataset distributions. The proposed model is compared against base learners using metrics based on the confusion matrix. In the latter half of this section, the performance of the proposed dynamic fitness function is compared to that of static fitness function. Lastly, the proposed BHO optimizer is compared with HGSO and BSO optimizers.

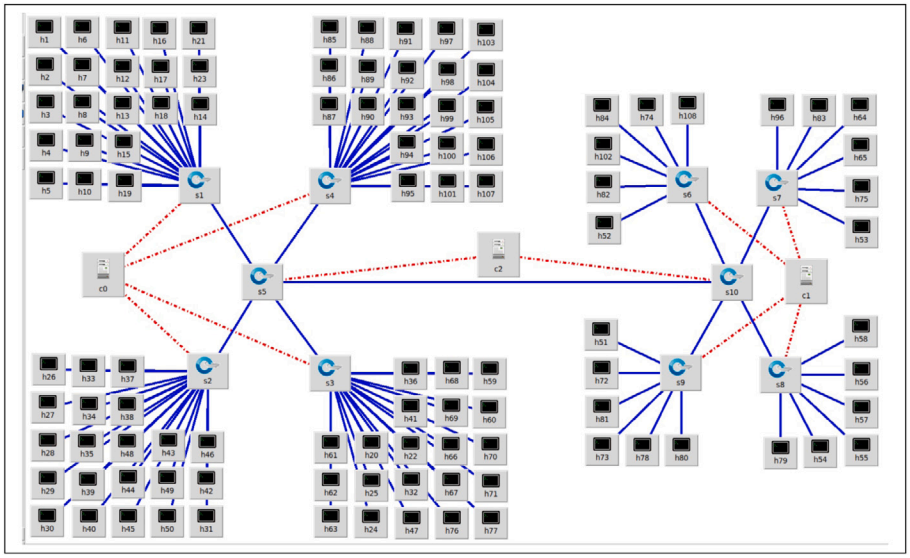


Fig. 5. Testbed.

5.1. Experimental setup

The experiments in this work are carried on a testbed created using mininet emulator and POX controller. Mininet is a network simulator that constructs a virtual network of hosts, switches, controllers, and links, while POX is an open source OpenFlow/Software Defined Networking (SDN) controller written in Python.

Fig. 5 shows our implemented testbed. It contains 108 hosts (h1-h108), 10 switches (s1 to s10), 3 controllers (c0-c2). 8 subnets are arranged in our testbed.

5.2. Performance evaluation

After creating the network setup, packets from CIC-DDoS2019 dataset were fed to the network. Fig. 6 shows the feature-wise-analysis of traffic on our SDN simulation.

To evaluate the performance of classifiers we have used the following metrics:

- Classification accuracy (Ac) = $\frac{TN+TP}{TP+FP+FN+TN}$
- Sensitivity/Recall (R) = $\frac{TP}{TP+FN}$
- Specificity (S) = $\frac{TN}{TN+FP}$
- Precision (P) = $\frac{TP}{TP+FP}$
- F-measure (F)= $2(\frac{PR}{P+R})$
- Detection rate (DR) = $100P$
- False alarm rate (FAR) = $100(\frac{FP}{FP+TN})$

where, TN is true negative, TP is true positive, FP is false positive and FN is false negative.

5.2.1. Comparison with base learners and majority voting ensemble (MVE)

Table 4 shows the comparison of base learners, MVE (majority voting ensemble) and the proposed OWVE (optimized weighted voting ensemble) in terms of above discussed performance metrics. Our proposed model is tested on CIC-DDoS2019 and CAIDA-2007 datasets and achieved an accuracy of 99.41% and 99.36% respectively.

SVM base learners performed poorly with a high false alarm rate (FAR). Random forest (RF) and Gradient boosted machines (GBM) performed much better than SVM base learners but still suffers from the problem of high FAR. MVE combines the strength of all base learners and outperformed each individual base learners in terms of classification accuracy as well as FAR. The proposed-OWVE model outperformed MVE and achieved very low FAR and high accuracy while maintaining a proper balance between Sensitivity and Specificity. The high performance of our proposed ensemble model on multiple datasets indicates the low variance of the framework.

Table 4
Base learners comparison.

Classifier	Ac (%)	R (%)	S (%)	F (%)	DR	FAR
SVM1	92.9609	96.3364	88.8576	93.7569	91.312	11.1424
SVM2	92.9648	96.3329	88.8706	93.7600	91.3209	11.1294
RF1	98.6802	99.7470	97.3834	98.8085	97.8876	2.61664
RF2	98.9279	99.1055	98.7117	99.0243	98.9433	1.28834
GBM1	98.9891	99.2766	98.6397	99.0806	98.8854	1.36031
GBM2	99.0751	99.3264	98.7697	99.1586	98.9913	1.23034
MVE	99.1201	99.3549	98.8346	99.1994	99.0443	1.16536
OWVE (CAIDA)	**99.3591**	**99.392**	**99.3221**	**99.4249**	**99.4595**	**0.6778**
OWVE (CIC-DDoS)	**99.4163**	**99.347**	**99.5022**	**99.4719**	**99.5971**	**0.4978**

Table 5
Performance of the proposed model.

Dataset distribution (CIC-DDoS)	Accuracy (%)
50–50	99.3896
60–40	99.3858
70–30	99.3432
80–20	99.4163
90–10	99.3155

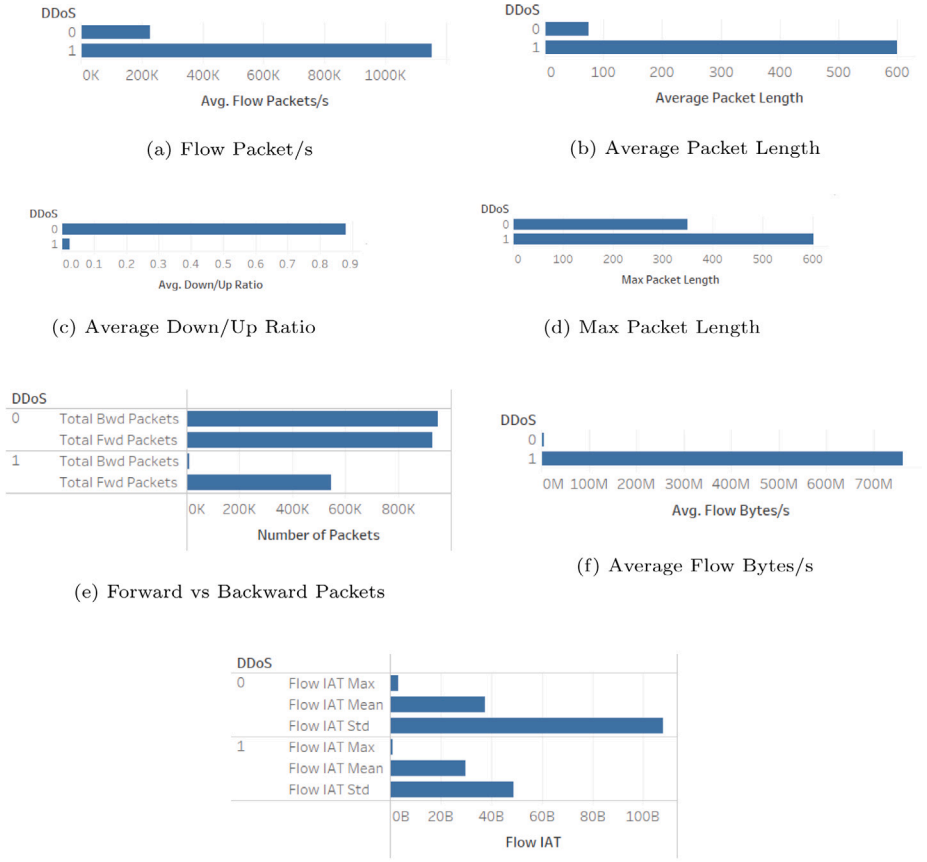
5.2.2. Dataset distribution

When it comes to assessing the effectiveness of a machine learning model, the distribution of dataset between training and testing is crucial. Table 5 shows how the performance of the proposed model varies with dataset distribution.

5.2.3. Comparison of proposed dynamic fitness function with static fitness function

The Figs. 7a, 7b shows the confusion matrix of the proposed model with static fitness function and dynamic fitness function for both CIC-DDoS2019, CAIDA-2007 datasets respectively.

The proposed design of the dynamic fitness function helps the model to increase the overall accuracy by decreasing the false negatives.



(a) Flow Packet/s

(b) Average Packet Length

(c) Average Down/Up Ratio

(d) Max Packet Length

(e) Forward vs Backward Packets

(f) Average Flow Bytes/s

(g) Packet Flow Inter Arrival Time

Fig. 6. Feature-wise analysis of traffic.

(a) Confusion Matrix using CIC-DDoS2019

	Positive	Negative
Positive	55.12	0.213
Negative	0.407	44.25

(a) Static fitness function

	Positive	Negative
Positive	55.97	0.222
Negative	0.361	44.44

(a) Proposed dynamic fitness function

(b) Confusion Matrix using CAIDA-2007

	Positive	Negative
Positive	54.49	0.287
Negative	0.383	43.85

(a) Static fitness function

	Positive	Negative
Positive	55.97	0.300
Negative	0.341	44.04

(a) Proposed dynamic fitness function

Fig. 7. Obtained confusion matrix (%).

Table 6
Comparison with previous studies.

Study	Year	Type	Method	Detection	Mitigation	Accuracy (%)
Myintoo et. al.	2019	ML	ASVM	Yes	No	97
Liu et. al.	2019	DL	Entropy + PSO-BP	Yes	No	97.7
Huseyin et. al.	2020	ML	Wrapped feature selection + KNN	Yes	No	98.3
Hannache et. al.	2020	DL	ANN + Traffic flow classifier	Yes	Yes	96.1
Makuvaza et. al.	2021	DL	DNN	Yes	No	97.6
Ahuja et. al.	2021	ML	SVM + RF	Yes	No	98.8
Ramprasath et. al.	2021	ML	Multinomial regression	Yes	Yes	98.7
Our proposed model	**2021**	**ML**	**OWVE**	**Yes**	**Yes**	**99.41**

With the static fitness function, the model achieved an accuracy of 99.384% and with dynamic fitness function the accuracy of the model is 99.416%.

5.2.4. Comparison with previous studies

The accuracy of the results obtained from our proposed model are compared to the accuracy of some recently proposed machine learning and deep learning based models (from 2019 to 2021) as shown in Table 6. A few of the models have been effective in enhancing levels of accuracy. In the field of machine learning based models, Ramprasath et al. [61] proposed a multinomial regression based approach to know the traffic alterations responsible for anomalies and achieved a notable

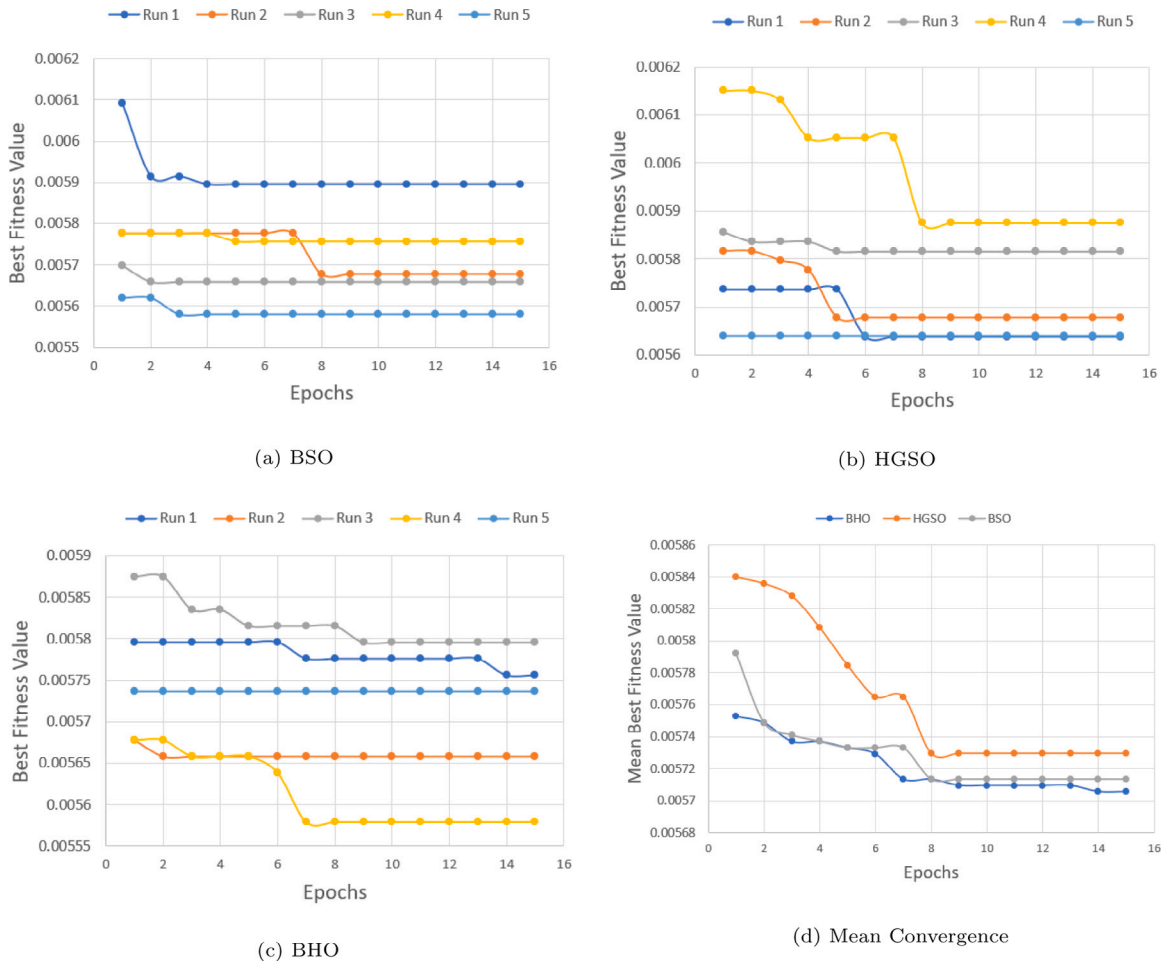


Fig. 8. Performance of the algorithms over 5 runs.

accuracy of 98.7% and Ahuja et al. [62] proposed a hybrid approach which utilizes support vector classifier followed by a random forest filtering and attained an accuracy of 98.8%. Coequally deep learning based models such as Hannache et al. [35] proposed an artificial neural network based traffic flow classifier and achieved an accuracy of 96.1% with mitigation capabilities.

Our proposed model, with an accuracy of 99.41% and 99.36% on CIC-DDoS, CAIDA datasets respectively, outperformed other shown recent models and has capabilities to detect and mitigate DDoS attacks on SDN with low variance.

5.2.5. Comparison of proposed BHO with HGSO and BSO

To compare the proposed hybrid metaheuristic optimizer (BHO) with the individual optimizers HGSO [60] and BSO [59], best fitness function value v/s iteration curve is drawn for 5 runs. Using the best fitness value attained in each run, a statistical table is generated which shows the performance of the algorithm over 5 runs (Table 7). Fig. 8: a, b and c shows the performance of BSO, HGSO and BHO respectively.

With the minimum best fitness value over 5 runs while maintaining a minimum mean, median, and standard deviation, BHO outperformed both BSO and HGSO. In the worst case scenario, BHO does not suffer too much loss and outperformed both BSO and HGSO.

The graph in Fig. 8d is a mean convergence curve (over 5 runs) which, at an iteration, takes mean of best fitness values achieved in 5 runs. As can be seen from the graph, in terms of best mean fitness value and convergence rate, BHO outperformed both HGSO and BSO.

Table 7
Statistical analysis to compare BHO with BSO and HGSO (derived using best fitness value over 5 runs).

Statistical quantity	Algorithm		
	BSO	HGSO	BHO
Best	0.005579323	0.005638468	0.005579323
Worst	0.006091911	0.006151056	0.005875047
Mean	0.005728370	0.005764140	0.005721530
Median	0.005697613	0.005737000	0.005737042
Standard deviation	0.000114556	0.000137538	7.9806E−05

6. Conclusion and future work

In this paper, we proposed a new approach to the detection and prevention of DDoS attacks in the environment of software defined network. The proposed optimized weighted voting ensemble model includes 6 base learners (2 SVMs, 2 RFs, 2 GBMs) differentiated by hyperparameters. To find an optimal set of weights for the proposed ensemble, a novel dynamic fitness function is proposed along with a novel hybrid metaheuristic optimization algorithm (BHO). The design of proposed dynamic fitness function helped the model to reduce false negatives and increasing overall accuracy. For training the ensemble framework we used CIC-DDoS2019 dataset. The model's efficiency is assessed on two datasets, CIC-DDoS2019 and CAIDA-2007 which are fed to a testbed created using mininet and POX controller. The proposed model achieves an accuracy of 99.4163% and 99.3591% respectively. High accuracy on multiple datasets shows low variance of the proposed ensemble framework.

In the future, we will explore novel ways of designing dynamic fitness functions for minimization problem and novel ways to create efficient hybrids that can combine strengths of existing metaheuristic optimization algorithms and use them in different domains to create efficient systems.

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.

References

[1] Internet growth usage statistics, 2019, https://www.clickz.com/internetgrowth-usage-stats-2019-time-online-devices-users/235102/.

[2] J. Singh, S. Behal, Detection and mitigation of DDoS attacks in SDN: A comprehensive review, research challenges and future directions, Comp. Sci. Rev. 37 (2020) 100279.

[3] T. Benson, A. Akella, D.A. Maltz, Unraveling the complexity of network management, in: NSDI, 2009, pp. 335–348.

[4] H. Uppal, D. Brandon, OpenFlow based load balancing, in: CSE561: Networking Project Report, University of Washington, 2010.

[5] D. Kreutz, F.M. Ramos, P.E. Verissimo, C.E. Rothenberg, S. Azodolmolky, S. Uhlig, Software-defined networking: A comprehensive survey, Proc. IEEE 103 (1) (2014) 14–76.

[6] S. Scott-Hayward, G. O'Callaghan, S. Sezer, SDN security: A survey, in: 2013 IEEE SDN for Future Networks and Services (SDN4FNS), IEEE, 2013, pp. 1–7.

[7] M. Feily, A. Shahrestani, S. Ramadass, A survey of botnet and botnet detection, in: 2009 Third International Conference on Emerging Security Information, Systems and Technologies, IEEE, 2009, pp. 268–273.

[8] M. Abu Rajab, J. Zarfoss, F. Monrose, A. Terzis, A multifaceted approach to understanding the botnet phenomenon, in: Proceedings of the 6th ACM SIGCOMM Conference on Internet Measurement, 2006, pp. 41–52.

[9] B. Saha, A. Gairola, Botnet: An Overivew. CERT-In White Paper, Tech. rep., CIWP-2005-05, 2005, June.

[10] Crippling cyber-attacks, 1998, https://www.bbc.com/news/technology35376327.

[11] S. Jajodia, K. Kant, P. Samarati, A. Singhal, V. Swarup, C. Wang, Secure Cloud Computing, Springer, 2014.

[12] P. Zhang, H. Wang, C. Hu, C. Lin, On denial of service attacks in software defined networks, IEEE Netw. 30 (6) (2016) 28–33.

[13] W. Xia, Y. Wen, C.H. Foh, D. Niyato, H. Xie, A survey on software-defined networking, IEEE Commun. Surv. Tutor. 17 (1) (2014) 27–51.

[14] S. Bu, F.R. Yu, X.P. Liu, H. Tang, Structural results for combined continuous user authentication and intrusion detection in high security mobile ad-hoc networks, IEEE Trans. Wireless Commun. 10 (9) (2011) 3064–3073.

[15] S. Sezer, S. Scott-Hayward, P.K. Chouhan, B. Fraser, D. Lake, J. Finnegan, N. Viljoen, M. Miller, N. Rao, Are we ready for SDN? Implementation challenges for software-defined networks, IEEE Commun. Mag. 51 (7) (2013) 36–43.

[16] M. Goossens, F. Mittelbach, A. Samarin, Open Networking Specifications 1.5.1, Vol. 3, Open Networking Foundation, 2015.

[17] OpenFlow switch, 2020, https://www.opennetworking.org/wp-content/uploads/2014/10/openflow-switch-v1.5.1.pdf. (Accessed on 11 March 2020).

[18] K. Giotis, C. Argyropoulos, G. Androulidakis, D. Kalogeras, V. Maglaris, Combining OpenFlow and sFlow for an effective and scalable anomaly detection and mitigation mechanism on SDN environments, Comput. Netw. 62 (2014) 122–136.

[19] S.-C. Tsai, I.-H. Liu, C.-T. Lu, C.-H. Chang, J.-S. Li, Defending cloud computing environment against the challenge of DDoS attacks based on software defined network, in: Advances in Intelligent Information Hiding and Multimedia Signal Processing, Springer, 2017, pp. 285–292.

[20] K. Kalkan, L. Altay, G. Gür, F. Alagöz, JESS: Joint entropy-based DDoS defense scheme in SDN, IEEE J. Sel. Areas Commun. 36 (10) (2018) 2358–2372.

[21] N.Z. Bawany, J.A. Shamsi, SEAL: SDN based secure and agile framework for protecting smart city applications from DDoS attacks, J. Netw. Comput. Appl. 145 (2019) 102381.

[22] J. Cui, M. Wang, Y. Luo, H. Zhong, DDoS detection and defense mechanism based on cognitive-inspired computing in SDN, Future Gener. Comput. Syst. 97 (2019) 275–283.

[23] N.M. AbdelAzim, S.F. Fahmy, M.A. Sobh, A.M.B. Eldin, A hybrid entropy-based DoS attacks detection system for software defined networks (SDN): A proposed trust mechanism, Egypt. Inform. J. 22 (1) (2021) 85–90.

[24] J. Li, Z. Zhao, R. Li, H. Zhang, Ai-based two-stage intrusion detection for software defined iot networks, IEEE Internet Things J. 6 (2) (2018) 2093–2102.

[25] T.V. Phan, M. Park, Efficient distributed denial-of-service attack defense in SDN-based cloud, IEEE Access 7 (2019) 18701–18714.

[26] M. Myint Oo, S. Kamolphiwong, T. Kamolphiwong, S. Vasupongayya, Advanced support vector machine-(ASVM-) based detection for distributed denial of service (DDoS) attack on software defined networking (SDN), J. Comput. Netw. Commun. 2019 (2019).

[27] H. Polat, O. Polat, A. Cetin, Detecting DDoS attacks in software-defined networks through feature selection methods and machine learning models, Sustainability 12 (3) (2020) 1035.

[28] B. Jia, X. Huang, R. Liu, Y. Ma, A DDoS attack detection method based on hybrid heterogeneous multiclassifier ensemble learning, J. Electr. Comput. Eng. 2017 (2017).

[29] N. Ahuja, G. Singal, D. Mukhopadhyay, N. Kumar, Automated DDOS attack detection in software defined networking, J. Netw. Comput. Appl. (2021) 103108.

[30] Y. Cui, L. Yan, S. Li, H. Xing, W. Pan, J. Zhu, X. Zheng, SD-Anti-DDoS: Fast and efficient DDoS defense in software-defined networks, J. Netw. Comput. Appl. 68 (2016) 65–79.

[31] C. Li, Y. Wu, X. Yuan, Z. Sun, W. Wang, X. Li, L. Gong, Detection and defense of DDoS attack–based on deep learning in OpenFlow-based SDN, Int. J. Commun. Syst. 31 (5) (2018) e3497.

[32] M.P. Novaes, L.F. Carvalho, J. Lloret, M.L. Proença, Long short-term memory and fuzzy logic for anomaly detection and mitigation in software-defined network environment, IEEE Access 8 (2020) 83765–83781.

[33] F. Gharvirian, A. Bohlooli, Neural network based protection of software defined network controller against distributed denial of service attacks, Int. J. Eng. 30 (11) (2017) 1714–1722.

[34] Z. Liu, Y. He, W. Wang, B. Zhang, DDoS attack detection scheme based on entropy and PSO-BP neural network in SDN, China Commun. 16 (7) (2019) 144–155.

[35] O. Hannache, M.C. Batouche, Neural network-based approach for detection and mitigation of DDoS attacks in SDN environments, Int. J. Inf. Secur. Priv. (IJISP) 14 (3) (2020) 50–71.

[36] B. Wang, Y. Zheng, W. Lou, Y.T. Hou, DDoS attack protection in the era of cloud computing and software-defined networking, Comput. Netw. 81 (2015) 308–319.

[37] K. Bhushan, B.B. Gupta, Distributed denial of service (DDoS) attack mitigation in software defined network (SDN)-based cloud computing environment, J. Ambient Intell. Humaniz. Comput. 10 (5) (2019) 1985–1997.

[38] A. AlEroud, I. Alsmadi, Identifying cyber-attacks on software defined networks: An inference-based intrusion detection approach, J. Netw. Comput. Appl. 80 (2017) 152–164.

[39] J. Wang, R. Wen, J. Li, F. Yan, B. Zhao, F. Yu, Detecting and mitigating target link-flooding attacks using sdn, IEEE Trans. Dependable Secure Comput. 16 (6) (2018) 944–956.

[40] R. Burduk, Recognition task with feature selection and weighted majority voting based on interval-valued fuzzy sets, in: N.-T. Nguyen, K. Hoang, P. Jędrzejowicz (Eds.), Computational Collective Intelligence. Technologies and Applications, Springer Berlin Heidelberg, Berlin, Heidelberg, 2012, pp. 204–209.

[41] A. Rojarath, W. Songpan, Cost-sensitive probability for weighted voting in an ensemble model for multi-class classification problems, Appl. Intell. (2021) 1–25.

[42] J. Large, J. Lines, A. Bagnall, A probabilistic classifier ensemble weighting scheme based on cross-validated accuracy estimates, Data Min. Knowl. Discov. 33 (6) (2019) 1674–1709.

[43] L.I. Kuncheva, J.J. Rodríguez, A weighted voting framework for classifiers ensembles, Knowl. Inf. Syst. 38 (2) (2014) 259–275.

[44] Y. Zhang, H. Zhang, J. Cai, B. Yang, A weighted voting classifier based on differential evolution, in: Abstract and Applied Analysis, Vol. 2014, Hindawi, 2014.

[45] A. Onan, S. Korukoğlu, H. Bulut, A multiobjective weighted voting ensemble classifier based on differential evolution algorithm for text sentiment classification, Expert Syst. Appl. 62 (2016) 1–16.

[46] C. Cortes, V. Vapnik, Support-vector networks, Mach. Learn. 20 (3) (1995) 273–297, http://dx.doi.org/10.1007/BF00994018.

[47] B.E. Boser, I.M. Guyon, V. Vapnik, A Training Algorithm for Optimal Margin Classifiers, ACM, New York, 1992, pp. 144–152.

[48] J.R. Quinlan, Simplifying decision trees, Int. J. Man-Mach. Stud. 27 (3) (1987) 221–234.

[49] J. Friedman, T. Hastie, R. Tibshirani, et al., The Elements of Statistical Learning, Vol. 1, Springer series in statistics New York, 2001.

[50] L. Breiman, Random forests, Mach. Learn. 45 (1) (2001) 5–32, http://dx.doi.org/10.1023/A:1010933404324.

[51] L. Breiman, Bagging predictors, Mach. Learn. 24 (2) (1996) 123–140.

[52] Y. Amit, D. Geman, K. Wilder, Joint induction of shape features and tree classifiers, IEEE Trans. Pattern Anal. Mach. Intell. 19 (11) (1997) 1300–1305.

[53] J. Friedman, T. Hastie, R. Tibshirani, et al., Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors), Ann. Statist. 28 (2) (2000) 337–407.

[54] J.H. Friedman, Greedy function approximation: a gradient boosting machine, Ann. Statist. (2001) 1189–1232.

[55] J.H. Friedman, Stochastic gradient boosting, Comput. Statist. Data Anal. 38 (4) (2002) 367–378.

[56] I. Sharafaldin, A.H. Lashkari, S. Hakak, A.A. Ghorbani, Developing realistic distributed denial of service (DDoS) attack dataset and taxonomy, in: 2019 International Carnahan Conference on Security Technology (ICCST), IEEE, 2019, pp. 1–8.
[57] CAIDA dataset, 2007, https://www.caida.org/catalog/datasets/ddos-20070804_dataset.
[58] N. Littlestone, M. Warmuth, The weighted majority algorithm, in: 30th Annual Symposium on the Foundations of Computer Science, 1994.
[59] Y. Shi, Brain storm optimization algorithm, in: International Conference in Swarm Intelligence, Springer, 2011, pp. 303–309.
[60] F.A. Hashim, E.H. Houssein, M.S. Mabrouk, W. Al-Atabany, S. Mirjalili, Henry gas solubility optimization: A novel physics-based algorithm, Future Gener. Comput. Syst. 101 (2019) 646–667.
[61] A. Makuvaza, D.S. Jat, A.M. Gamundani, Deep neural network (DNN) solution for real-time detection of distributed denial of service (DDoS) attacks in software defined networks (SDNs), SN Comput. Sci. 2 (2) (2021) 1–10.
[62] J. Ramprasath, V. Seethalakshmi, Improved network monitoring using software-defined networking for DDoS detection and mitigation evaluation, Wirel. Pers. Commun. 116 (3) (2021) 2743–2757.



Aastha Maheshwari obtained her Master's degree from NIT Hamirpur and currently pursuing Ph.D. in Computer Science and Engineering Department from Delhi Technological University, Delhi. Her specializations Internet of Things, sensor network and wireless communication.
Email: aasthamaheshwari.am@gmail.com



Burhan Mehraj is currently pursuing a Bachelors' degree in Computer Science and Engineering from Delhi Technological University, Delhi. His interests are in Machine Learning, Deep learning, Metaheuristic optimization, data mining.
Email: burhanmrj@gmail.com



Mohd Shaad Khan is currently pursuing a Bachelors' degree in Computer Science and Engineering from Delhi Technological University, Delhi. His interests are in Machine Learning, Deep learning, Metaheuristic optimization, data mining.
Email: shaadk7865@gmail.com



Mohd Shaheem Idrisi is currently pursuing a Bachelors' degree in Computer Science and Engineering from Delhi Technological University, Delhi. His interests are in Machine Learning, Web development and App Development.
Email: mohdshaheem91@gmail.com

Thai Journal of **Math**ematics
Volume 19 Number 4 (2021)
Pages 1187–1197

http://thaijmath.in.cmu.ac.th

ISSN 1686-0209

Approximation of Jain Operators by Statistical Convergence

Nurhayat İspir[1], **Naokant Deo**[2] **and Neha Bhardwaj**[3,*]

[1] *Gazi University, Sciences & Arts Faculty Department of Mathematics, 06500 Teknikokullar/Ankara, Turkey*
e-mail : nispir@gazi.edu.tr

[2] *Department of Applied Mathematics, Delhi Technological University*
(Formerly Delhi College of Engineering), Bawana Road, Delhi 110042, India
e-mail : naokantdeo@dce.ac.in

[3] *Department of Mathematics, School of Basic Sciences and Research*
Sharda University, Greater Noida 201310, India.
e-mail : neha_bhr@yahoo.co.in

Abstract In this paper, we consider a positive linear operators $P_n^{[\beta]}$ introduced by Jain [G.C. Jain, Approximation of functions by a new class of linear positive operators, Jour. Austral. Math. Soc. 13 (3) (1972) 271–276] with the help of Poisson type distribution and study the Voronovskaya type result of the operator then obtain an error estimate in terms of the higher order modulus of continuity of the function being approximated and its A-statistical convergence. We also compute the corresponding rate of A-statistical convergence for these operators.

MSC: 41A10; 41A25; 41A30; 41A63; 26A15
Keywords: Poisson distribution; Voronovskaya; A-statistical convergence

Submission date: 18.02.2017 / Acceptance date: 25.10.2021

1. Introduction and Definitions

Several extension and generalization of Bernstein polynomials have been given by various mathematician like Szâsz [1], Meyer-König and Zeller [2], Meir and Sharma [3], Stancu [4] and Balázs [5]. In [6], Kajla and Acar have studied the blending type approximation properties by generalized Bernstein-Durrmeyer type operators. Deo and Bhardwaj [7] considered the multidimensional Bernstein operators and its Durrmeyer variants on simplex. Mirakyan [8] has also given another modification with the help of the Poisson distribution.

Later on in the same way with the help of a Poisson type distribution,

$$w_\beta(k,\alpha) = \frac{\alpha}{k!}(\alpha+k\beta)^{k-1}e^{-(\alpha+k\beta)},\ k \in N_0 = \{0\} \cup \mathbb{N},$$

*Corresponding author.

Published by The Mathematical Association of Thailand.
Copyright © 2021 by TJM. All rights reserved.

for $0 < \alpha < \infty$ and $|\beta| < 1$, Jain [9] defined the following class of positive linear operators,

$$P_n^{[\beta]}(f;x) = \sum_{k=0}^{\infty} w_\beta(k;nx)\, f\left(\frac{k}{n}\right),\ x \geq 0, \tag{1.1}$$

where $\beta \in [0,1)$ and $f \in C(\mathbb{R}_+)$, the space of all real valued continuous functions defined on $\mathbb{R}_+ = [0,\infty)$. Notice that the parameter β may depend on the natural number n. Original Szâsz-Mirakyan operator can easily be obtained for $\beta = 0$. Deo et al. [10, 11] studied another modification of Bernstein operators and Gupta [12] introduced q analogue of Bernstein operator. For better approximation of well-known Szâsz-Mirakyan operators [1], Jain and Pethe [13] generalized the operators as:

$$S_n^{[\gamma]}(f;x) = (1+n\gamma)^{-x/\gamma} \sum_{k=0}^{\infty} \left(\gamma + \frac{1}{n}\right)^{-k} \frac{x^{(k,-\gamma)}}{k!} f\left(\frac{k}{n}\right),$$

where $x^{(k,-\gamma)} = x(x+\gamma)\ldots(x+(k-1)\gamma), x^{(0,-\gamma)} = 1$ and f is any function of exponential type such that $|f(t)| \leq Ke^{Mt}\ (t \geq 0)$ for some finite constants $K, M > 0$ and $\gamma = (\gamma_n)_{n\in\mathbb{N}}$.

Very recently, Bodur [14] introduced a new modification of modified Lupas-Jain operators based on function ϕ having some properties. Various researchers worked on Jain's operators, one can refer ([15], [16], [17]).

Now the following lemmas follow from [9], for the operators $P_n^{[\beta]}$ mentioned by (1.1).

Lemma 1.1 ([9]). *Let $e_i(x) = x^i$, $i = 0,1,2$, then for fixed $x \in [0,\infty)$, $n \in \mathbb{N}$ and $\beta \in [0,1)$, we have*

(*i*) $P_n^{[\beta]}(e_0;x) = 1$,
(*ii*) $P_n^{[\beta]}(e_1;x) = \frac{1}{1-\beta}x$,
(*iii*) $P_n^{[\beta]}(e_2;x) = \frac{1}{(1-\beta)^2}x^2 + \frac{1}{n(1-\beta)^3}x$.

Lemma 1.2. *For $x \in [0,\infty)$, $n \in \mathbb{N}$, $\beta \in [0,1)$ and $\varphi_x(t) = t - x$, we have*

(*i*) $P_n^{[\beta]}(\varphi_x;x) = \frac{\beta}{1-\beta}x$,
(*ii*) $P_n^{[\beta]}(\varphi_x^2;x) = \frac{\beta^2}{(1-\beta)^2}x^2 + \frac{1}{n(1-\beta)^3}x$.

This paper is organised in four sections. The first section is devoted to definitions and auxiliary properties of operators $P_n^{[\beta]}$ given by (1.1). In the second section, we obtain the Voronovskaya type result of the operator and an error estimate of the function being approximated. In the third section, we give some definitions and properties of statistical and A-statistical convergence followed by statistical approximation theorems. In the last section, we compute the corresponding A-statistical rates of approximation and its order.

2. Voronovskaya Type Results & Error Estimation

In this section we compute the Voronovskaya type results of these operators $P_n^{[\beta]}$ given by (1.1).

Let $f \in C_B[0,\infty)$ be the space of all real valued functions with bounded and uniformly continuous on $[0,\infty)$, equipped with the norm $\|f\| = \sup_{x\in[0,\infty)} |f(x)|$. The Peetre's K_2-functional is defined by

$$K_2 = \inf\left\{\|f-g\| + \delta\|g''\| : g \in C_B^2[0,\infty)\right\}, \quad \delta > 0$$

where $C_B^2[0,\infty) = \{g \in C_B[0,\infty) : g', g'' \in C_B[0,\infty)\}$ and the norm

$$\|f\|_{C_B^2[0,\infty)} = \|f\|_{C_B} + \|f'\|_{C_B} + \|f''\|_{C_B}. \tag{2.1}$$

From [18], there exists a positive constant C such that

$$K_2(f,\delta) \le C\omega_2\left(f,\sqrt{\delta}\right) \tag{2.2}$$

and

$$\omega_2(f,\sqrt{\delta}) = \sup_{0<h\le\delta}\sup_{x\in[0,\infty)} |f(x+2h) - 2f(x+h) + f(x)|.$$

Theorem 2.1. *Let $f \in C_B[0,\infty)$, and $\beta \to 0$ as $n \to \infty$, then for every $x \in [0,\infty)$, $\beta \in [0,1)$ and for $C > 0$, we have*

$$\left|P_n^{[\beta]}(f;x) - f(x)\right| \le C\omega_2\left(f, \sqrt{\frac{\beta}{1-\beta}x}\right).$$

Proof. Let $g \in C_B^2$. Using Taylor's expansion

$$g(t) = g(x) + g'(x)(t-x) + \int_x^t (t-u)g''(u)du.$$

From Lemma 1.2, we have

$$P_n^{[\beta]}(g;x) - g(x) = P_n^{[\beta]}\left(g'(x)(t-x);x\right) + P_n^{[\beta]}\left(\int_x^t (t-u)g''(u)du;x\right).$$

We know that

$$\left|\int_x^t (t-u)g''(u)du\right| \le (t-u)^2\|g''\|.$$

Therefore

$$\begin{aligned}
\left|P_n^{[\beta]}(g;x) - g(x)\right| &\le P^{[\beta]}\left((t-x);x\right)\|g'\| + P_n^{[\beta]}\left((t-u)^2;x\right)\|g''\| \\
&= \frac{\beta}{1-\beta}x\|g'\| + \left(\frac{\beta^2x^2}{(1-\beta)^2} + \frac{x}{n(1-\beta)^3}\right)\|g''\| \\
&\le \frac{\beta}{1-\beta}x\{\|g'\| + \|g''\|\} \\
&\le \frac{\beta}{1-\beta}x\|g''\|.
\end{aligned}$$

By Lemma 1.1, we have

$$\left|P_n^{[\beta]}(g;x)\right| \le \sum_{k=0}^{\infty}\frac{\alpha}{k!}(\alpha+k\beta)^{k-1}e^{-(\alpha+k\beta)}g\left(\frac{k}{n}\right) \le \|g\|.$$

Hence

$$\left|P_n^{[\beta]}(g;x)-g(x)\right| \leq \left|P_n^{[\beta]}((g-f);x)-(g-f)(x)\right| + \left|P_n^{[\beta]}(g;x)-g(x)\right|$$
$$\leq 2\|g-f\| + \left(\frac{\beta}{1-\beta}x\right)\|g''\|.$$

Taking the infimum on the right side over all $g \in C_B^2$ and using (2.2), we get the required result. ∎

3. A-Statistical Convergence

In this section of the paper, we use concept of statistical convergence and study the Korovkin type approximation theorem for the operators $P_n^{[\beta]}$. Before we present the main results, we shall recall some definitions and results on the statistical and A-statistical convergence. In the year 1951, Fast [19] studied the definition of statistical convergence for sequences of real numbers as:

Let K be a subset of natural numbers $\mathbb{N}$ and $K_j = \{n \leq j : n \in \mathbb{N}\}$. If limit exists, then the natural density of K is defined by

$$\delta(K) = \lim_j \frac{1}{j}|K_j|,$$

where vertical bars indicating the number of elements in the enclosed set. A sequence $(x_n)_n$ is said to be statistically convergent to a number L, if for every $\varepsilon > 0$, the natural density of the set

$$K_\varepsilon = \{n \in \mathbb{N} : |x_n - L| \geq \varepsilon\}$$

is zero, i.e.,

$$\lim_n \frac{1}{n}|\{n \leq j : |x_n - L| \geq \varepsilon\}| = 0.$$

We denote this statistical limit by $st - \lim_n x_n = L$ (see [19–22]).

Let $A = (a_{jn})$, $j, n = 1, 2, ...,$ be a non-negative regular summability matrix. A sequence $(x_n)_n$ is said to be $A-$statistically convergent to a number L, if for every $\varepsilon > 0$,

$$\lim_j \sum_{n:|x_n-L|\geq\varepsilon} a_{jn} = 0,$$

holds. We denote this limit by $st_A - lim_n x_n = L$. We note that by taking $A = C_1$, the Cesàro matrix, $A-$statistical convergence reduce to ordinary convergence. Kolk [21] studied $A-$statistical convergence is stronger than ordinary convergence.

$A-$statistical convergence may also be given in normed spaces. Suppose $(X, \|.\|)$ is a normed space and $v = (v_k)$ is a sequence. This sequence is $A-$statistically convergent to $v_0 \in X$, if for every $\varepsilon > 0, \delta_A\{k \in \mathbb{N} : \|v_k - v_0\| \geq \varepsilon\} = 0$ (see [23, 24]).

Now $A = [a_{jn}]$ $(j, n \in \mathbb{N})$ is a non-negative regular summability matrix. Assume that for each $t \in [0, \infty)$, $(\alpha_n^*(t))_{n\in\mathbb{N}}$ is a sequence in $[0, \infty)$ satisfying

$$st_A - \lim_n \alpha_n^*(t) = t$$

then we have

$$st_A - \lim_n (t - \alpha_n^*(t)) = 0.$$

Gadjiev and Orhan [25] have explored a new research area, the use of statistical convergence in approximation theory and they investigated the Korovkin type approximation theorems via statistical convergence. Later on Korovkin type approximation theorems are introduced for A-statistical convergence by Duman et al. [26]. İspir and Gupta [27] studied approximation properties and estimate the rates of $A-$statistical convergence of Kantorovich variant of generalized Bernstein type operators. Using A-statistical and statistical convergence some approximating operators are studied by several mathematicians [28, 29].

In this section $A = [a_{jn}]$ is a non-negative regular summability matrix and β, given in definition of (1.1), will denote the sequence $\beta := \{\beta_n\}$ such that

$$st_A - \lim_n \beta_n = 0 \tag{3.1}$$

with $\beta_n \in [0, 1)$. Notice that, the present condition (3.1) includes the given condition for β of Theorem (2.1) in [9], its converse is not true. For example, if we define a sequence (β_n) such that

$\beta_n := \begin{cases} \sqrt{n}, & \text{if } n \text{ is a square,} \\ n & \text{otherwise} \end{cases}$ then observe that (3.1) holds true with the choice of $A = C_1$, the Cesaro matrix; but β_n is a non-convergent sequence in the usual sense.

Theorem 3.1. *Let $A = [a_{jn}]$ is a non-negative regular summability matrix and let $\beta := \{\beta_n\}$ be a sequence satisfying the condition (3.1) with $\beta_n \in [0, 1)$ for all $n \in \mathbb{N}$. Then, for every $f \in C[0, \infty)$ and for each compact interval $[0, b] \subset [0, \infty)$ uniformly*

$$st_A - \lim_n \left| P_n^{[\beta]}(e_i; x) - e_i(x) \right| = 0; \ e_i(t) = t^i, i = 0, 1, 2.$$

Proof. From Lemma 1.1, obviously $st_A - \lim_n \left| P_n^{[\beta]}(e_0; x) - 1 \right| = 0.$

$$\left| P_n^{[\beta]}(e_1; x) - x \right| = \left| \frac{\beta_n x}{1 - \beta_n} \right| \leq b \frac{\beta_n}{1 - \beta_n}, \quad x \in [0, b] \subset [0, \infty), \tag{3.2}$$

and from hypothesis we get $st_A - \lim_n \frac{\beta_n}{1-\beta_n} = 0.$
Now, for a given $\varepsilon > 0$, we define

$$V = \left\{ n : \frac{\beta_n}{1 - \beta_n} \geq \frac{\varepsilon}{b} \right\},$$

Therefore, by (3.2), we obtain

$$\sum_{n: \left| P_n^{[\beta]}(e_1, x) - x \right| \geq \varepsilon} a_{jn} \leq \sum_{n \in V} a_{jn}.$$

Taking limit as $j \to \infty$ and from the above inequality we get the result.
Similarly, for each $x \in [0,b] \subset [0,\infty)$

$$\begin{aligned}\left|P_n^{[\beta]}(e_2;x)-x^2\right| &= \left|\frac{x^2}{(1-\beta_n)^2}+\frac{x}{n(1-\beta_n)^3}-x^2\right| \\ &= \left|\frac{x^2\beta_n(2-\beta_n)}{(1-\beta_n)^2}+\frac{x}{n(1-\beta_n)^3}\right| \\ &\leq b^2\frac{\beta_n(2-\beta_n)}{(1-\beta_n)^2}+b\frac{1}{n(1-\beta_n)^3} \\ &\leq A\left\{\frac{\beta_n(2-\beta_n)}{(1-\beta_n)^2}+\frac{1}{n(1-\beta_n)^3}\right\}, \end{aligned} \tag{3.3}$$

where $A=\max\left\{b^2,b\right\}$. Now, for a given $\varepsilon>0$, define

$$T=\left\{n:\frac{\beta_n(2-\beta_n)}{(1-\beta_n)^2}+\frac{1}{n(1-\beta_n)^3}\geq\frac{\varepsilon}{A}\right\},$$

$$T_1=\left\{n:\frac{\beta_n(2-\beta_n)}{(1-\beta_n)^2}\geq\frac{\varepsilon}{2b^2}\right\}$$

$$T_2=\left\{n:\frac{1}{n(1-\beta_n)^3}\geq\frac{\varepsilon}{2b}\right\}.$$

It is easy to see that $T\subseteq T_1\cup T_2$. Since $st_A-\lim_n\beta_n=0$ we get $st_A-\lim_n\frac{\beta_n(2-\beta_n)}{(1-\beta_n)^2}$, $st_A-\lim_n\frac{1}{n(1-\beta_n)^3}$. Therefore by (3.3), we obtain

$$\sum_{n:\left|P_n^{[\beta]}(e_2,x)-x^2\right|\geq\varepsilon} a_{jn}\leq\sum_{n\in T}a_{jn}\leq\sum_{n\in T_1}a_{jn}+\sum_{n\in T_2}a_{jn}.$$

Taking limit as $j\to\infty$, we reach the result. ∎

Similarly, from Lemma 1.1 and by $st_A-\lim_n\beta_n=0$ we have for each compact interval $[0,b]\subset[0,\infty)$

$$st_A-\lim_n\left\|P_n^{[\beta]}(e_1-xe_0)^j\right\|_{C[0,b]}=0,\ \ j=1,2.$$

Now we give a Korovkin type theorem for the operators $P_n^{[\beta]}(f;x)$ via A-statistical convergence.

Theorem 3.2. *Let $A=[a_{jn}]\,(j,n\in N)$ is a non-negative regular summability matrix and let us suppose that $\beta:=\{\beta_n\}$ is a sequence satisfying the condition (3.1) with $\beta_n\in[0,1)$ for all $n\in N$. Then, for each $x\in[0,b]\subset[0,\infty)$ and for every $f\in C[0,\infty)$, we have*

$$st_A-\lim_n\left\|P_n^{[\beta]}(f;x)-f(x)\right\|_{C[0,b]}=0.$$

Proof. The result follows by Theorem 3.1 and from Theorem 1 in [25] (see also [26]). We recall that Theorem 1 in [25] is given for statistical convergence but the proof also works for A-statistical convergence. ∎

If we take $A = I$, identity matrix, then we have Theorem 2.1 in [9].

Remark. *We now present an example of a sequence of positive linear operators satisfying the conditions of Theorem 3.2 but that does not satisfy the conditions of Theorem 2.1 of* [9].

Assume now that $\{u_n\}$ *is an A-statistically null sequence but not convergent. Notice that, if* $A = (a_{jn})$ *is a non-negative regular matrix such that* $\lim_j \max_n \{a_{jn}\} = 0$, *then A-statistical convergence is stronger than convergence* [21]. *Without loss of generality we may assume that* $\{u_n\}$ *is a non-negative; otherwise we would replace* $\{u_n\}$ *by* $\{|u_n|\}$. *Now define* $\{L_n\}$ *on* $C[0,\infty)$ *by*

$$L_n(f;x) = (1+u_n)P_n^{[\beta]}(f;x),$$

where $\left\{P_n^{[\beta]}\right\}$ *is the sequence of the operators given with* (1.1). *Now observe that* $\left\{P_n^{[\beta]}\right\}$ *being convergent and* $\{u_n\}$ *being A-statistical null, their product will also be A-statistical null. Hence* $\{L_n\}$ *will not be convergent to* f *but will be A-statistically convergent to* f.

4. Rate of A-Statistical Convergence

Now we give the rate of A-statistical convergence for the operators $P_n^{[\beta]}(f;x)$ by using the Peetre's K-functional in the space $C_B^2[0,\infty)$.

Theorem 4.1. *Let* $A = [a_{jn}]\,(j,n \in N)$ *is a non-negative regular summability matrix and let* $\beta := \{\beta_n\}$ *is a sequence satisfying the condition* (3.1) *with* $\beta_n \in [0,1)$ *for all* $n \in N$. *For each* $f \in C_B[0,\infty)$

$$\left\|P_n^{[\beta]}(f;x) - f(x)\right\|_{C_B} \le \kappa\left(f;\gamma_{n,x}\right),$$

where $\kappa(f;\gamma_{n,x})$ *is the sequence of Peetre's K-functional and*

$$\gamma_{n,x} = \left\|P_n^{[\beta]}((e_1-x);x)\right\|_{C_B} + \left\|P_n^{[\beta]}((e_1-x)^2;x)\right\|_{C_B}$$

and $st_A - \lim_n \gamma_{n,x} = 0$ *for each fixed* $x \in [0,\infty)$.

Proof. Applying Taylor expansion to the function $f \in C_B^2[0,\infty)$, we get

$$P_n^{[\beta]}(f,x) - f(x) = f'(x)P_n^{[\beta]}((e_1-x);x) + \frac{1}{2}f''(\xi)P_n^{[\beta]}((e_1-x)^2;x), \quad \xi \in (t,x).$$

Hence

$$\begin{aligned}&\left\|P_n^{[\beta]}(f;x) - f(x)\right\|_{C_B}\\ &\le \|f'\|_{C_B}\left\|P_n^{[\beta]}((e_1-x);x)\right\|_{C[0,\infty)} + \|f''\|_{C_B}\left\|P_n^{[\beta]}((e_1-x)^2;x)\right\|_{C[0,\infty)}. \qquad (4.1)\end{aligned}$$

Using (2.1) and (4.1), for each $g \in C_B^2[0,\infty)$,

$$\begin{aligned}\left\|P_n^{[\beta]}(g;x)-g(x)\right\|_{C_B^2} &= \left(\left\|P_n^{[\beta]}((e_1-x);x)\right\|_{C[0,\infty)} + \left\|P_n^{[\beta]}((e_1-x)^2;x)\right\|_{C[0,\infty)}\right)\|g\|_{C_B^2}\\ &= \gamma_{n,x}\|g\|_{C_B^2}.\end{aligned}$$

For each $f \in C_B[0,\infty)$ and $g \in C_B^2[0,\infty)$,

$$\begin{aligned}\left\|P_n^{[\beta]}(f;x)-f(x)\right\|_{C_B^2} &\leq \left\|P_n^{[\beta]}(f;x)-P_n^{[\beta]}(g;x)\right\|_{C_B}+\left\|P_n^{[\beta]}(g;x)-g(x)\right\|_{C_B^2}+\|g-f\|_{C_B}\\ &\leq 2\|g-f\|_{C_B}+\left\|P_n^{[\beta]}(g;x)-g(x)\right\|_{C_B^2}\\ &\leq 2\|g-f\|_{C_B}+\gamma_{n,x}\|g\|_{C_B^2} \leq 2\left(\|g-f\|_{C_B}+\gamma_{n,x}\|g\|_{C_B^2}\right).\end{aligned}$$

Taking infimum over $g \in C_B^2[0,\infty)$, we get

$$\left\|P_n^{[\beta]}(f,x)-f(x)\right\|_{C_B^2} \leq \kappa\left(f;\gamma_{n,x}\right).$$

Since $st_A - \lim_n \beta_n = 0$, with $\beta_n \in [0,1)$,
we get $st_A - \lim_n \gamma_{n,x} = 0$, therefore $st_A - \lim_n \kappa(f;\gamma_{n,x}) = 0$. ∎

Now following [26], we would like to find the order of A-statistical approximation for the sequence of operators $\left\{P_n^{[\beta]}\right\}$ given by (1.1).

We recall some definitions and notations. Let $A = (a_{nk})$ is a non-negative regular summability matrix and let (b_j) be positive non-increasing sequence. A sequence $x = (x_n)$ is called $A-$statistical convergent to the number L with the rate of $o(b_j)$ if for every $\varepsilon > 0$, $\lim_j \frac{1}{b_j} \sum_{n:\,|x_n-L|\geq\varepsilon} a_{jn} = 0$. In this case we write $x_n - L = st_A - o(b_n)$, (as $n \to \infty$). If for every $\varepsilon > 0$, $\sup_j \frac{1}{b_j}\sum_{n:|x_n|\geq\varepsilon} a_{jn} < \infty$, then x is $A-$statistically bounded with the rate of $O(b_n)$ and it is denoted by $x_n = st_A - O(b_n)$, as $n \to \infty$, [26].

In the above two definitions the "rate" is controlled more by the entries of the summability method than by the terms of the sequence $x = (x_n)$. For example if $A = (a_{jn}) = I$ the identity matrix and if $a_{nn} = o(b_n)$ then $x_n - L = st_A - o(b_n)$ for any convergent sequence $(x_n - L)$ regardless of how slowly it goes to zero. To avoid such an unfortunate situation, considering the concept of convergence in measure from measure theory, the following two extra definitions are introduced in [26].

The sequence $x = (x_n)$ is $A-$statistically convergent to L with the rate of $o_\mu(b_n)$, denote by $x_n - L = st_A - o_\mu(b_n)$, (as $n \to \infty$), if for every $\varepsilon > 0$, $\lim_j \sum_{n:\,|x_n-L|\geq\varepsilon b_n} a_{jn} = 0$. Finally, the sequence $x = (x_n)$ is $A-$statistically bounded with the rate of $O\mu(b_n)$ provided that there is a positive number M such that $\lim_j \sum_{n:\,|x_n|\geq M\ b_n} a_{jn} = 0$. In this case we write $x_n = st_A - O\mu(b_n)$, as $n \to \infty$. In this case we write

$$x_k - L = st_A - o(a_n), \quad (\text{as } k \to \infty).$$

Theorem 4.2. *Let $A = (a_{jn})$ be a non-negative regular summability matrix and $(b_n(x))$ non-increasing sequence. If the sequence of positive linear operators $\left\{P_n^{[\beta]}\right\}$ defined by (1.1) and*

$$\omega(f;\delta_{n,x}) = st_A - o(b_n(x)) \text{ with } \delta_{n,x} = \sqrt{P_n^{[\beta]}((e_1-x)^2;x)} \tag{4.2}$$

then

$$P_n^{[\beta]}(f;x) - f(x) = st_A - o(b_n(x)),$$

where $\omega(f, \delta_n(x))$ *is the usual modulus of continuity of the function* f *for fixed* $x \geq 0$. *Similar results hold with little "o" replaced by big "O".*

Proof. Since $P_n^{[\beta]}(e_0;x) = 1 = e_0(x)$ using Cauchy-Schwarz inequality we can write

$$\left|P_n^{[\beta]}(f;x) - f(x)\right| \leq \left[P_n^{[\beta]}(e_0;x) + \frac{1}{\delta_{n,x}}\left(P_n^{[\beta]}((e_1{-}x)^2;x)\right)^{1/2}\right]\omega(f,\delta_{n,x}).$$

Hence with choosing $\delta_{n,x} = \sqrt{L_n((e_1{-}x)^2;x)}$ we obtain

$$\left|P_n^{[\beta]}(f;x) - f(x)\right| \leq 2\omega(f,\delta_{n,x}).$$

This implies that

$$\frac{1}{b_n(x)} \sum_{n:\left|P_n^{[\beta]}(f,x)-f(x)\right|\geq\varepsilon} a_{jn} \leq \frac{1}{b_n(x)} \sum_{n:\,2\omega(f;\delta_{n,x})\geq\varepsilon/2} a_{jn}.$$

The equality (4.2) gives the proof. ∎

Replacing "o" by "o_μ" one can get the following result immediately.

Theorem 4.3. *Let* $A = (a_{jn})$, $(b_n(x))$, $\beta := \{\beta_n\}$ *and* $\left\{P_n^{[\beta]}\right\}$ *be the same as in Theorem 3.4. Then*

$$P_n^{[\beta]}(f;x) - f(x) = st_A - o_\mu(b_n(x))$$

where $\omega(f;\delta_{n,x}) = st_A - o_\mu(b_n(x))$ *with* $\delta_{n,x} = \sqrt{P_n^{[\beta]}((e_1{-}x)^2;x)}$ *for fixed* $x \in [0,\infty)$. *Similar conclusions hold when little "o_μ" is replaced by big "O_μ".*

Remark. Notice that when $A = (a_{jn})$ is replaced by the identity matrix in Theorem 3.4 or Theorem 3.5 by choosing $\beta' := \left\{\beta_n'\right\} = \left\{\frac{1}{n}\right\}$, we get Theorem 3.1 in [9]. It is clear that $st_A - \lim_n \beta_n' = 0$.

Acknowledgements

We would like to thank the referees for their comments and suggestions on the manuscript. The third author is extremely thankful to Science and Engineering Research Board (SERB), Govt. of India, for providing the financial support under Teachers Associateship for Research Excellence (TARE) Award (TAR/2018/000356).

References

[1] O. Szâsz, Generalization of S. Bernstein's polynomials to the infinite interval, J. Res. Nat. Bur. Standards 45 (1950) 239–245.

[2] W. Meyer-König, K. Zeller, Bernsteinsche Potenzreihen, Studia Math. 19 (1960) 89–94.

[3] A. Mei, A. Sharma, Approximation methods by polynomials and power series, Indg. Math. 29 (1967) 395–403.

[4] D.D. Stancu, Approximation of functions by a new class of linear polynomial operators, Rev. Roumaine Math. Pures Appl. XIII (8) (1968) 1173–1194.

[5] K. Balázs, Approximation by Bernstein type rational functions, Acta Math. Acad. Sci. Hungar. 26 (1–2) (1975) 123–134.

[6] A. Kajla, T. Acar, Blending type approximation by generalized Bernstein-Durrmeyer type operators, Miskolc Math. Notes 19 (1) (2018) 319–336.

[7] N. Deo, N. Bhardwaj, Some approximation theorems for multivariate Bernstein operators, Southeast Asian Bull. Math. 34 (2010) 1023–1034.

[8] G.M. Mirakyan, Approximation of continuous functions with the aid of polynomials (Russian), Docl. Akad. Nauk. SSSR 31 (1941) 201–205.

[9] G.C. Jain, Approximation of functions by a new class of linear positive operators, J. Aust. Math. Soc. 13 (3) (1972) 271–276.

[10] N. Deo, N. Bhardwaj, S.P. Singh, Simultaneous approximation on generalized Bernstein-Durrmeyer operators, Afr. Mat. 24 (1) (2011) 77–82.

[11] N. Deo, M.A. Noor, M.A. Siddiqui, On approximation by a class of new Bernstein type operators, Appl. Math. Comput. 201 (2008) 604–612.

[12] V. Gupta, Some approximation properties on q-Durrmeyer operators, Appl. Math. Comput. 197 (2008) 172–178.

[13] G.C. Jain, S. Pethe, On the generalizations of Bernstein and Szâsz-Mirakyan operators, Nanta. Math. 10 (1977) 185–193.

[14] M. Bodur, Modified Lupas-Jain operators, Math. Slovaca. 70 (2) (2020) 431–440.

[15] V.N. Mishra, P. Patel, Some approximation properties of modified Jain-Beta operators, J. Calc. Var. 2013 (2013) Article ID 489249.

[16] S. Tarabie, On Jain-Beta linear operators, Appl. Math. Inf. Sci. 6 (2) (2012) 213–216.

[17] B. Tunca, G. Bodur, D. Soylemez, On Lupaş-Jain operators, Stud. Univ. Babes-Bolyai Math. 63 (4) (2018) 525–537.

[18] R.A. DeVore, G.G. Lorentz, Constructive Approximation, Springer, Berlin, 1993.

[19] H. Fast, Sur la convergence statistique, Colloq. Math. 2 (1951) 241–244.

[20] J.A. Fridy, H.L. Miller, A matrix characterization of statistical convergence, Analysis 11 (1991) 59–66.

[21] E. Kolk, Matrix summability of statistically convergent sequences, Analysis 13 (1993) 77–83.

[22] H.I. Miller, A measure theoretical subsequence characterization of statistical convergence, Trans. Amer. Math. Soc. 347 (1995) 1811–1819.

[23] E. Kolk, Statistically convergent sequences in normed spaces, Reports of convergence, Methods of algebra and analysis, Acta Comment. Univ. Tartu. Math. (1998) 63–66.

[24] E. Kolk, The statistical convergence in Banach spaces, Acta Comment. Univ. Tartu. Math. 928 (1991) 41–52.

[25] A.D. Gadjiev, C. Orhan, Some approximation theorems via statistical convergence, Rocky Mountain J. Math. 32 (1) (2002) 129–138.

[26] O. Duman, M.K. Khan, C. Orhan, A-Statistical convergence of approximating operators, Math. Ineq. Appl. 6 (4) (2003) 689–699.

[27] N. İspir, V. Gupta, A-Statistical approximation by the generalized Kantorovich-Bernstein type rational operators, Southeast Asian Bull. Math. 32 (2008) 87–97.

[28] O. Dogru, O. Duman, C. Orhan, Statistical approximation by generalized Meyer-Köning and Zeller type operators, Studia Sci. Math. Hung. 40 (2003) 359–371.

[29] M. Mursaleen, V. Karakaya, M. Ertürk, F. Gürsoy, Weighted statistical convergence and its application to Korovkin type approximation theorem, Appl. Math. Comput. 218 (18) (2012) 9132–9137.

[Downloaded free from http://www.jcnonweb.com on Tuesday, January 25, 2022, IP: 14.139.251.100]

Original Article

Assessment of Gestational Age by New-born Joint Angles

Arnab Nandy, Tanushree Mondal[1], Divyoshanu Mondal Ivan[2], Tapti Sengupta[3], Sankar Kumar Das[4], D. Goldar[5], Avijit Hazra[6], Rakesh Mondal[7]

Department of Pediatrics, NB Medical College, Siliguri, Departments of [1]Community Medicine, [4]Pediatrics, and [7]Pediatric Medicine, Medical College, [6]Department of Pharmacology, IPGMER and SSKM Hospital, Kolkata, [3]Department of Microbiology, WBSU, Barasat, West Bengal, [2]Department of Biotechnology, Delhi Technological University, [5]Ex-Expert Experimental Mechanics, Delhi Technological University Delhi, New Delhi, India

ABSTRACT

Objective: We aimed to measure the joint angles of major joints in newborn babies to explore their association with gestational age (GA). **Materials and Methods:** This observational study was conducted to evaluate joint angles in healthy newborn infants, born between 28 and 41 completed weeks of gestation. Healthy and hemodynamically stable babies were selected by purposive sampling up to the 2nd day of life. Six major joints of upper and lower limbs (shoulder, elbow, wrist, hip, knee, and ankle) were assessed on either side. The joint angles and range of motion (ROM) achieved through predefined passive movements were measured by a single observer using a manual acrylic goniometer. The positioning of the joints during measurement was standardized. **Results:** A total of 433 newborn babies, belonging to different GA, were evaluated. The joint angles and ROM did not show any significant differences between right and left sides and between male and female babies. The ROM of flexion-extension and adduction-abduction movement at shoulder, angle of palmer-flexion at wrist, and angle of dorsiflexion at ankle, were observed to correlate strongly with GA (correlation coefficient *r* of –0.75, –0.74, –0.73, and –0.73, respectively). The relationships were inverse. A mathematical model based upon multiple regression analysis is proposed to predict GA from these four measurements. **Conclusions:** Structured clinical assessment of major joint angles in newborn babies has the potential to provide relatively precise estimation of GA, which may be used to add accuracy to modified New Ballard Score.

Keywords: *Articular goniometry, gestational age, joint angles, prediction model*

Submitted: 20-Sep-2021
Accepted: 19-Nov-2021
Published: 03-Jan-2022

Introduction

Clinical assessment of different major joints angle in newborn babies has been conventionally utilized in the verification of gestational age (GA). Accordingly, several maneuvers and signs are mentioned in the literature which have laid down the foundation of the scheme of GA evaluation.[1,2] These describe passive movements involving the major joints in the newborn. The joint angles which are formed by various passive movements can be greatly influenced by the gradual development of muscle tone. This is the physiologic basis for the corroboration of joint angles with newborn maturity which is reflected in the GA at birth. However, in this context, a lacuna can be noticed about the nonavailability of precise quantification of joint angles in Indian newborns as per GA in completed weeks.[3] We sought to address the lacuna by generating corresponding data and exploring the potential association between newborn joint angles and their GA.

Materials and Methods

Ours was an institution-based single center analytical observational study conducted with healthy and hemodynamically stable newborn infants aged up to 2 days of life. The participants were stratified according to their period of gestation in completed weeks, ranging between 28 and 41 weeks. Consecutive subjects who

Access this article online

Quick Response Code:

Website: www.jcnonweb.com

DOI: 10.4103/jcn.jcn_116_21

Address for correspondence: Dr. Rakesh Mondal, Department of Pediatrics, Medical College, Kolkata, West Bengal, India.
E-mail: ivanrakesh2001@gmail.com

This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as the author is credited and the new creations are licensed under the identical terms.

For reprints contact: WKHLRPMedknow_reprints@wolterskluwer.com

How to cite this article: Nandy A, Mondal T, Ivan DM, Sengupta T, Das SK, Goldar D, *et al.* Assessment of gestational age by new-born joint angles. J Clin Neonatol 2022;11:23-9.

[Downloaded free from http://www.jcnonweb.com on Tuesday, January 25, 2022, IP: 14.139.251.100]

were delivered at our center or attended the outpatient clinic and met the inclusion and exclusion criteria were included if parental written informed consent was available. Sick newborns and babies with obvious congenital anomaly were not included. Furthermore, small for GA and large for GA newborn infants were screened with the help of Fenton's preterm growth chart and excluded from the study. Due approval from the Institutional Ethics Committee was taken beforehand.

GA of the participants was recorded from the antenatal ultrasound reports. It was compared with the GA documented from the last menstrual period of mother. In instances, where the two did not corroborate, New-Ballard Score (NBS) recorded for the participants was taken into consideration to obviate the risk of classification bias.

An 8-inch 360° plastic manual International Standards of Measurement goniometer (Baseline Evaluation Instruments, Model No. 12–100) made up of transparent acrylic material was used in the study to measure different joint angles. Minimum interval of recording for this instrument was 1°. All readings were taken by a single observer with the assistance of nursing staff. Standard aseptic precautions were followed in handling the babies.

Six major joints of upper and lower limbs (shoulder, elbow, wrist, hip, knee, and ankle) were evaluated on each side. The joint angles achieved through specific passive movements at the respective joints were measured. Shoulder and hip joint were evaluated for both flexion-extension and adduction-abduction movements, elbow and knee joint for flexion-extension movement, wrist joint for palmer flexion-dorsiflexion movement, and ankle joint for plantarflexion-dorsiflexion movement. Gentle pressure in contrast to the normal handling of newborn babies was applied while performing the predefined passive and rolling movements. For recording the joint angle, two lever arms of the joint served as the reference lines of the two limbs of the goniometer, whereas the joint itself acted as fulcrum. The meeting point of the lateral or anterior bisecting lines of the lever arms was taken as the reference point to measure different joint angles as applicable to the study procedure. The visual goniometric assessments of joint angles were expressed in degrees. Assessments of all the joint angles were done in lateral posture except the passive adduction-abduction movement of hip joint which was performed in the supine position.

The flexion-extension range of motion (ROM) at the shoulder joint was recorded with goniometer at the height of complete passive flexion and extension movement in a vertical plane which was placed at 30° internally rotated from the sagittal plane. ROM of adduction-abduction movement was recorded as complete passive adduction and abduction movement in a vertical plane which was placed at 30° internally rotated from the frontal plane, around the longitudinal axis of the articular interface. Owing to the natural angle of torsion of humeral upper end and scapular anteversion, the antero-posterior axis in the transverse plane of the articular interface of the shoulder joint was placed in such a manner which was approximately rotated 30° internally from the sagittal plane.[4,5]

In the hip joint, the ranges of motion were recorded at the height of complete passive flexion and extension movement in the sagittal plane, and full adduction and abduction passive movement in the frontal plane, respectively, around the medio-lateral and antero-posterior axis of the articular interface. It is noteworthy that normal acetabular anteversion and angular torsion of femoral neck in the transverse plane had apparently nullified each other to place the antero-posterior axis of the articular interface of hip joint in the sagittal plane.[6,7]

The joint angles formed in the sagittal plane at the elbow and knee joints were recorded after exercising complete passive flexion and extension movement. Similarly, the joint angles formed by complete passive palmer flexion and dorsiflexion movements at the wrist, and complete passive plantarflexion and dorsiflexion movements ankle joint were recorded successively in the sagittal plane. On certain instances where complete passive movement could not be performed, the angle of limitation for the movement was documented.

Statistical considerations

We did not go in for a formal sample size calculation for this descriptive study. Instead, based upon case load and logistical considerations, we aimed to recruit at least 30 participants for each completed week of gestation between 28 and 41 weeks or 420 subjects overall, irrespective of gender, through purposive sampling. Joint angle data were found to be normally distributed by Kolmogorov − Smirnov goodness-of-fit test. Descriptive statistics have been generated as mean and standard deviation and 95% confidence interval ranges, stratified by GA in completed weeks. Subgroup comparisons (left versus right; male versus female) were done by Student's paired samples or independent samples *t*-test, as appropriate, with cutoff for statistical significance at $P < 0.05$. Strength of association between angle or ROM of specific passive movements for individual joints (in degrees) and GA (in weeks) was quantified by calculating Pearson's product

[Downloaded free from http://www.jcnonweb.com on Tuesday, January 25, 2022, IP: 14.139.251.100]

moment correlation coefficient (r), after verifying the linearity assumption through scatterplots. Statistica version 6 (Tulsa, Oklahoma: StatSoft Inc.; 2001) and MedCalc version 11.6 (Mariakerke, Belgium: MedCalc Software bvba; 2011) software were used for the statistical analysis.

Results

Overall, 487 newborn babies were screened as probable participants of the study. For some (n = 16) parents did not give consent for study. On occasions (n = 35) all the predefined joint angles could not be assessed, whereas for few subjects (n = 3), inadmissible fallacy were noticed at the time of documenting the GA. Thus, data of 433 newborns were analyzed, which included 205 male (47.3%) and 228 female (52.7%) babies. Participants were stratified according to GA in completed weeks.

Table 1 summarizes all the joint angles assessed in the 433 newborns with comparison between left and right sides. This summary is presented without stratification. Left-right comparisons by Student's paired samples t-test did not indicate any statistically significant differences. The comparisons between male and female babies (data not shown) by independent samples t-test also did not reveal any significant differences. This table also presents the strength of association between angles or ranges of individual movements (in degrees) and GA (in completed weeks) as Pearson's correlation coefficient (r) values. Evidently, strong inverse correlations ($r > 0.65$) are observed for all movements at shoulder joint, flexion at elbow, palmar flexion at wrist, palmar flexion-dorsiflexion at wrist, and plantar flexion and dorsiflexion at ankle. Most movements at hip and knee show good but not strong correlation ($0.5 < r < 0.65$).

Table 2 presents the joint angle measurements stratified by GA for the four movements that show the strongest correlations, namely passive flexion-extension ROM (r = –0.75) and passive adduction-abduction ROM (r = –0.74) at shoulder, passive palmer-flexion at wrist (r = –0.73), and passive dorsiflexion (r = –0.73) at ankle. Figures 1-4 present the scatter plots corresponding to these associations. Notably, all these associations are linear but inverse indicating the decline of joint flexibility with increasing GA.

Multiple linear regression analysis was performed to explore whether any consistent mathematical relationship can be deduced between these four strongly correlated joint angles and GA in completed weeks. The resultant regression equation was: GA (in completed weeks)= 51.054 – (0.033*A) –

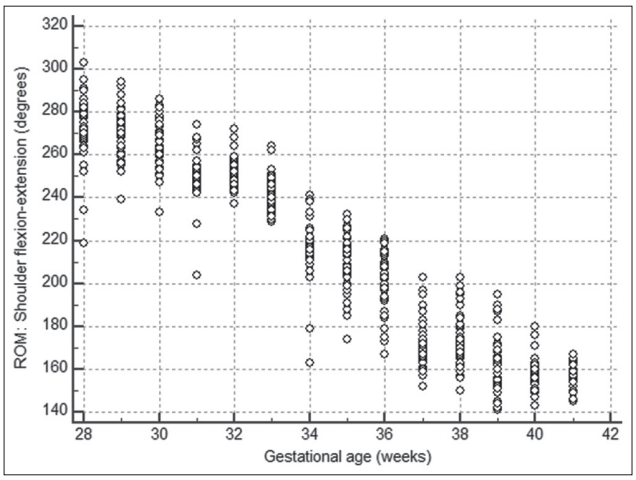


Figure 1: Scatter diagram showing the linear relationship of angle of passive flexion-extension range of motion at shoulder joint with newborn gestational age in completed weeks

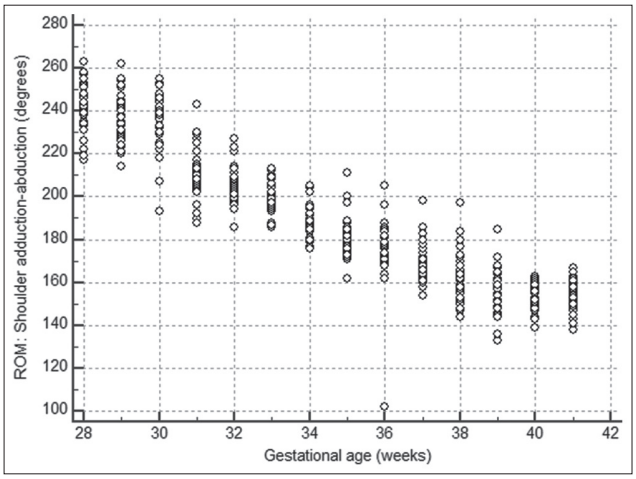


Figure 2: Scatter diagram showing the linear relationship of angle of passive adduction-abduction range of motion at shoulder joint with newborn gestational age in completed weeks

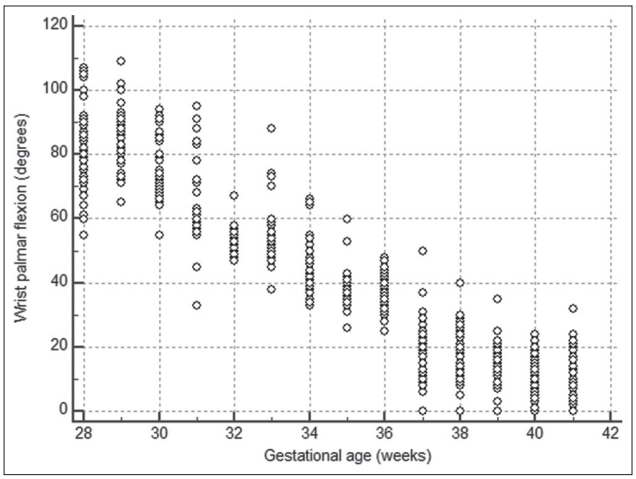


Figure 3: Scatter diagram showing the linear relationship between angle of passive palmer-flexion at wrist and newborn gestational age in completed weeks

[Downloaded free from http://www.jcnonweb.com on Tuesday, January 25, 2022, IP: 14.139.251.100]

Table 1: Goniometric joint angle measurements at 6 major synovial joints in newborn babies studied

Joint	Angle measured	Side	Mean (SD)	95% CI	*P* value for right-left comparison	Correlation (*r*) with gestational age (weeks)
Shoulder	Passive flexion	Right	127.38 (30.33)	124.86-129.89	0.180	−0.73
		Left	127.38 (30.33)	124.86-129.90		
	Passive extension	Right	86.85 (16.19)	85.50-88.19	0.763	−0.68
		Left	86.85 (16.18)	85.50-88.19		
	Range of passive flexion-extension	Right	214.22 (44.26)	210.55-217.90	0.593	−0.75
		Left	214.23 (44.24)	210.56-217.90		
	Passive adduction	Right	75.87 (16.25)	74.52-77.22	0.180	−0.69
		Left	75.88 (16.24)	74.53-77.23		
	Passive abduction	Right	114.27 (17.03)	112.86-115.68	0.180	−0.71
		Left	114.26 (17.02)	112.85-115.68		
	Range of passive adduction-abduction	Right	190.14 (31.73)	187.51-192.78	1.000	−0.74
		Left	190.14 (31.72)	187.51-192.78		
Elbow	Passive flexion	Right	22.05 (8.88)	21.32-22.79	0.564	−0.70
		Left	22.06 (8.88)	21.32-22.79		
	Passive extension	Right	16.11 (8.68)	15.39-16.83	0.058	+0.53
		Left	16.12 (8.68)	15.40-16.84		
	Range of passive flexion-extension	Right	141.84 (7.30)	141.23-142.44	0.052	+0.22
		Left	141.82 (7.29)	141.22-142.43		
Wrist	Passive palmer-flexion	Right	43.18 (27.09)	40.93-45.43	0.513	−0.73
		Left	43.19 (27.09)	40.94-45.43		
	Passive dorsiflexion	Right	90.21 (8.06)	89.54-90.87	1.000	+0.11
		Left	90.21 (8.06)	89.54-90.87		
	Range of passive palmer flexion - dorsiflexion	Right	226.61 (27.50)	224.33-228.90	0.549	+0.68
		Left	226.61 (27.50)	224.33-228.89		
Hip	Passive flexion	Right	28.65 (11.50)	27.70-29.61	0.249	−0.61
		Left	28.66 (11.50)	27.70-29.61		
	Passive extension	Right	17.24 (9.82)	16.43-18.06	1.000	+0.57
		Left	17.24 (9.82)	16.43-18.06		
	Range of passive flexion-extension	Right	134.11 (9.28)	133.34-134.88	0.394	+0.15
		Left	134.10 (9.28)	133.33-134.87		
	Passive adduction	Right	35.17 (12.09)	34.17-36.18	0.763	−0.64
		Left	35.18 (12.09)	34.17-36.18		
	Passive abduction	Right	36.02 (7.34)	35.41-36.63	0.366	−0.57
		Left	36.02 (7.35)	35.41-36.63		
	Range of passive adduction-abduction	Right	71.19 (17.74)	69.72-72.67	0.670	−0.67
		Left	71.19 (17.75)	69.72-72.66		
Knee	Passive flexion	Right	28.60 (9.18)	27.84-29.36	0.655	−0.63
		Left	28.60 (9.17)	27.84-29.36		
	Passive extension	Right	15.34 (8.79)	14.61-16.07	0.655	+0.54
		Left	15.34 (8.79)	14.61-16.07		
	Range of passive flexion-extension	Right	136.06 (8.77)	135.33-136.79	0.564	+0.12
		Left	136.06 (8.76)	135.34-136.79		
Ankle	Passive plantarflexion	Right	144.05 (21.58)	142.26-145.84	1.000	−0.71
		Left	144.05 (21.59)	142.25-145.84		
	Passive dorsiflexion	Right	39.39 (17.82)	37.91-40.87	0.051	−0.73
		Left	39.41 (17.82)	37.93-40.88		
	Range of passive plantarflexion - dorsiflexion	Right	104.66 (10.55)	103.78-105.53	0.088	−0.22
		Left	104.64 (10.54)	103.77-105.52		

CI – Confidence interval; SD – Standard deviation

(0.035*B) – (0.031*C) – (0.037*D), where A represents angle of passive flexion-extension ROM at shoulder; B passiveadduction-abduction ROM at shoulder, C angle of passive palmer-flexion at wrist, and D angle of passive dorsiflexion at ankle, with all four variables being significant predictors. The overall model quality

[Downloaded free from http://www.jcnonweb.com on Tuesday, January 25, 2022, IP: 14.139.251.100]

Table 2: Selected goniometric joint angle measurements in newborns stratified by gestational age

Gestational age (completed weeks)	Mean (SD)	10th percentile	25th percentile	Median	75th percentile	90th percentile
Passive flexion-extension range of motion at the shoulder						
28 weeks	272.40 (14.91)	258.00	267.50	272.50	281.50	288.00
29 weeks	269.70 (11.35)	256.00	262.00	270.50	276.50	283.00
30 weeks	265.72 (11.25)	252.00	257.00	267.00	274.00	279.00
31 weeks	249.82 (11.38)	243.00	245.50	249.00	254.50	266.00
32 weeks	250.32 (7.24)	243.00	245.50	248.50	254.50	258.50
33 weeks	242.53 (8.17)	232.00	235.50	243.00	247.50	251.00
34 weeks	216.13 (13.79)	206.00	212.00	216.00	222.00	232.00
35 weeks	210.83 (12.87)	193.00	203.50	213.50	219.00	226.00
36 weeks	201.95 (13.42)	181.50	195.50	203.5	213.00	216.00
37 weeks	172.67 (11.69)	160.00	165.50	170.50	179.50	192.50
38 weeks	172.80 (12.78)	156.50	164.00	171.50	179.50	194.00
39 weeks	160.37 (14.65)	144.00	150.00	155.50	168.50	185.00
40 weeks	156.80 (7.37)	148.00	151.00	157.00	159.50	164.00
41 weeks	157.10 (5.84)	147.50	155.00	157.00	162.00	164.50
Passive adduction-abduction range of motion at shoulder						
28 weeks	239.95 (10.66)	226.50	233.00	241.00	246.00	254.00
29 weeks	235.88 (10.53)	223.50	228.50	234.50	242.50	251.50
30 weeks	236.03 (11.96)	223.00	230.00	238.00	244.00	247.50
31 weeks	209.88 (10.90)	196.00	203.50	209.00	214.00	226.00
32 weeks	203.80 (7.97)	196.50	198.50	202.50	207.50	213.50
33 weeks	200.48 (6.51)	193.50	196.50	199.50	204.50	209.50
34 weeks	187.32 (6.92)	178.00	184.00	186.00	190.50	195.50
35 weeks	179.83 (8.69)	172.50	175.00	177.50	183.50	188.50
36 weeks	175.52 (14.56)	168.00	171.00	177.00	181.00	187.00
37 weeks	168.07 (8.51)	160.00	161.50	167.50	171.50	179.00
38 weeks	162.83 (10.59)	150.50	155.50	162.50	168.00	175.00
39 weeks	156.00 (9.77)	144.00	150.00	158.00	161.50	166.00
40 weeks	152.78 (5.96)	144.00	148.00	152.50	157.50	160.00
41 weeks	153.65 (6.78)	145.00	149.00	154.00	159.00	161.50
Angle of passive palmar flexion at wrist						
28 weeks	81.83 (13.21)	64.00	72.50	81.50	90.00	102.00
29 weeks	83.70 (9.16)	71.50	77.50	84.00	88.50	94.50
30 weeks	76.03 (9.49)	65.00	68.50	74.00	84.50	90.00
31 weeks	63.18 (12.02)	55.50	56.50	60.00	65.50	83.50
32 weeks	53.50 (4.11)	48.00	51.00	52.50	57.00	58.00
33 weeks	55.65 (8.93)	47.00	49.00	55.00	59.50	65.00
34 weeks	44.23 (8.78)	34.50	38.00	42.00	47.00	59.50
35 weeks	38.82 (6.47)	33.00	34.00	38.00	43.00	43.00
36 weeks	37.35 (5.97)	30.00	32.50	37.00	42.50	45.00
37 weeks	18.73 (9.26)	8.50	12.50	18.00	24.00	28.00
38 weeks	15.60 (9.10)	2.50	10.00	14.00	22.00	27.50
39 weeks	11.25 (7.81)	0.00	7.50	11.00	16.50	20.50
40 weeks	12.20 (6.50)	2.00	7.50	13.00	17.50	20.00
41 weeks	12.48 (7.58)	3.50	6.00	12.00	19.00	22.00
Angle of passive dorsiflexion at ankle						
28 weeks	67.42 (6.18)	60.50	64.00	67.50	73.00	74.00
29 weeks	59.97 (5.20)	53.00	55.00	60.00	64.50	67.50
30 weeks	60.85 (6.30)	55.00	57.00	60.00	64.00	67.50
31 weeks	50.83 (6.43)	43.00	45.00	52.00	53.50	56.00
32 weeks	47.23 (4.53)	41.50	43.00	50.00	51.00	53.50

Contd...

[Downloaded free from http://www.jcnonweb.com on Tuesday, January 25, 2022, IP: 14.139.251.100]

Table 2: Contd...

Gestational age (completed weeks)	Mean (SD)	10th percentile	25th percentile	Median	75th percentile	90th percentile
33 weeks	49.15 (5.24)	40.00	45.00	47.00	53.00	55.00
34 weeks	42.22 (6.49)	34.00	38.50	42.50	45.50	49.00
35 weeks	37.13 (5.13)	31.00	33.50	36.00	41.00	44.50
36 weeks	36.67 (7.39)	30.00	32.00	36.00	43.00	45.00
37 weeks	23.40 (7.59)	11.50	19.00	24.00	29.00	31.00
38 weeks	21.80 (8.41)	12.50	16.00	20.00	28.00	32.00
39 weeks	18.10 (7.46)	12.00	13.00	17.50	21.00	23.50
40 weeks	18.55 (5.98)	12.00	14.00	16.50	24.50	27.50
41 weeks	18.18 (5.90)	10.50	14.00	18.00	23.00	27.00

SD – Standard deviation

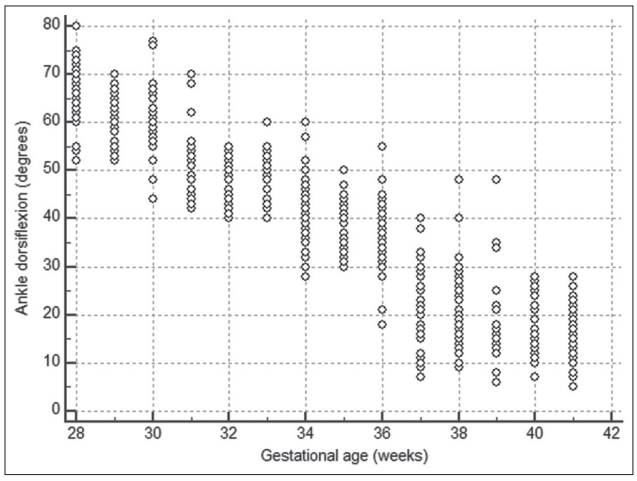


Figure 4: Scatter diagram showing the linear relationship between angle of passive dorsiflexion at ankle and newborn gestational age in completed weeks

was very good with coefficient of determination $R^2 = 0.946$, indicating that most of the variability in the outcome (GA) can be explained by the four predictors in the model.

Discussion

Assessment of GA is a salient feature of clinical evaluation of newborn babies. The modified NBS helps considerably in this purpose. The method makes use of a composite scoring system to estimate newborn GA in a wide range of completed weeks. For neuromuscular maturity assessment, it has described approximate estimation of different joints angle in broad range of degrees with certain interval under the bare eye examination.[8] However, joint angle being a quantitative parameter, it is possible to get more precise value of the measurement. Evidently, these measurements in numbers have the potential to indicate newborn GA not only in completed weeks but possibly also in days if we are able to measure accurately and follow the principles of mathematics.

Different clinical methods have been postulated and assessed over the time to determine newborn infant maturity accurately.[9] The proposals from Amiel-Tison, Dubowitz *et al*., and Ballard *et al*. have left their trail on the subsequent works that are conducted in this field.[2,10,11] It was Claudine Amiel-Tison who first described from her observations on developmental physiology that muscle tone could have a direct association with the range of movements performed at the different joints in newborn infants.[2] The acquisition of muscle tone and motor functions is documented to follow caudo-cephalic pattern of progression during the 28–40 weeks of gestation where the development of flexor tone usually precedes extensor tone.[12,13] Thus, measurements of the joints angle being achieved by passive movements can be used for evaluation of newborn maturity. However, the conundrum persists regarding estimation of the joints angle which are graded by scores to provide an impression on the status of newborn maturity. Instead, if the joint angles are measured precisely and comprehensively in newborns, the definite values of these measurements can reciprocate into explicit determination of the GA.

In our study, predefined passive movements at the major joints were evaluated in newborn babies of different GA using goniometric assessments in a standardized manner for the various joints. This has yielded precise data. Single observer assessment reduced the possibility of interobserver bias. As expected, laterality or gender did not introduce and differences. Correlation analysis indicated that passive flexion-extension ROM and passive adduction-abduction ROM at shoulder, angle of passive palmer flexion at wrist and angle of passive dorsiflexion at ankle had the strongest associations with newborn GA. The regression model obtained from the study indicates that goniometric assessment of joint angles could be utilized as a clinical tool to find GA of the newborn babies conforming to completed weeks.

Reference data of different newborn joint angles according to the completed weeks of newborn GA are scarce in the literature.[14-16] Some of the studies are conducted with similar objectives but their results cannot

[Downloaded free from http://www.jcnonweb.com on Tuesday, January 25, 2022, IP: 14.139.251.100]

be compared as the subjects are recruited in cluster of weeks instead of specific groups of GA in completed weeks.[17-20]

The primary limitation of this study is the fact that newborn babies lack bony landmarks which poses considerable difficulty in joint angle measurements. Although the possibility of interobserver bias was reduced in this by a single observer conducting the measurements, in actual clinical practice this difficulty can introduce subjective variability in measurements. Influence of other factors which can possibly affect the measurement of joints angle like volume of soft tissues and maternal antenatal medications have not been taken into the consideration.

Conclusion

Despite these limitations, it is reasonable to conclude that standardized manual goniometric measurement of just four joint angles can help in precise estimation of newborn GA in completed weeks, which even may be used to add accuracy to modified NBS.

Financial support and sponsorship

Nil.

Conflicts of interest

There are no conflicts of interest.

References

1. Usher R, McLean F, Scott KE. Judgment of fetal age. II. Clinical significance of gestational age and an objective method for its assessment. Pediatr Clin North Am 1966;13:835-62 contd.
2. Amiel-Tison C. Neurological evaluation of the maturity of newborn infants. Arch Dis Child 1968;43:89-93.
3. Nandy A, Guha A, Datta D, Mondal R. Evolution of clinical method for new-born infant maturity assessment. J Matern Fetal Neonatal Med 2020;33:2852-9.
4. Norkin CC, White DJ. Measurement of Joint Motion: A Guide to Goniometry. 3rd ed. New Delhi: Jaypee Brothers Medical Publishers Private Limited; 2004. p. 3-12.
5. Ludewig PM, Borstad JD. The shoulder complex. In: Levangie PK, Norkia CC, editors. Joint Structure and Function: A Comprehensive Analysis. 5th ed. New Delhi: Jaypee Brothers Medical Publishers Private Limited; 2012. p. 244-8.
6. Wientroub S, Tardiman R, Green I, Salama R, Weissman SL. The development of the normal infantile hip as expressed by radiological measurements. Int Orthop 1981;4:239-41.
7. McKibbin B. Anatomical factors in the stability of the hip joint in the newborn. J Bone Joint Surg Br 1970;52:148-59.
8. Ballard JL, Khoury JC, Wedig K, Wang L, Eilers-Walsman BL, Lipp R. New Ballard score, expanded to include extremely premature infants. J Pediatr 1991;119:417-23.
9. Wariyar U, Tin W, Hey E. Gestational assessment assessed. Arch Dis Child Fetal Neonatal Ed 1997;77:F216-20.
10. Dubowitz LM, Dubowitz V, Goldberg C. Clinical assessment of gestational age in the newborn infant. J Pediatr 1970;77:1-10.
11. Ballard JL, Novak KK, Driver M. A simplified score for assessment of fetal maturation of newly born infants. J Pediatr 1979;95:769-74.
12. Illingworth RS. Normal Development. In: Nair MK, Russell PS, editors. Illingworth's The Development of the Infant and Young Child: Normal and Abnormal. 10th ed. New Delhi: Reed Elsevier India Private Limited; 2012. p. 92-3.
13. Robinson RJ. Assessment of gestational age by neurological examination. Arch Dis Child 1966;41:437-47.
14. Jayakrishnan TT, Sharma S, Gulati S, Pandey RM, Wadhwa S, Paul VK. Agreement between visual and goniometric assessments of adductor and popliteal angles in infants. J Pediatr Neurosci 2013;8:93-6.
15. Farmania R, Sitaraman S, Das RR. Goniometric assessment of muscle tone of preterm infants and impact of gestational age on its maturation in Indian setting. J Neurosci Rural Pract 2017;8:S44-8.
16. Waugh KG, Minkel JL, Parker R, Coon VA. Measurement of selected hip, knee, and ankle joint motions in newborns. Phys Ther 1983;63:1616-21.
17. Forero N, Okamura LA, Larson MA. Normal ranges of hip motion in neonates. J Pediatr Orthop 1989;9:391-5.
18. Vidhi S, Parmar L, Khandare S, Palekar T, Shah V. Measurement of elbow range of motion in full term neonates. Int J Sci Res 2017;7:7721-31.
19. Harris MB, Simons CJ, Ritchie SK, Mullett MD, Myerberg DZ. Joint range of motion development in premature infants. Pediatr Phys Ther 1990;2:185-91.
20. Broughton NS, Wright J, Menelaus MB. Range of knee motion in normal neonates. J Pediatr Orthop 1993;13:263-4.

Environmental Science and Pollution Research
https://doi.org/10.1007/s11356-021-17695-w

RESEARCH ON SUSTAINABLE DEVELOPMENTS FOR ENVIRONMENT MANAGEMENT

Assessment of the effect of the judicial prohibition on firecracker celebration at the Diwali festival on air quality in Delhi, India

Shailendra Kumar Yadav[1] · Rajeev Kumar Mishra[1] · Bhola Ram Gurjar[2]

Received: 6 April 2021 / Accepted: 18 November 2021
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

Abstract
Diwali (the festival of lights and crackers) is celebrated grandly, resulting in a significant drop in the city's air quality. To study the impact of the judicial prohibition in Delhi to improve air quality, a comprehensive and comparative analysis was conducted over two consecutive years, namely 2015–2016 (when no significant regulations on the sale or usage of firecrackers were imposed) and 2017–2018 (when radically different regulations were implemented). Data on PM_{10}, $PM_{2.5}$, NO_x, and CO were analysed, and their trends and levels with various regulations in place were compared. In 2017, the concentrations of PM_{10}, $PM_{2.5}$, NO_x, and CO were reduced by 50%, 50%, 71%, and 64%, respectively, compared to 2016. However, in 2018, there was an increase of 32% in PM_{10} and $PM_{2.5}$ concentrations, as well as a 25% increase in CO concentrations, with the exception of NOx, which decreased to 25% on Diwali day. The data was also examined in conjunction with the entire timeline of the various court rulings and regulations imposed in Delhi. The questionnaire survey study revealed that, despite the legislation in place, ambient air quality continued to deteriorate, necessitating a deeper dive into the policy's structure and implementation to fine-tune its feasibility and applications. Air pollution-related health effects were recognized by 82% of participants. Despite this, only 13% of people were observed without a mask, and only 12% of people were aware of green crackers as of 2018. To combat this deteriorating situation, the national capital must enact radical and well-thought-out legislation and regulations governing firecrackers, as well as raise public awareness amongst its citizens.

Keywords Air pollution · Air pollution policy · Firecrackers · Judicial prohibition · Megacities · Questionnaire survey

Introduction

Diwali is one of the most important festivals in the Indian subcontinent. It is celebrated with grandeur and opulence, with an extensive display of firecrackers and pyrotechnics on a national scale. The national capital, New Delhi, is no exception to this tradition. The festival of Diwali falls annually in October or November. Several research groups studied firecrackers as an integral part of the festival, associated with the ambient air quality in Diwali for their concentration of pollutants, chemical speciation, and health effects (Sarkar et al. 2010; Singh et al. 2010, 2019; Perrino et al. 2011; Singh and Dey 2012; Tiwari et al. 2012; Beig et al. 2013; Kumar et al. 2016b; Rastogi et al. 2019). These firecrackers are made up of paper tubes filled with a mixture of metal salts, charcoal, sulphur, and a binding agent that when ignited emits thick layers of smoke and brilliant lights (Pervez et al. 2016). The extensive burning of crackers results in a largescale emission of several hazardous chemical pollutants and particulate matter ranging from oxides of sulphur, nitrogen, carbon dioxide, carbon monoxide, metals, and metal ions to charcoal-containing compounds (Singh et al. 2010; Nigam et al. 2016; Gautam et al. 2018; Ghei and Sane 2018). Also, large amounts of these pollutants in the atmosphere have been observed to catalyse the formation of secondary particulate matter (Kumar et al. 2016a).

The air quality of Delhi from the last few years reaches a more critical level in the Diwali period and puts pressure on the policy of the concerned agencies (Dholakia et al. 2013; Polash et al. 2016; Gurjar et al. 2016; Balakrishnan et al. 2019). New Delhi, along with the National Capital Region

Responsible Editor: Philippe Garrigues.

✉ Rajeev Kumar Mishra
rajeevkumarmishra@dtu.ac.in

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

[2] Department of Civil Engineering, Indian Institute of Technology, Roorkee, Uttrakhand 247667, India

Published online: 04 January 2022

(NCR), has an approximate population of 22.2 million in 2011 with a decennial increase of 7.96% (Ministry of home affairs 2011) and a projected vehicle population of 25.6×10^6 with an annual increase of 5.83% based on business as usual (BAU) estimates (Kumar et al. 2011), the second-largest urban area in the world. However, in addition to the unique geographical and landlock conditions in Delhi, the situation of winter and low dispersion worsens the ambient air quality (Chaudhari et al. 2016; Kumar et al. 2017; Guttikunda et al. 2019) and public health conditions (Health Effects Institute 2019; Singh et al. 2019). For the reasons set out above, in November 2016, which Kanawade et al. (2020) described as a Sever Air Pollution Episode (SAPE), the Air Quality Index (AQI) reached an abysmal level (Sati and Mohan 2014) (Source—National Air Quality Index; https://app.cpcbccr.com/AQI_India/). Many experts have also named Delhi the "Gas Chamber" (Shukla and Anisha 2016). The mean concentration of $PM_{2.5}$ in new Delhi before, during, and after the SAPE were 142 µg/m^3, 563 µg/m^3, and 240 µg/m^3 respectively. The Government, central public sector enterprises and state-level public enterprises (CPSU and SLPE), and other concerned organizations have taken a variety of measures to reduce air pollution, such as a ban on old vehicles (Goyal and Gandhi 2016; Kumar et al. 2017), potential construction activities, biomass burning, diesel generator, and Parali burning in and around Delhi (Srivastava and Jain 2008). As vehicle emissions are one of the major sources of ambient air pollution in Delhi (Nagpure et al. 2013), the Government of Delhi introduced an odd–even scheme in which an odd and even number of vehicles are allowed on an alternative day basis. Many experts have documented the positive effects of this study and have suggested that it is not a permanent solution (Kumar et al. 2017; Mohan et al. 2017; Chandra et al. 2018; Mishra et al. 2019). A judiciary prohibition was also implemented on firecrackers for buying, selling, and its bursting timing in and around the Delhi region in 2017 (TOI 2016) since, as mentioned earlier, Diwali also comes early winter from October to November. The judicial ban also includes green crackers instead of regular crackers (TOI 2020). Eco-friendly crackers (green crackers) have been formulated by Indian scientists at the Council of Scientific and Industrial Research (CSIR) laboratories to minimize noise and air pollution emissions. As recorded by the CSIR, it reduces the PM_{10} and $_{2.5}$ through 30–35% emissions and sulphur and nitrogen oxides by 35–40% and has a sound level of less than 120 dB (decibel unit) (Srishti 2016).

The objective of this study is to determine the effect of the legal framework on the concentration of Diwali air pollution using secondary data analysis (Diwali 2015 to 2018) obtained from the Government of India's Central Pollution Control Board (CPCB). Data was collected from only a few representative stations to assess the impact and effectiveness of two drastically different policies implemented in the National Capital Territory of Delhi to reduce pollutants emitted during the Diwali festivities and to investigate more feasible mitigation and management strategies for achieving the desired results. The selection of stations in this study was based on the availability of data and their locations being as spread out in the city as possible in order to produce a sample that was as close to a representative as possible. Along with the opinion-based online, offline, and face-to-face surveys, a random survey was also conducted in Delhi to analyse and assess the impact of the honourable Supreme Court policy on Diwali air pollution and attempt to identify gaps in regulation to combat air pollution caused by firecrackers. This research is divided into three major sections: (1) the worldwide legal prohibition on firecrackers, including in India, (2) secondary real-time monitoring of air pollution data in Delhi and the NCR areas to assess pollutant concentrations before and after policy implementation, and (3) a questionnaire survey of Delhi residents to assess judicial prohibition in the Delhi and NCR areas before and after policy implementation.

Literature on firecracker ban around the world and in India

As the state of ambient air quality deteriorated, the judiciary in many countries (UK, USA, Canada, Germany, Ireland, India, Netherlands, Norway, Belgium, and Australia) was forced to develop complex prohibition policies regarding firework celebrations over time, whilst the Indian judiciary is not the first in the world to regulate the sale, release, and bursting of firecrackers. Other countries have also prohibited the sale and purchase of firecrackers. In 2004, the UK passed the "Firecrackers Law," which regulated this. This was done to reduce illegal firecracker selling and possession, noise and air pollution, and illegal firecracker sales (Instruments 2005). Despite the fact that firecrackers are a common Chinese practise of welcoming the New Year, the sale of firecrackers has been banned in more than 444 towns and cities in the People's Republic of China, including 10 provincial capitals, 91 prefecture-level cities, and 343 counties; 764 cities restrict the distribution of firecrackers and firecrackers, including 4 municipalities and 15 provincial capitals. In China, the imposition of a ban resulted in a reduction of $PM_{2.5}$ levels in the ambient air by ~ 14% in 2017 and ~ 31% in 2014 (Wen and Thomas 2018; Silver et al. 2018). Firework displays during New Year's celebrations are only permitted in Germany between 6 p.m. on New Year's Evening and 7 a.m. the following morning. Similarly, different laws apply to different states in the USA (Samy 2018). Most states allow the use of different types of firecrackers, such as novelty items and sparklers, whilst others, such as Delaware and Massachusetts, prohibit the sale of all consumer firecrackers, and still others have placed restrictions on firecracker

commerce without a permit on both the retailer and customer end (APA 2018). Other than novelty and sparklers, firecrackers can be sold throughout the year in Finland, but only on New Year's Eve until the next morning. Consumer firecrackers are strictly controlled in Ireland, and their use is supervised by a licenced firecracker operator (EC 2006; Fitzgibbon 2014). Fireworks cannot be sold to anyone under the age of 16 in the Netherlands, and they can only be sold between 6 p.m. on New Year's Eve and 2 a.m. the following day (Pieters 2018).

Similarly, Norway has segregated the firecrackers into Class 1, 2, and 3, which are sold for 4 days till before New Year's Eve and can only be burst only during the same time slot as the Netherlands (EC 2006). In Belgium, firecrackers cannot be sold to anyone under the age of 16. They are available for purchase all year but cannot be burst without the police's permission (EC 2006; Hope 2015). In Canada, firecracker laws vary across provinces. Some areas like Alberta, British Columbia, and others have implemented a total ban on crackers' bursting. At the same time, states like Ontario, Atlantic Canada, and Prince Edward Island, and many more allow firecrackers year-round (Singh et al. 2019). Fireworks are prohibited in the majority of Australian states. They can only be sold and used on July 1, one day a year in the Northern Territory (NT), with time restrictions. NT governs the sale of firework cakes and novelty items to the general public. In certain circumstances, Tasmanians can also apply for a permit to use firecrackers (Queensland Government 2018). In New Zealand, it may only be sold between November 2 and November 5. However, firecrackers can be used at any time throughout the year. Sales of firecrackers have been restricted increasingly recently. Skyrockets and other aerial firecrackers were banned in 1994. As of 2007, the limit on the number of sparklers bought per capita has been reduced to 50 dues to their popularity. In Chile, firecrackers are restricted for consumer use and can only be seen in professional shows. The prohibition was enacted at the request of organizations that help burnt children, with many of the accidents caused by the unsupervised use of firecrackers. Improper usage of firecrackers leads to hefty fines and imprisonment in all countries (Fitzgibbon 2014).

The trial of the Indian judiciary against firecrackers in Delhi began when the Supreme Court dismissed the petition to ban firecrackers during Diwali 2015. In 2016, the ambient air quality of New Delhi rose to alarmingly severe levels due to the excessive bursting of firecrackers in the national capital and the burning of crops in neighbouring states. Whilst the observed peak of 2016 was an aberration, the decade's general trend proved that pollution levels increased rapidly during the festive season of Diwali. Since then, various government policies have sought to minimize emissions in the city during the festival. On November 25, 2016, the Supreme Court of India (SCI) (Arjun Gopal vs Union of India 2019) passed a ruling, banning the firecracker's sales within Delhi—NCR to combat the deteriorating air quality. According to this order, all the existing licenses of the existing firecracker traders in the city were suspended, and no new licenses were issued. It also directed the CPCB to monitor the harmful effects of the chemical used in firecrackers and report back within 3 months. On September 12, 2017, the SCI reversed its decision, stating that it was "too radical a step" and instead recommended a graded gradual approach to a complete ban that would take effect on November 1, 2017 (Arjun Gopal vs Union of India 2019).

On October 1, 2017, the Delhi High Court issued an order prohibiting the burning of agricultural waste. It also hired companies and plants (from Delhi, Haryana, Punjab, and Uttar Pradesh) to fulfil corporate social responsibility by collecting agricultural waste and funding farmers to dispose of it. The Ministry of Environment, Forestry, and Climate Change was directed to compile weekly environmental reports from special committees established by the governments of each concerned state (Apoorva 2017; Arjun Gopal vs Union of India 2019).

On October 9, 2017, the SCI reverted to its original ruling from November 2016 to test whether a festive season without firecrackers would positively impact the city's ambient air quality and, as a result, the citizens' health (Krishnadas 2017). Diwali 2017 data showed that poor conditions persisted, and the SCI ruling could not produce the desired results. On Diwali, the AQI ranges from 327 in 2015 to 621–642 in 2018. The highest AQI was recorded in 2018, when regulations were less stringent, and more firecrackers were lit than in 2017. The strict guidelines on the ban of firecrackers resulted in the lowest AQI in 2017. The AQI has clearly improved, ranging from 327, 426, 326–367, 621–642 in 2015, 2016, 2017, and 2018 (Table S1; Source—National Air Quality Index; https://app.cpcbccr.com/AQI_India/). Notably, the ban on the sale of firecrackers reduced particulate matter levels significantly when compared to previous years' trends. As a result, it is critical to understand and further analyse Delhi's population's exposure levels to particulate matter (PM_{10}, $PM_{2.5}$), oxides of nitrogen (NO_x), oxides of sulphur (SO_x), and carbon monoxide (CO) during the Diwali festive season, as well as to implement a viable public policy to ensure the populace's safety and health whilst reducing morbidity due to these pollutants.

Materials and methods

Location and topography

The experiment was carried out in New Delhi (28°36′36″N, 77°13′48″E), at an altitude of about 216 m above mean sea level and covering an area of 1484 km^2. The Himalayas

connect the metropolitan area to the north, the Thar Desert in Rajasthan to the west, and the central Indian plains to the south. Owing to its position, Delhi experiences semi-arid climatic conditions with a maximum of 46 °C in summer and a minimum of 2 °C temperature in the winter. The India Meteorological Department defines four distinct seasons for the city of Delhi, namely winter (December–March), summer (April–June), monsoon (July–September), and post-monsoon (October–November). The city receives an average rainfall of 670 mm annually; most of it (75–80%) occurs in the monsoon season, characterized by high humidity levels. Westerly or North-Westerly winds dominate Delhi during the months of Diwali (As per the Hindu calendar, Diwali is observed on Amavasya (or new moon)—the 15th day of the month of Kartik, every year, which comes in October and November of the Gregorian calendar). The temperature begins to fall in the city starting in October. In November, lower temperatures are combined with foggy conditions that reduce visibility and air masses that contain many pollutants that pass-through Delhi during this season. Lower temperature can be attributed to the low mixing heights observed in these months that render mixing pollutants with the upper boundary layer difficult (Tiwari et al. 2012). According to some studies, the reduced visibility is also due to the high frequency of Western disturbances during the foggy months, low wind speed, and temperature inversion (Shivani et al. 2019).

The data was collected from CPCB monitoring stations located at Mandir Marg, R.K. Puram, Netaji Subhash Institute of Technology (NSIT), Lodhi Road, Delhi Technological University (DTU), and Indira Gandhi International (IGI) Airport. Figure 1 depicts all of the selected locations on a Google Earth map of Delhi. Monitoring stations for the study were chosen based on data availability and demographic presentation during the Diwali period at CPCB or India Meteorological Department (IMD) monitoring stations in Delhi. The Particulate Matter (PM_{10}, $PM_{2.5}$) Dataset was collected on a 24-h temporal resolution for the festive season of Diwali in 2015, 2016, 2017, and 2018. Throughout the study, the data was analysed for three monitoring periods: Pre Diwali (the 7 days preceding the Diwali festival); Diwali (the day of the festival); and Post-Diwali (the 7 days

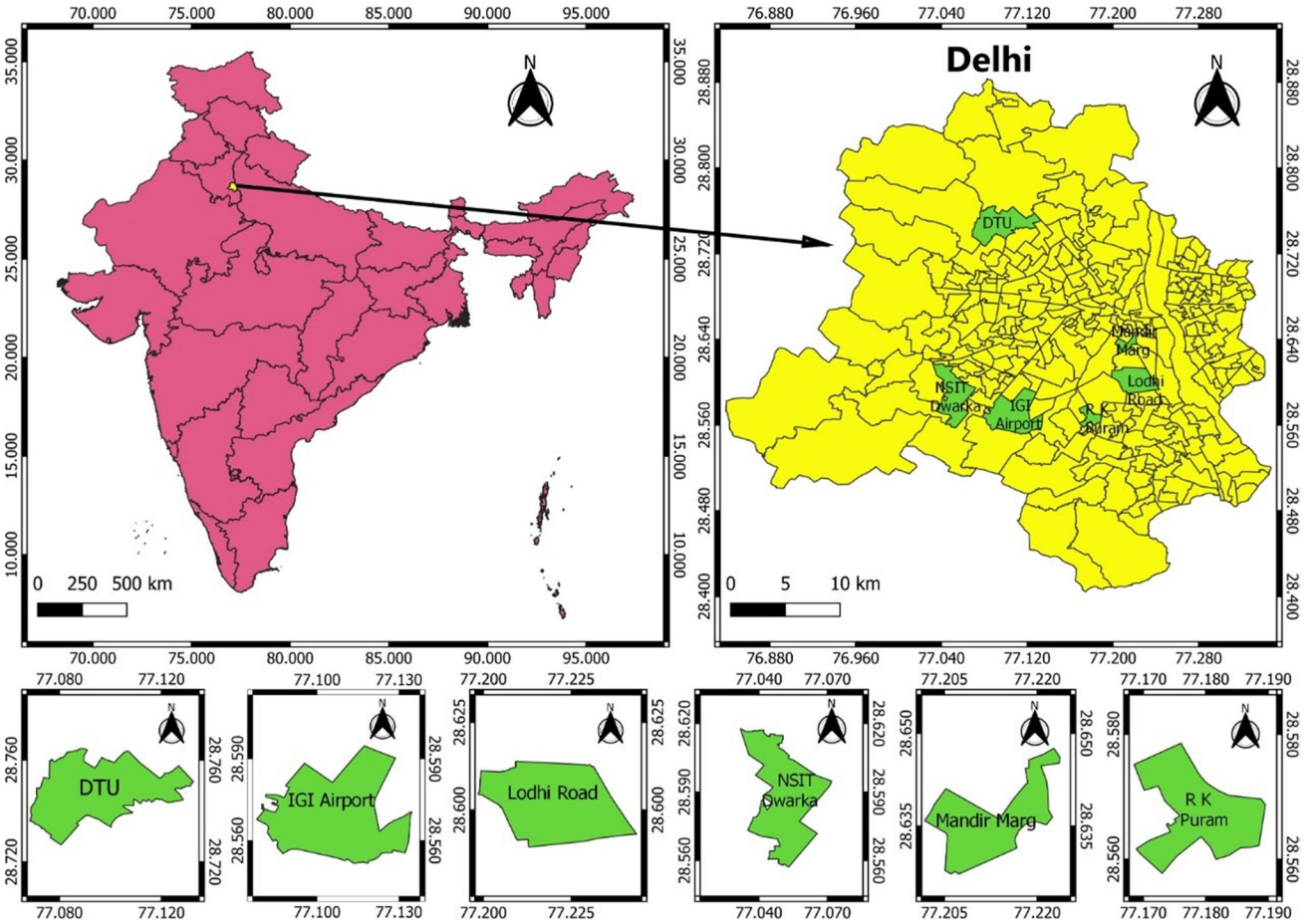


Fig. 1 Representative CPCB monitoring stations in New Delhi

following the festival). The days were classified as described above, despite the fact that their dates differed in the Gregorian Calendar due to the use of the Hindu Lunar Calendar at the start of Diwali.

Statistics

The dataset normality was tested using a combination of the Exploratory Data Analysis (EDA) plot and the Shapiro–Wilk (S-W) test. The Kruskal–Wallis (K-W) test, a non-parametric version of the analysis of variance (ANOVA) test, was performed on air quality data. The Kruskal–Wallis test (rank based non-parametric test) was used to determine the statistical differences between years (2015, 2016, 2017, 2018) and independent variable pollutant (PM10, PM2.5, NOx, CO) of representative location in Delhi. The Kruskal–Wallis test denotes with H test:

$$\boldsymbol{H} = \left[\frac{12}{\boldsymbol{n}(\boldsymbol{n}+1)}\sum_{j=1}^{c}\frac{\boldsymbol{T}_i^2}{\boldsymbol{n}_i}\right] - 3(\boldsymbol{n}+1) \quad (1)$$

where, n = sum of sample size for all samples; c = number of samples; T_j = sum of rank in the j^{th} sample; n_j = size of the j^{th} sample. In particular, the H test examines whether the null hypothesis is valid that there are no statistical mean variations in population variances; if not true, then an alternative hypothesis is accepted. A 0.05 significance level was used in both cases. Therefore, if the p value was lower than the significance threshold (< 0.05), the null hypothesis was rejected, and there was a statistically meaningful difference between the mean results (supplementary materials). Percentage change of pollutant with respect of year is calculated by

$$\%\boldsymbol{Changeof Con.} = \frac{\boldsymbol{Con.TY} - \boldsymbol{Con.BY}}{\boldsymbol{Con.BY}} \times 100 \quad (2)$$

where, $Con.TY$ = concentration of pollutant in the target year; $Con.BY$ = concentration of pollutants in the base year. Here, base year (BY) is the previous year of the target year (TY).

The Diwali festival, which is celebrated by each economic class of the greater part of Hindus and a few other societies (Sikh, Jain, Buddh, etc.) in Delhi, provides a diverse population for the questionnaire survey. As a result, probabilistic sampling, i.e., the stratified random sampling method, was used in this study. The population was divided into groups (strata) based on an economic factor (low, medium, and high), and data were collected on each sampling unit that was randomly selected from each group (stratum). The questionnaire was divided into three sections. The first section included demographic information such as gender, age, occupation. The second section included questions about pollution, pollutants, and air quality to evaluate the awareness, and the third section included questions about the health effects experienced by the locals. The questionnaire survey was conducted and validated by the "Cronbach alpha test" for the questionnaire reliability test to determine public opinion on air pollution awareness and experience-based state of air pollution judgement and the effect of the legal ban on firecrackers during the Diwali season on air pollution reduction. A general rule of thumb is that 0.6–0.7 indicates an acceptable level of reliability, and 0.8 or higher indicates a very good level. However, values greater than 0.95 are not always desirable because they may indicate redundancy (Ursachi et al. 2015). Statistical tool "IBM-SPSS" software has been used in all statistical analysis.

Result and discussion

Pollution trend

The primary sources of air pollutants in Delhi are vehicle pollution (Goyal et al. 2013; Nagpure et al. 2013), biomass burning (NAAQS 2012; Banerjee et al. 2015), industrial emissions (Srivastava and Jain 2008; Mathew et al. 2015; Rathi et al. 2016), and many other minor sources, which further exacerbated during Diwali by cracker bursting. Many entities, including the judiciary, have taken exemplary measures to improve air quality in the Delhi and NCR regions of Diwali. Their influence can be seen in the time-series graph (Fig. 2) of the representative pollutants (PM_{10}, $PM_{2.5}$, NO_x, CO). R K Puram and Mandir Marg monitoring stations show a marked decline in contrast (2017) with severe air pollution episodes (SAPE; 2016), whilst the rest of the monitoring locations cannot be compared to SAPE due to data unavailability. The data from CPCB was plotted linearly by dividing the data into three phases, as explained earlier. Graphs for understanding the parameters obtained from different stations have been shown in the form of a graphical matrix below which Diwali Day has an annotation "D" (Fig. 2). However, in the case of AQI on Diwali Days, it decreased from 13.85 to 23.48% in 2017 compared to 2016 due to the enforcement of the SCI firecracker prohibition. However, it rises to a high level of > 500 AQI (see SI Table of AQI). Figure 1 indicates a higher concentration of pollutants in 2018 at all five representative locations compared to 2017. This may be due to flexibility in the prohibition (SCI 2018) and general perception of SCI's ban on firecrackers, leading to inactive public involvement.

The Kruskal–Wallis test analysed the statistically median difference in pollutant concentration with a p value of 0.05 by testing air pollution data from representative monitoring locations separately over a group of years as 2015, 2016, 2017, and 2018. Table 1 shows the H test value, degree of freedom (df), and significance level (p value) with the K-W

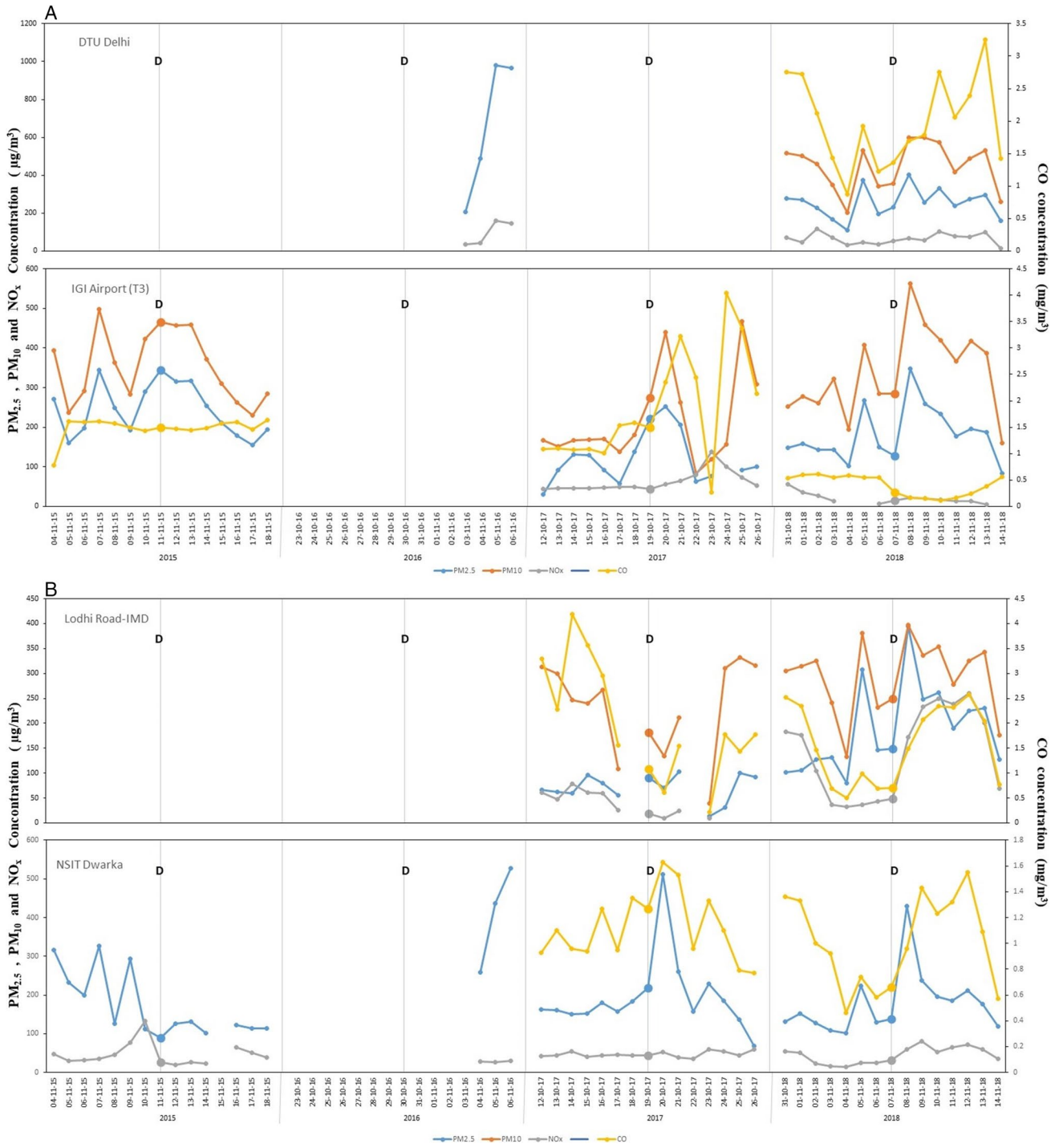


Fig. 2 The levels of PM_{10}, $PM_{2.5}$, NO_x, CO during the festive period of Diwali (2015–2018) at different representative stations (A, B, and C)

decision. At all locations, the concentrations of PMs and gaseous pollutants show a statistically significant difference. However, CO concentration at two locations (NSIT, Lodhi Road) due to unavailability of data and NOx concentration at NSIT retains the K-W test hypothesis for the year 2015–2018. However, significant concentration reduction can be seen in Fig. 2, especially in comparison to 2017. It clearly shows the ban's impact on firecrackers during the Diwali period, but it does not show the same magnitude of pollution reduction in the following year, 2018.

Pairwise comparison between years in relation to the concentration of pollutants does not show a linear pattern of



Fig. 2 (continued)

Table 1 *H* test, df, and *p* value of pollution data

Station	Parameter	*H* test	df	*p* value (0.05)	Decision
Mandir Marg	PM_{10}	15.732	3	0.001287	Rejected
	$PM_{2.5}$	8.7947	3	0.03215	Rejected
	NO_x	16.273	3	0.000997	Rejected
	CO	17.705	3	0.000506	Rejected
R K Puram	PM_{10}	10.918	3	0.01218	Rejected
	$PM_{2.5}$	31.134	3	$7.967*10^{-7}$	Rejected
	NO_x	18.717	3	0.0003129	Rejected
	CO	23.936	3	$2.576*10^{-5}$	Rejected
DTU	PM_{10}	-	-	-	-
	$PM_{2.5}$	9.6937	2	0.007853	Rejected
	NO_x	17.879	2	0.0001311	Rejected
	CO	-	-	-	-
NSIT	PM_{10}	-	-	-	-
	$PM_{2.5}$	8.4529	3	0.03752	Rejected
	NO_x	4.0367	3	0.2575	Retain
	CO	0.62204	1	0.4303	Retain
Lodhi Road	PM_{10}	3.8329	1	0.05026	Rejected
	$PM_{2.5}$	17.38	1	$3.061*10^{-5}$	Rejected
	NO_x	7.6923	1	0.005546	
	CO	0.65022	1	0.42	Retain
IGI Airport	PM_{10}	12.03	2	0.002442	Rejected
	$PM_{2.5}$	17.288	2	0.0001762	Rejected
	NO_x	25.921	2	$2.352*10^{-6}$	Rejected
	CO	26.953	2	$1.404*10^{-6}$	Rejected

significant level (*p* value) between all locations (Table S2). However, in the case of the average median rank, the year 2017 shows a high average difference in rank between 2015, 2016, 2018, and these three years do not offer the same, although these years have a lesser difference (2015, 2016, 2018). It is well known that the Diwali celebration in North India begins early in the evening and ends late at night, with firecrackers bursting and lighting being directly proportional to people's economic status and locality. It may be one of the main reasons behind the pairwise comparison, but it does not show a clear pattern concerning the concentration of pollutants. Delhi's demographic and urban strata are very diverse, so the celebration varies accordingly.

Diwali is a holiday that marked a drop in the traffic flow volume (Devara et al. 2015) and, to some extent, small-scale enterprises. Taking the pollutant concentration on Diwali day into account aided in determining the judiciary prohibition on Diwali, as shown in Table 2. At all the selected monitoring stations (Mandir Marg, RK Puram, DTU, NSIT, Lodhi Road, IGI Airport), the data for 2015 to 2018 was analysed for 15 days average and 24 h average on the day of Diwali with bracket in Table 2. Observing the PM_{10}, $PM_{2.5}$ trend, it was discovered that the concentration in 2018 was higher than in 2017, but it was not the same in the case of gaseous. NO_x and CO concentration was reduced at the same location also (Table 2 and Fig. 2). The year 2016 shows an unprecedented spike which can be explained by the absence of any significant regulatory policy and the coincidence of

Table 2 Percentage change in parameters at the selected stations

Years	PM_{10} (%)	$PM_{2.5}$ (%)	NO_x (%)	CO (%)
Mandir Marg				
2016 with BY 2015	+58.5 (+30.12)	+33.33 (+6.38)	+28.05 (+308)	+57.69 (+74.13)
2017 with BY 2016	−58.6 (−44.7)	−55.47 (−47.41)	−57.30 (−78.43)	−52.03 (−62.37)
2018 with BY 2017	+49.01 (+0.023)	+83.8 (+26.18)	+9.21 (−23.88)	+138.13 (+214.5)
RK Puram				
2016 with BY 2015	+46.50 (+12.5)	+44.97 (−17.09)	+1.2 (+159.4)	+22.34 (+74.88)
2017 with BY 2016	−37.82 (−30.93)	−87.81 (−77.62)	+8.35 (−63.69)	−50.68 (−66.84)
2018 with BY 2017	+5.82 (−13.07)	+459.09 (+262.3)	−63.9 (−35.47)	−2.77 (+40.99)
DTU				
2018 with BY 2017	+38.95 (+11.65)	+82.2 (+67.48)	−49.2 (−37.98)	+46.15 (0.74)
NSIT				
2018 with BY 2017	——	+19.07 (−4.56)	+0.32 (+10.96)	+12.08 (−9.58)
Lodhi Road				
2018 with BY 2017	+23.36 (−7.5)	+81.3 (+25.34)	+10.23 (−39.80)	+26.98 (−16.67)
IGI				
2018 with BY 2017	+57.0 (+18.05)	+86.0 (+5.19)	——-	-65.38 (−75.23)

crop burning with Diwali (Hutton et al. 2007; Perrino et al. 2011; Tiwari et al. 2012; Devara et al. 2015; Parkhi et al. 2016; Nigam et al. 2016; Li et al. 2017). Both PM_{10} and $PM_{2.5}$ showed an average reduction of 50% in the concentration for 15 days of monitoring data, and a 59% concentration reduction was observed on the day of Diwali in 2017 as compared to 2016 when there was no restriction on sale or purchases of firecrackers in Delhi like 2017. The selected location's result showed a 48.21% and 37.82% reduction in concentration for 15 days of monitoring and on the day of Diwali, respectively, for PM_{10} and $PM_{2.5}$. PM_{10} increased by 43.8% after 15 days of monitoring and a 1.8% increase on Diwali. $PM_{2.5}$ has also increased 135.2% for 15 days of monitoring and 63.65% for the Diwali event. There was more increment in $PM_{2.5}$ concentration as compared to PM_{10} in the year 2018. In 2018, an increasing concentration of $PM_{2.5}$ was observed mostly at all the locations except a few sites, which indicates the impact of local response towards fire bursting in that particular area (Table 1). It is depicted from Fig. 2 that only RK Puram and Lodhi Road have less PM_{10} concentration in the year 2018 as compared to 2017 on the day of Diwali. However, the annual percentage of PM_{10} concentration was found higher at these two monitoring locations, namely, RK Puram (5.8%) and Lodhi road (23.4%).

Regarding $PM_{2.5}$ concentration levels (Fig. 2), 2016 and 2017 are considered the years of global maximum and global minima, respectively. Various monitoring stations showed slight variations in the peaks, which can be explained due to the presence or absence of residential suburbs. The observed readings of 2018 are more in lieu of those of 2015 when, again, there was no regulation in place. Even after accounting for the annual population growth rate of India, which is 1.1%, the fact that the data from 2018 follows the same trend as 2015 despite the presence of a comprehensively revised court ruling in place which restricted the bursting of crackers to 2 h in a day shows that there might have been lapses in the implementation. Table 2 shows that $PM_{2.5}$ concentration is slightly higher at IGI airport on the day of Dewali compared to the RK Puram monitoring station. However, at both stations, during 2018, the annual percentage change was observed as more than 80%. Reports indicate up to a 100% increase in PM_{10} concentrations during festivals that involve firecracker bursting. It was estimated in a study that in Italy, half of the observed PM_{10} level is attributable to firecrackers, whilst about 30% of Delhi's total PM_{10} could be attributed to firecrackers (Rastogi et al. 2019; Singh et al. 2019).

Oxides of nitrogen are primarily generated from the combustion of fuels. The smog blanket in the city of Delhi after the festival of Diwali has been found to contain high concentrations of oxides of nitrogen. Similar to the case of particulate matter, 2016 had the highest observed concentrations of NO_x, again probably due to the absence of any regulations and crop burning (Verma and Deshmukh 2014; You et al. 2017). After implementing the court ruling of 2017, there was a significant drop in the concentration of oxides of nitrogen compared to 2016 (24.47% and 71.06% reduction in the concentration of 15-day monitoring and on Diwali day, respectively) and 2015 as well. However, the revised policy of 2018, which restricted the burning of firecrackers, seems to have further caused a drop in the concentration of NO_x when compared to 2017 (18.66% and 25.23% reduction in the concentration of 15-day monitoring and on Diwali day, respectively), despite the same not being observed in the case of ambient air quality. Besides this, CO concentration shows the same trend as PM_{10}, $PM_{2.5}$ concentration in 2017

and 2018. In comparison with 2018, its concentration was reduced to 51.35% and 64.60% for 15 days of monitoring and on the day of Diwali respectively in 2017, but its concentration increased ~ 25% during the period of monitoring. A reduction of 71.6% for 15 days of monitoring and 62.5% for the Diwali festival stands positively with the impact of the Hon'ble Supreme Court ruling on Delhi's air quality during the Diwali festival celebration. However, this scenario of concentration reduction has changed in the year 2018 with the concern of PM_{10}, $PM_{2.5}$, NO_x, CO compared with the year 2017.

Resident questionnaire survey for the assessment of judiciary prohibition

A survey consisting of 20 questions was conducted to gauge the public awareness regarding the Supreme Court ruling on the regulations on firecracker usage during Diwali and the proper mitigation methods to tackle the worsening ambient air quality in Delhi. This was circulated mostly amongst the students of universities, colleges, and schools in Delhi. They were assigned to complete the same and collect responses from 3–4 individuals from different economic strata in their circles.

The survey form was created on the free platform provided by Google (Google Forms) and circulated amongst the students. It was explained that the responsibility of collecting the responses from a technologically illiterate individual was that of the particular student. The entire set of questions present in the form are furnished on the following pages.

After a previously decided circulation and a waiting period of a month, a total of 1364 responses were received, of which 1167 were found considerable for further analysis, and the same has been compiled and analysed. The questionnaire reliability test (Cronbach's alpha test) was done with an IBM SPSS statistical tool. Cronbach's alpha test is found at 0.656, higher than the 0.500 reliability test score with a number of items 23. Some of the essential findings have been elaborated upon in the following paragraphs.

A total of 1167 responses were analysed. About 65% of the responders were from the student community, around 8% were public sector employees, 16% in the private sector, 1.2% of army personnel, 2.5% of retired professionals. In comparison, 7% was made up of unemployed individuals. A clear majority of the responses were from males (74.7%) and 24.4% from females, whilst 0.9% preferred not to disclose their gender. When ranked the most harmful forms of pollution in Delhi, 794 responders seemed to agree that air pollution had the most negative impact; 492 opined that water pollution came next, and 385 felt that solid waste disposal and related pollution was next in line. It is also interesting to note that almost an equal number of responders were tied between the impact of noise and sound pollution. Consequently, most responders know that vehicular emissions are the primary source of air pollution in their locality. As per most rankings, the other sources are industrial pollution, road/construction debris, and open waste burning (Figures S1 and S2). From the survey, it is evident that despite there being a consensus that air pollution is a significant pollution hazard for Delhi, only 13.2% of responses indicated the continuous usage of air filtering media like masks whilst 24.2% alluded to using them often; 27.8% of the responses were of occasional usage, whilst a large chunk (34.9%) of the responders rarely or never used them, and 22.8% of the responders who used protective masks did not know its specifications, whilst 19.3% were aware of it. Interestingly, a significant percentage (32.2%) of the responses were from individuals who used kerchiefs to protect themselves from Delhi's air pollution. The majority of the responders (58.2%) revealed that they did not celebrate the festival of Diwali in Delhi despite being residents of the city. This demonstrates that, despite the fact that a large portion of the population is not present in the city during the festival, the staggering number of crackers burst by the remaining percentage severely degrades the city's air quality; 72.3% of the responders' families spent less than Rs. 500 on crackers, whilst 16.6% spent between Rs. 500 and 1000, 7.2% between Rs. 1000 and 2000, 2.5% between Rs. 2000 and 5000, 0.5% between Rs. 5000 and 10,000, and 0.9% spent more than 10,000 rupees. Moreover, 88.9% of the responses have pointed to expenditures below Rs—1000 on firecrackers, which points to a gaping disparity with the levels of pollution observed during the festival. Possible explanations could be that most responders of the survey belong to the section of Delhi's population who do not indulge in celebrating with firecrackers, or it could be due to the furnishing of incorrect information in the survey. A clear majority, 57.9% of the responders, reported to be strongly aware of the effects of firecrackers on the air quality, 23.8% were moderately aware, 11.3% reported satisfactory awareness, whilst the rest were unaware or lacked proper awareness. This indicates that despite almost 82% of the responders reporting good levels of awareness, Delhi repeatedly succumbs to stale air every year during the festival. This could be either due to an absence of proper awareness amongst the public or an even more dangerous eventuality. Individuals use firecrackers despite knowing their ill effects on the environment and subsequently themselves (Figure S3). According to Fig. 3, 43.4% of the responders reported being strongly aware of the Supreme Court's regulations on firecrackers in 2018; 29.8% were moderately aware, 16.3% had satisfactory work, and 10.55% lacked proper awareness. Despite a clear majority of the survey takers reportedly being aware of the policy, ambient air quality deterioration during the festival paints a different picture. Only 12.9% of the responders believed that the court ruling was implemented according to the 2-h

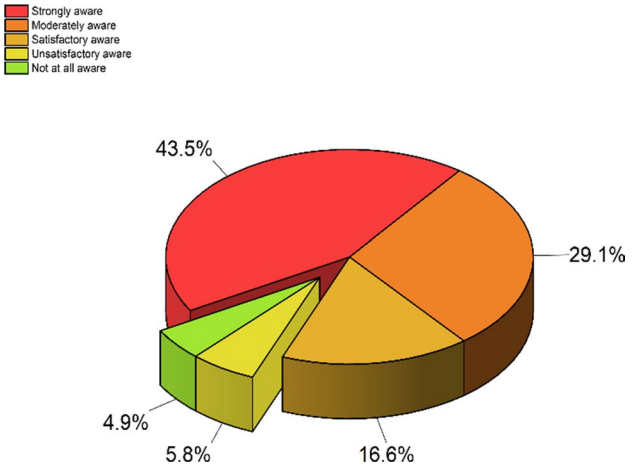


Fig. 3 Information status about the Supreme Court's rulings regarding firecracker's sale and bursting during Diwali 2017 and Diwali 2018 (total respondent—1082)

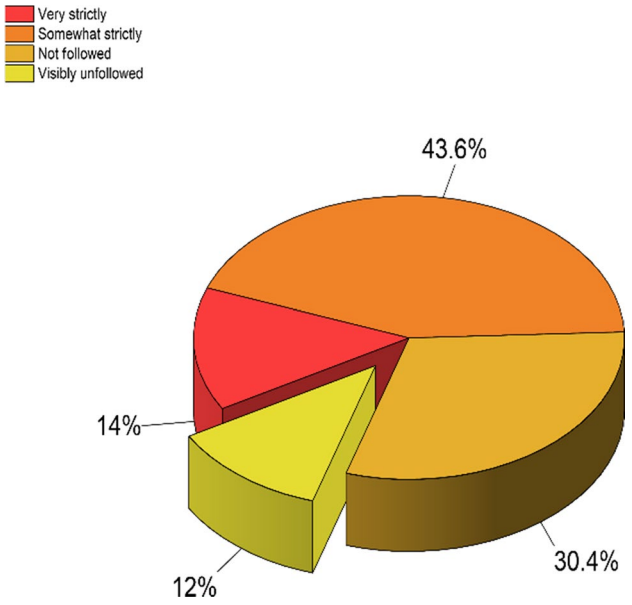


Fig. 4 The follow-ups regarding the SCI ruling of bursting firecrackers in the given time slot by locality

window in their locality, 42.8% opined that it was enforced somewhat strictly, whilst 44.3% of respondents stood with the opinion of mismanagement and poor enforcement of the ruling. Only 10.6% of the responses agreed over high police surveillance regarding implementing the Supreme Court ruling, 33.2% reported moderate surveillance, whilst 56.3% reported low or no surveillance. This suggests a pressing need for better surveillance by the Police Department (Figs. 4 and S4).

Despite higher levels of particulate matter in the ambient air in 2018 compared to 2017, 74.6% of respondents predicted a decrease in the number of firecrackers set off in 2017. Green crackers were an essential feature of the Supreme Court ruling of 2018. Despite 43.4% reporting to be actively aware of the same, only 12.9% were aware of green crackers, again pointing towards a blatant lack of awareness; 87.6% of the responders reported that they were suffering from health ailments after the Diwali celebrations. The health effects mentioned varied from eye irritation to bronchitis, sleeplessness, headaches, breathing difficulty, coughing bouts, and allergies.

According to the Supreme Court ruling of 2018, enforcing the cracker regulation was allocated to the respective Station House Officers (S.H.O.), i.e., the Delhi Police. There are about 100,000 police officers in Delhi Police for the city with a population of 190 million, which means there is one police officer for every 1900 people. This could have created difficulties in the implementation of the regulation discussed earlier. It has been reported that the Delhi Police registered 550 cases of violation during Diwali (Times of India 2018). However, the number seems to be a little too low considering the population and the fact that several reports of a considerable section of the populace flouting the ruling as indicated by the rise in the particulate matter levels (Ghose 2018).

The ruling regarding the green firecrackers was issued merely a fortnight before the festival, leaving minimal time for certifying and ensuring that only green firecrackers were produced and sold in the city. About 34% of the population of Delhi has an annual income of 150,000 rupees, which would reduce the feasibility of an extravagant splurge on green firecrackers during Diwali (Statista 2019). Instead, this marker of the population might resort to locally made firecrackers that are toxic. The verification of green crackers can only be done at the time of sale. It is nearly impossible for a patrol officer to identify whether the firecrackers being burst are toxic or otherwise. Hence, if a section of the population could acquire and bring in locally made firecrackers into the city, it would be improbable that any enforcing authority would prevent their usage. It is noteworthy to understand that post the extensive firecracker regulations in Chinese states, there has been a 73% decline in firecrackers' contribution to ambient $PM_{2.5}$. It has been reported that $PM_{2.5}$ levels in Shanghai dropped from 79 to 32 μg/m^3 between 2013 and 2017 (Vecchi et al. 2008; Singh et al. 2019). In contrast, there has been a noticeable increment in the same standards in Delhi's dynamics despite the above-discussed policies' prevalence. This further underlines the necessity for a well-structured plan that integrates both short-term and long-term benefits for the city's ambient air quality.

The absence of relevant data during specific periods at particular stations of CPCB is also a significant issue related to any related regulation's success. The CPCB must ensure

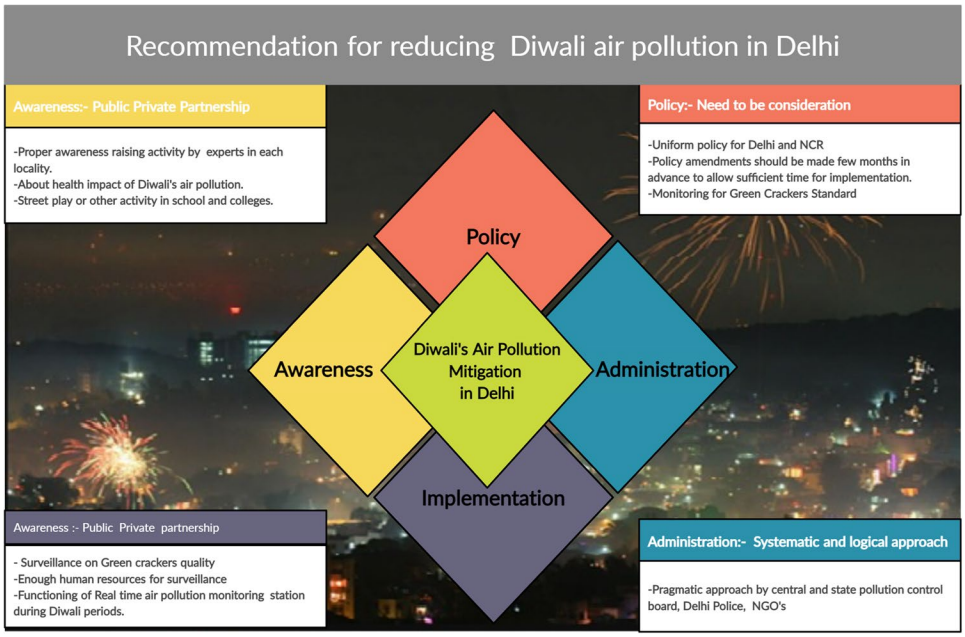


Fig. 5 Recommendation for curbing Diwali's air pollution impact in Delhi

continuous monitoring during this festive period and provide timely reports to the policymakers to evaluate the legislation's implementation and efficiency. Comprehensive meteorological data should also be collected over the festive timeline. Remote sensing can also be used to study the ambient quality where the physical monitoring stations are not present (Singh et al. 2019).

Besides, central, state, and local governments must arrange awareness drives and campaigns to inform people about the various severe hazards associated with using any type of firecrackers. It should be ensured that this message should not come across as a threat to their cultural and religious ethos but a pathway to ensure the amelioration of the prevailing health hazards faced by the majority of Delhi's population.

Conclusion

The data collected from CPCB resources suggest that, despite existing regulations on using firecrackers, the high levels of particulate matter and NO_x may indicate policy implementation flaws. It suggests that the general public is unaware of the dangers of poor air quality. As a result, there is an urgent need for radical policy changes to ensure greater efficacy, improved continuous monitoring by the CPCB, and increased public awareness to combat falling ambient air quality levels. The questionnaire survey demonstrates the importance of policy, administration, awareness, and implementation status in reducing Diwali air pollution, which requires more proper coordination about the SCI ruling's goal. Figure 5 depicts the suggested recommendations.

The air quality index (AQI) 2018 and PM_{10}, $PM_{2.5}$, NO_x, and CO concentrations on Diwali day in Delhi were discovered to be higher than the AQI 2017, demonstrating the impact of SCI ruling readings of firecrackers bursting during the Diwali period of 2017 and 2018.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11356-021-17695-w.

Acknowledgements The authors are incredibly grateful to Advance Air and Acoustics Research Laboratory, Delhi Technological University (DTU), Delhi (India), to encourage to conduct this research and provide all the facility to compile this work. The authors are also thankful to Mr Ashish Garg, undergraduate student of Department of Environmental Engineering and project student of Advance Air and Acoustics Research lab (DTU), for his support in data management.

Author contribution **SKY:** conceptualization, investigation, formal analysis, writing—original draft preparation; **RKM:** conceptualization, investigation, writing—review and editing, project administration; **BRG:** investigation, writing—review and editing.

Data availability The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate "Not applicable".

Consent for publication "Not applicable".

Conflict of interest The authors declare no competing interests.

References

APA (2018) State fireworks control law directory for the USA. In: Am Pyrotech Assoc

Apoorva M (2017) HC issues directions to Delhi, neighboring states to check air pollution though crop burning [read order]. In: Live Law

Arjun Gopal vs Union of India (2019) Judgments | Supreme Court of India

Balakrishnan K, Dey S, Gupta T et al (2019) The impact of air pollution on deaths, disease burden, and life expectancy across the states of India: the Global Burden of Disease Study 2017. Lancet Planet Heal 3:e26–e39. https://doi.org/10.1016/S2542-5196(18)30261-4

Banerjee T, Murari V, Kumar M, Raju MP (2015) Source apportionment of airborne particulates through receptor modeling: Indian scenario. Atmos Res 164–165:167–187

Beig G, Chate DM, Ghude SD et al (2013) Evaluating population exposure to environmental pollutants during Deepavali fireworks displays using air quality measurements of the SAFAR network. Chemosphere 92:116–124. https://doi.org/10.1016/j.chemosphere.2013.02.043

Chandra BP, Sinha V, Hakkim H et al (2018) Odd-even traffic rule implementation during winter 2016 in Delhi did not reduce traffic emissions of VOCs, carbon dioxide, methane and carbon monoxide. Curr Sci 114:1318–1325. https://doi.org/10.18520/cs/v114/i06/1318-1325

Chaudhari PR, Verma SR, Singh DK (2016) Experimental implementation of odd-even scheme for air pollution control in Delhi. India 2:57–65

Devara PCS, Vijayakumar K, Safai PD et al (2015) Celebration-induced air quality over a tropical urban station, Pune, India. Atmos Pollut Res 6:511–520. https://doi.org/10.5094/APR.2015.057

Dholakia HH, Purohit P, Rao S, Garg A (2013) Impact of current policies on future air quality and health outcomes in Delhi. India 75:241–248. https://doi.org/10.1016/j.atmosenv.2013.04.052

EC (2006) European measures for safer fireworks. Eur Comm

Fitzgibbon P (2014) The law on fireworks—Pierse Fitzgibbon solicitors. In: Pierse.ie

Gautam S, Yadav A, Pillarisetti A, et al (2018) Short-term introduction of air pollutants from fireworks during Diwali in Rural Palwal, Haryana, India: a case study. In: IOP Conference Series: Earth and Environmental Science

Ghei D, Sane R (2018) Estimates of air pollution in Delhi from the burning of firecrackers during the festival of Diwali. PLoS ONE 13:e0200371. https://doi.org/10.1371/journal.pone.0200371

Ghose D (2018) Supreme Court order on firecrackers violated: Delhi pollution soars, but who cares when sab chalta hai?—Firstpost. In: Firstpost. https://www.firstpost.com/india/supreme-court-order-on-firecrackers-violated-delhi-pollution-soars-but-who-cares-when-sab-chalta-hai-5518631.html. Accessed 25 Mar 2021

Goyal P, Gandhi G (2016) Assessment of air quality during the 'odd-even scheme' of vehicles in Delhi. Indian J Sci Technol 9:1–7. https://doi.org/10.17485/ijst/2016/v9i48/105801

Goyal P, Mishra D, Kumar A (2013) Vehicular emission inventory of criteria pollutants in Delhi. Springerplus 2:1–11. https://doi.org/10.1186/2193-1801-2-216

Gurjar BR, Ravindra K, Nagpure AS (2016) Air Pollution Trends over Indian Megacities and Their Local-to-Global Implications. Atmos Environ 142:475–495

Guttikunda SK, Nishadh KA, Jawahar P (2019) Air pollution knowledge assessments (APnA) for 20 Indian cities. Urban Clim 27:124–141. https://doi.org/10.1016/j.uclim.2018.11.005

Health Effects Institute (2019) State of global air 2019. Heal Eff Institute 24. https://www.stateofglobalair.org/sites/default/files/soga_2019_report.pdf

Hope A (2015) Fireworks law gets a rethink after 50 years | The Bulletin. In: Bull

Hutton G, Rehfuess E, Tediosi F (2007) Evaluation of the costs and benefits of interventions to reduce indoor air pollution. Energy Sustain Dev 11:34–43. https://doi.org/10.1016/S0973-0826(08)60408-1

Instruments S (2005) Fireworks The fireworks regulations 2004. Interpret A J Bible Theol 11:1–10

Kanawade VP, Srivastava AK, Ram K et al (2020) What caused severe air pollution episode of November 2016 in New Delhi? Atmos Environ 222:117125. https://doi.org/10.1016/j.atmosenv.2019.117125

Krishnadas R (2017) Supreme Court brings back ban on firecrackers in Delhi-NCR—the Hindu. In: The Hindu

Kumar M, Singh RKS, Murari V et al (2016a) Fireworks induced particle pollution: a spatio-temporal analysis. Atmos Res 180:78–91. https://doi.org/10.1016/j.atmosres.2016.05.014

Kumar P, Gulia S, Harrison RM, Khare M (2017) The influence of odd–even car trial on fine and coarse particles in Delhi. Environ Pollut 225:20–30. https://doi.org/10.1016/j.envpol.2017.03.017

Kumar P, Gurjar BR, Nagpure AS, Harrison RM (2011) Preliminary estimates of nanoparticle number emissions from road vehicles in megacity Delhi and associated health impacts. Environ Sci Technol 45:5514–5521. https://doi.org/10.1021/es2003183

Kumar P, Kumar R, Yadav S (2016b) Water-soluble ions and carbon content of size-segregated aerosols in New Delhi, India: direct and indirect influences of firework displays. Environ Sci Pollut Res 23:20749–20760. https://doi.org/10.1007/s11356-016-7313-x

Li J, Xu T, Lu X et al (2017) Online single particle measurement of fireworks pollution during Chinese New Year in Nanning. J Environ Sci (china) 53:184–195. https://doi.org/10.1016/j.jes.2016.04.021

Mathew J, Goyal R, Taneja KK, Arora N (2015) Air pollution and respiratory health of school children in industrial, commercial and residential areas of Delhi. Air Qual Atmos Heal 8:421–427. https://doi.org/10.1007/s11869-014-0299-y

Ministry of home affairs (2011) Census of India website: Office of the Registrar General & Census Commissioner, India. Government of India 388:Cahpter 2, Pg 22–3

Mishra RK, Pandey A, Pandey G, Kumar A (2019) The effect of odd-even driving scheme on PM 2.5 and PM 1.0 emission. Transp Res Part D Transp Environ 67:541–552. https://doi.org/10.1016/j.trd.2019.01.005

Mohan D, Tiwari G, Goel R, Lahkar P (2017) Evaluation of odd-even day traffic restriction experiments in Delhi, India. Transp Res Rec 2627:9–16. https://doi.org/10.3141/2627-02

NAAQS (2012) National Ambient Air Quality Status & Trends in India-2010. P R Div Cent Pollut Control Board 17–20

Nagpure AS, Sharma K, Gurjar BR (2013) Traffic induced emission estimates and trends (2000–2005) in megacity Delhi. Urban Clim 4:61–73. https://doi.org/10.1016/j.uclim.2013.04.005

Nigam S, Kumar N, Mandal NK et al (2016) Real time ambient air quality status during Diwali Festival in Central, India. J Geosci Environ Prot 04:162–172. https://doi.org/10.4236/gep.2016.41017

Parkhi N, Chate D, Ghude SD et al (2016) Large inter annual variation in air quality during the annual festival "Diwali" in an Indian megacity. J Environ Sci (china) 43:265–272. https://doi.org/10.1016/j.jes.2015.08.015

Perrino C, Tiwari S, Catrambone M et al (2011) Chemical characterization of atmospheric PM in Delhi, India, during different periods of the year including Diwali festival. Atmos Pollut Res 2:418–427. https://doi.org/10.5094/APR.2011.048

Pervez S, Chakrabarty RK, Dewangan S et al (2016) Chemical speciation of aerosols and air quality degradation during the festival of lights (Diwali). Atmos Pollut Res 7:92–99. https://doi.org/10.1016/j.apr.2015.09.002

Pieters J (2018) Netherlands to ban firecrackers before New Year's: report | NL Times. In: NL Times
Polash Mukerjee, Shambhavi Shukla AR (2016) Through the thick air. Down to earth
Queensland Government (2018) Cyclones and severe storms | emergency services and safety | Queensland Government. In: Queensl. Gov. https://www.qld.gov.au/emergency/dealing-disasters/disaster-types/cyclone-severe-storm
Rastogi N, Singh A, Satish R (2019) Characteristics of submicron particles coming from a big firecrackers burning event: implications to atmospheric pollution. Atmos Pollut Res 10:629–634. https://doi.org/10.1016/j.apr.2018.11.002
Rathi S, Niyogi S, Rathi K (2016) Air pollution: an initiative of Delhi government to curb air pollution by implementing odd-even car rationing. Int J Control Theory Appl 9:199–208
Samy A (2018) Fireworks in Germany: what you need to know about ending the year with a bang—the local. In: The Local
Sarkar S, Khillare PS, Jyethi DS et al (2010) Chemical speciation of respirable suspended particulate matter during a major firework festival in India. J Hazard Mater 184:321–330. https://doi.org/10.1016/j.jhazmat.2010.08.039
Sati AP, Mohan M (2014) Analysis of air pollution during a severe smog episode of November 2012 and the Diwali Festival over Delhi, India. Int J Remote Sens 35:6940–6954. https://doi.org/10.1080/01431161.2014.960618
SCI (2018) https://indiankanoon.org/doc/63024516/Supreme Court of India; Arjun Gopal vs Union Of India on 23 October, 2018. In: SCI. https://indiankanoon.org/doc/63024516/. Accessed 2 Mar 2021
Shivani GR, Saxena M et al (2019) Short-term degradation of air quality during major firework events in Delhi, India. Meteorol Atmos Phys 131:753–764. https://doi.org/10.1007/s00703-018-0602-9
Shukla S, Anisha R (2016) The gas chamber called Delhi. The Pioneer
Silver B, Reddington CL, Arnold SR, Spracklen DV (2018) Substantial changes in air pollution across China during 2015–2017. Environ Res Lett 13:114012. https://doi.org/10.1088/1748-9326/aae718
Singh A, Dey S (2012) Influence of aerosol composition on visibility in megacity Delhi. Atmos Environ 62:367–373. https://doi.org/10.1016/j.atmosenv.2012.08.048
Singh A, Pant P, Pope FD (2019) Air quality during and after festivals: aerosol concentrations, composition and health effects. Atmos Res 227:220–232
Singh DP, Gadi R, Mandal TK et al (2010) Study of temporal variation in ambient air quality during Diwali festival in India. Environ Monit Assess 169:1–13. https://doi.org/10.1007/s10661-009-1145-9
Srishti C (2016) CSIR scientists lead the way for Green Diwali, develop eco-friendly firecrackers. In: Mint. https://www.livemint.com/Science/2BaiCY0cE1WjhztFP1eDEL/CSIR-scientists-lead-the-way-for-Green-Diwali-develop-ecof.html. Accessed 22 Feb 2021
Srivastava A, Jain VK (2008) Source apportionment of suspended particulate matters in a clean area of Delhi: a note. Transp Res Part D Transp Environ 13:59–63. https://doi.org/10.1016/j.trd.2007.09.001
Statista (2019) India—share of Delhi's annual household income 2015 | Statistic. https://www.statista.com/statistics/658909/share-of-annual-income-in-delhi-india/. Accessed 25 Mar 2021
Times of India (2018) Firecracker Ban: Delhi Police register over 550 cases and arrest more than 300 on Diwali night for violating SC order | Delhi News—Times of India. In: Times of India. https://timesofindia.indiatimes.com/city/delhi/delhi-police-register-over-550-cases-and-arrest-more-than-300-on-diwali-night-for-violating-sc-order/articleshow/66545643.cms. Accessed 25 Mar 2021
Tiwari S, Chate DM, Srivastava MK et al (2012) Statistical evaluation of PM 10 and distribution of PM 1, PM 2.5, and PM 10 in ambient air due to extreme fireworks episodes (Deepawali festivals) in megacity Delhi. Nat Hazards 61:521–531. https://doi.org/10.1007/s11069-011-9931-4
TOI (2016) Supreme Court bans sale of firecrackers in Delhi, NCR—Times of India. In: Time India
TOI (2020) Supreme Court allows use of green crackers for 2 hours | India News—Times of India. Time India
Ursachi G, Horodnic IA, Zait A (2015) How reliable are measurement scales? External factors with indirect influence on reliability estimators. Procedia Econ Financ 20:679–686. https://doi.org/10.1016/s2212-5671(15)00123-9
Vecchi R, Bernardoni V, Cricchio D et al (2008) The impact of fireworks on airborne particles. Atmos Environ 42:1121–1132. https://doi.org/10.1016/j.atmosenv.2007.10.047
Verma C, Deshmukh DK (2014) The ambient air and noise quality in India during Diwali Festival : a review. Recent Res Sci Technol 6:203–210
Wen P, Thomas S (2018) China's millennium-old fireworks hub grapples with bans and shifting traditions—Reuters. In: Reuters
You ML, Shu CM, Chen WT, Shyu ML (2017) Analysis of cardinal grey relational grade and grey entropy on achievement of air pollution reduction by evaluating air quality trend in Japan. J Clean Prod 142:3883–3889. https://doi.org/10.1016/j.jclepro.2016.10.072

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Rendiconti del Circolo Matematico di Palermo Series 2
https://doi.org/10.1007/s12215-021-00695-7

Bézier variant of summation-integral type operators

Neha[1] · Naokant Deo[1] · Ram Pratap[2]

Received: 9 July 2021 / Accepted: 10 November 2021
© The Author(s), under exclusive licence to Springer-Verlag Italia S.r.l., part of Springer Nature 2022

Abstract
The motive of this article is to introduce the Bézier variant of a sequence of summation-integral type operators involving inverse Pólya-Eggenberger distribution and Păltănea operators [17]. For these operators, we estimate the approximation behaviour including first and second-order modulus of smoothness. Lastly, we establish the rate of convergence with a class of functions of derivatives of bounded variation.

Keywords Inverse Pólya-Eggenberger distribution · Rate of convergence · Modulus of continuity · Bounded variation

Mathematics Subject Classification 41A25 · 41A35

1 Introduction

In 1970, Stancu [19] presented the generalization of Baskakov operators using the concept of inverse Pólya-Eggenberger distribution (IPED) with subject to parameter $\beta \geq 0$. For $\varphi \in C[0, \infty)$, these operators are defined as follows:

✉ Ram Pratap
rampratapiitr@gmail.com

Neha
nehadtu03@gmail.com

Naokant Deo
naokantdeo@dce.ac.in

[1] Department of Applied Mathematics, Delhi Technological University, Bawana Road, Delhi 110042, India

[2] Department of Mathematics, Miranda House, University of Delhi, Delhi 110007, India

Published online: 20 January 2022

$$L_n^{(\beta)}(\varphi;x) = \sum_{r=0}^{\infty} \binom{n+r-1}{r} \frac{\prod_{i=0}^{r-1}(x+i\beta)\prod_{j=0}^{n-1}(1+j\beta)}{\prod_{k=0}^{n+r-1}(x+1+k\beta)} \varphi\left(\frac{r}{n}\right)$$
$$= \sum_{r=0}^{\infty} w_{n,r}^{(\beta)}(x)\varphi\left(\frac{r}{n}\right), \quad x \in [0,\infty). \tag{1.1}$$

In the current decade, several researchers proposed various operators and their variants based on PED as well as IPED and studied their approximation properties and interesting results. For more literature on this topic, one can refer to articles [6, 7, 9–12, 15].

In 2003, Srivastava-Gupta [18] introduced a widespread family of positive linear operators and contemplated their estimate properties and interesting results. In light of this work we consider summation-integral type operators involving inverse Pólya-Eggenberger distribution [19] and Păltănea operators [17].

For $\rho \geq 0$, we consider the integral-type operators, which is generalisation of the operators (1.1):

$$\hat{L}_n^{(\beta,\rho,c)}(\varphi;x) = \sum_{r=1}^{\infty} w_{n,r}^{(\beta)}(x) \int_0^{\infty} \chi_{n,r}^{\rho,c}(t)\varphi(t)dt + w_{n,0}^{(\beta)}(x)\varphi(0), \tag{1.2}$$

where

$$\chi_{n,r}^{\rho,c}(t) = \begin{cases} \frac{n\rho}{\Gamma(r\rho)} e^{-n\rho t}(n\rho t)^{r\rho-1}, & c = 0 \\ \frac{\Gamma\left(\frac{n\rho}{c}+r\rho\right)}{\Gamma(r\rho)\Gamma\left(\frac{n\rho}{c}\right)} \frac{c^{r\rho}t^{r\rho-1}}{(1+ct)^{\frac{n\rho}{c}+r\rho}}, & c \in \mathbb{N}, \end{cases}$$

and

$$\int_0^{\infty} \chi_{n,r}^{\rho,c}(t)t^j dt = \begin{cases} \frac{\Gamma(r\rho+j)}{\Gamma(r\rho)} \frac{1}{\prod_{i=1}^{j}(n\rho-ic)}, & j \neq 0 \\ 1, & j = 0. \end{cases}$$

The operators (1.2) also studied by Kajla et al. [15] for $c = 0$. Now we introduce the Bézier variant of (1.2) for $\theta \geq 1$ as follows:

$$\hat{L}_{n,\theta}^{(\beta,\rho,c)}(\varphi;x) = \sum_{r=1}^{\infty} w_{n,r}^{\theta,\beta}(x) \int_0^{\infty} \chi_{n,r}^{\rho,c}(t)\varphi(t)dt + w_{n,0}^{\theta,\beta}(x)\varphi(0), \tag{1.3}$$

where $w_{n,r}^{\theta,\beta}(x) = \left(J_{n,r}^{\beta}(x)\right)^{\theta} - \left(J_{n,r+1}^{\beta}(x)\right)^{\theta}$ and $J_{n,r}^{\beta}(x,c) = \sum_{j=r}^{\infty} w_{n,j}^{(\beta)}(x)$. It is obvious that the operators $\hat{L}_{n,\theta}^{(\beta,\rho,c)}(.;x)$ are the linear positive operators.

We may rewrite the operators (1.2) as:

$$\hat{L}_{n,\theta}^{(\beta,\rho,c)}(\varphi;x) = \int_0^{\infty} \mathcal{K}_{n,\theta}^{\beta,\rho}(x,t;c)\varphi(t)dt, \tag{1.4}$$

where $\mathcal{K}_{n,\theta}^{\beta,\rho}(x,t;c)$ is the kernel function given by

$$\mathcal{K}_{n,\theta}^{\beta,\rho}(x,t;c) = \sum_{r=1}^{\infty} w_{n,r}^{\beta,\theta}(x)\chi_{n,r}^{\rho}(t)dt + w_{n,0}^{\beta,\theta}(x)\delta(t), \tag{1.5}$$

where $\delta(t)$ is the Dirac delta function.

Special cases:

(1) For $c = 0$ and $\theta = 0$, the operators (1.3) reduce to Baskakov-Szász type operators based on inverse Polya-Eggenberger-distribution [15].
(2) For $c = \beta = 0$ and $\theta = \rho = 1$, the operators (1.3) include Baskakov-Szász operators (see [1, 13]).
(3) For $c = \beta = 0$, $\theta = 1$ and $\rho \to \infty$, the operators (1.3) reduce to Baskakov operators [3].
(4) For $c = 0$, $\theta = 1$, $\beta > 0$ and $\rho \to \infty$, the operators (1.3) include Stancu operators [19].

A lot of work has already been done on summation-integral type operators involving various linear positive operators and analysis of their convergence results. We refer to the readers some interesting articles (see [2, 4, 5, 14, 16]) for more information.

The vital target of the paper is to contemplate the approximation properties of operators (1.3) including K-functional and second-order modulus of smoothness. Lastly, we set up the rate of convergence for functions with a class of derivative of bounded variation.

2 Preliminaries

Lemma 2.1 *For $e_k(t) = t^k$, $k = 1$ and $n \in \mathbb{N}$, we have*

$$\hat{L}_n^{(\beta,\rho,c)}\big((e_1 - x);x\big) = \frac{(1-\beta)c + n\beta\rho}{(1-\beta)(n\rho - c)}x;$$

$$\hat{L}_n^{(\beta,\rho,c)}\big((e_1 - x)^2;x\big) = \frac{n\rho[1 + \rho + \beta\{(n-1)\rho - 2\}]}{(1-\beta)(1-2\beta)(n\rho - c)(n\rho - 2c)}x + \frac{2c^2(1 - 3\beta + 2\beta^2) + c(1 + \beta - 6\beta^2)n\rho + \{1 + \beta(1+2\beta)n\}n\rho^2}{(1-\beta)(1-2\beta)(n\rho - c)(n\rho - 2c)}x^2.$$

Further, for every $x \in [0,\infty)$, $\hat{L}_n^{(\beta,\rho,c)}\big((e_1 - x)^m;x\big)$ is a polynomial in x of degree m with

$$\hat{L}_n^{(\beta,\rho,c)}\big((e_1 - x)^m;x\big) = O\Big(\frac{1}{n^{-[(m+1)/2]}}\Big), \quad \text{as } n \to \infty.$$

Lemma 2.2 *If $\beta = \beta(n) \to 0$ as n is sufficient large $\lim_{n\to\infty} n\beta(n) = s \in \mathbb{R}$ then*

(i) $\lim_{n\to\infty} n\hat{L}_n^{(\beta,\rho,c)}\big((e_1 - x);x\big) = \Big(s + \frac{c}{\rho}\Big)x;$

(ii) $\lim_{n\to\infty} n\hat{L}_n^{(\beta,\rho,c)}\big((e_1 - x)^2;x\big) = \Big(1 + s + \frac{1}{\rho}\Big)x + \Big(1 + s + \frac{c}{\rho}\Big)x^2 \leqslant \Big(1 + s + \frac{c}{\rho}\Big)(1 + x)x.$

Lemma 2.3 *For every $\varphi \in C_B[0,\infty)$ then, we have*

$$\left|\hat{L}_{n,\theta}^{(\beta,\rho,c)}(\varphi(t);x)\right| \leq \theta \hat{L}_{n}^{(\beta,\rho,c)}(\|\varphi\|;x) \leq \theta\|\varphi\|.$$

Proof For $0 \leq l_1 \leq l_2 \leq 1, \theta \geq 1$, using the inequality

$$\left|l_1^{\beta} - l_2^{\beta}\right| \leq \theta\left|l_1 - l_2\right|,$$

from the definition of $w_{n,r}^{\theta,\beta}(x)$, we get

$$\begin{aligned} 0 &< \left(J_{n,r}^{\beta}(x)\right)^{\theta} - \left(J_{n,r+1}^{\beta}(x)\right)^{\theta} \\ &\leq \theta\left(J_{n,r}^{\beta}(x) - J_{n,r+1}^{\beta}(x)\right) \\ &= \theta J_{n,0}^{\beta}(x). \end{aligned}$$

Hence

$$\left|\hat{L}_{n,\theta}^{(\beta,\rho,c)}(\varphi(t);x)\right| \leq \theta \hat{L}_{n}^{(\beta,\rho,c)}(\|\varphi\|;x) \leq \theta\|\varphi\|.$$

□

3 Direct estimates

Suppose $C_B[0,\infty)$ denotes the class of continuous and bounded functions in $[0,\infty)$ with norm $\|\varphi\| = \sup_{x\in[0,\infty)} |\varphi(x)|$.

The usual modulus of smoothness for function $\varphi \in C_B[0,\infty)$ is defined as follows:

$$\omega\left(\varphi, \sqrt{\delta}\right) = \sup_{0<h\leq\sqrt{\delta}} \sup_{x\in[0,\infty)} |\varphi(x+h) - \varphi(x)|,$$

and let

$$\omega_2\left(\varphi, \sqrt{\delta}\right) = \sup_{0<h\leq\sqrt{\delta}} \sup_{x\in[0,\infty)} |\varphi(x+2h) - 2\varphi(x+h) + \varphi(x)|,$$

be the second order modulus of continuity and corresponding $K-$functional is defined as follows:

$$K_2(\varphi;\delta) := \inf_{x\in C_B^2[0,\infty)} \left\{\|\varphi - g\| + \delta\left\|g'\right\| + \delta^2\left\|g''\right\| : g \in C_B^2[0,\infty)\right\}, \quad \delta > 0,$$

where $C_B^2[0,\infty) = \left\{g \in C_B[0,\infty) : g', g'' \in C_B[0,\infty)\right\}$.

From DeVore and Lorentz([8], p.177. Theorem 2.4), the following inequality holds for $M > 0$:

$$K_2(\varphi;\delta) \leqslant M\omega_2\left(\varphi;\sqrt{\delta}\right). \tag{3.1}$$

Theorem 3.1 *For $\varphi \in C_B[0,\infty)$, we have*

$$\left|\hat{L}_{n,\theta}^{(\beta,\rho,c)}(\varphi;x)-\varphi(x)\right| \leq \left\{1+\left(\left(1+s+\frac{c}{\rho}\right)\theta x(1+x)\right)^{\frac{1}{2}}\right\}\omega\left(f;\frac{1}{\sqrt{n}}\right).$$

Proof By linearity property of $\hat{L}_{n,\theta}^{(\beta,\rho,c)}(.;x)$, we have

$$\begin{aligned}\left|\hat{L}_{n,\theta}^{(\beta,\rho,c)}(\varphi;x)-\varphi(x)\right| &\leq \int_0^\infty \mathcal{K}_{n,\theta}^{\beta,\rho}(x,t;c)|\varphi(t)-\varphi(x)|dt\\ &\leq \int_0^\infty \mathcal{K}_{n,\theta}^{\beta,\rho}(x,t;c)\left(1+\frac{|t-x|}{\delta}\right)dt,\end{aligned}$$

$\mathcal{K}_{n,\theta}^{\beta,\rho}(x,t;c)$ given in (1.5). Using Holder's inequality in right hand side of the above expression, we have

$$\begin{aligned}\left|\hat{L}_{n,\theta}^{(\beta,\rho,c)}(\varphi;x)-\varphi(x)\right| &\leq \left\{1+\frac{1}{\delta}\left(\hat{L}_{n,\theta}^{(\beta,\rho,c)}((t-x)^2;x)\right)^{\frac{1}{2}}\right\}\omega(\varphi;\delta)\\ &\leq \left\{1+\frac{1}{\delta}\left(\theta\hat{L}_{n}^{(\beta,\rho,c)}((t-x)^2;x)\right)^{\frac{1}{2}}\right\}\omega(\varphi;\delta)\end{aligned}$$

By using Lemma 2.2, Lemma 2.3 and choosing $\delta=\frac{1}{\sqrt{n}}$, we obtain the result. □

Theorem 3.2 *For $\varphi \in C_B[0,\infty)$, we have*

$$\left|\hat{L}_{n,\theta}^{(\beta,\rho,c)}(\varphi(t);x)-\varphi(x)\right| \leq M\omega_2\left(\varphi;\sqrt{\theta^{\frac{1}{2}}\delta_{n,\rho}^{c,s}(x)}\right),$$

where $\delta_{n,\rho}^{c,s}(x)=\sqrt{\hat{L}_{n}^{(\beta,\rho,c)}(t-x)^2;x)}$.

Proof Using Taylor's expansion for $g \in C_B^2[0,\infty)$ and $x,t\in[0,\infty)$, we get

$$g(t)=g(x)+(t-x)g'(x)+\int_x^t (t-u)g''(u)du.$$

Applying $\hat{L}_{n,\theta}^{(\beta)}$ on both sides of the above expansion, we get

$$\hat{L}_{n,\theta}^{(\beta,\rho,c)}(g(t);x)-g(x)=g'(x)\hat{L}_{n,\theta}^{(\beta,\rho,c)}((t-x);x)+\hat{L}_{n,\theta}^{(\beta,\rho,c)}\left(\int_x^t (t-u)g''(u)du;x\right).$$

Using Cauchy-Schwarz inequality, Lemma 2.2 and Lemma 2.3, we get

$$\begin{aligned}\left|\hat{L}_{n,\theta}^{(\beta,\rho,c)}(g(t);x)-g(x)\right| \leq & \left|g'(x)\hat{L}_{n,\theta}^{(\beta,\rho,c)}((t-x);x)\right| + \left|\hat{L}_{n,\theta}^{(\beta,\rho,c)}\left(\int_x^t (t-u)g''(u)du;x\right)\right| \\ \leq & \|g'\|\hat{L}_{n,\theta}^{(\beta,\rho,c)}(|t-x|;x) + \frac{\|g''\|}{2}\hat{L}_{n,\theta}^{(\beta,\rho,c)}\left((t-x)^2;x\right) \\ \leq & \|g'\|\left(\hat{L}_{n,\theta}^{(\beta,\rho,c)}\left((t-x)^2;x\right)\right)^{\frac{1}{2}} + \frac{\|g''\|}{2}\hat{L}_{n,\theta}^{(\beta,\rho,c)}\left((t-x)^2;x\right) \\ \leq & \sqrt{\theta}\|g'\|\delta_{n,\rho}^{c,s}(x) + \frac{\theta\|g''\|}{2}\left(\delta_{n,\rho}^{c,s}(x)\right)^2.\end{aligned} \tag{3.2}$$

Using Lemma 2.3 and (3.2), we obtain

$$\begin{aligned}\left|\hat{L}_{n,\theta}^{(\beta,\rho,c)}(\varphi(t);x)-\varphi(x)\right| \leq & \left|\hat{L}_{n,\theta}^{(\beta,\rho,c)}(\varphi(t)-g(t);x)\right| + \left|\hat{L}_{n,\theta}^{(\beta,\rho,c)}(g(t);x)-g(x)\right| + |\varphi(x)-g(x)| \\ \leq & (\theta+1)\|\varphi-g\| + \sqrt{\theta}\|g'\|\delta_{n,\rho}^{c,s}(x) + \frac{\theta\|g''\|}{2}\left(\delta_{n,\rho}^{c,s}(x)\right)^2\end{aligned}$$

Taking infimum on the right hand side of the above inequality over $g \in C_B^2$, we get

$$\left|\hat{L}_{n,\theta}^{(\beta,\rho,c)}(\varphi(t);x)-\varphi(x)\right| \leq (\theta+1)K_2\left(\varphi;\theta^{\frac{1}{2}}\delta_{n,\rho}^{c,s}(x)\right).$$

Using relation (3.1), we obtain the result. □

4 Rate of convergence

In this section we discuss the rate of convergence by means of the decomposition technique of functions with a derivative of bounded variation (DBV).

Let $\varphi \in DBV_\gamma \in [0,\infty)$, $\gamma \geqslant 0$ be the class of all functions defined on $[0,\infty)$, having a derivative of bounded variation on every finite subintervals of $[0,\infty)$ and $|\varphi(t)| \leqslant Mt^\gamma$, $M > 0$.

The function $\varphi \in DBV_\gamma[0,\infty)$, we may write

$$\varphi(x) = \int_0^x g(t)dt + \varphi(0),$$

where $g(t)$ is a function of bounded variation on each finite subinterval of $[0,\infty)$.

We denote the auxiliary function φ_x' by

$$\varphi_x'(t) = \begin{cases} \varphi'(t)-\varphi'(x-), & 0 \leqslant t < x; \\ 0, & t = x; \\ \varphi'(t)-\varphi'(x+), & 0 < t < \infty. \end{cases} \tag{4.1}$$

Lemma 4.1 *Let $\beta = \beta(n) \to 0$ as $n \to \infty$ and $\lim_{n\to\infty} n\beta(n) = s \in \mathbb{R}$. For all $x \in (0,\infty)$ and n is sufficiently large, we have*

(i) $\varsigma_{n,\theta}^{\beta,\rho}(x,y;c) = \int_0^y \mathcal{K}_{n,\theta}^{\beta,\rho}(x,t;c)dt \leqslant \frac{((1+s)\rho+c)}{n\rho}\frac{\theta x(1+x)}{(x-y)^2}, \quad 0 \leqslant y < x;$

(ii) $1-\varsigma_{n,\theta}^{\beta,\rho}(x,z;c) = \int_z^\infty \mathcal{K}_{n,\theta}^{\beta,\rho}(x,t;c)dt \leqslant \frac{((1+s)\rho+c)}{n\rho}\frac{\theta x(1+x)}{(z-x)^2}, \quad x < z < \infty.$

Proof In view of Lemma 2.2 and Lemma 2.3, first we prove (i)

$$\begin{aligned}\int_0^y \mathcal{K}_{n,\theta}^{\beta,\rho}(x,t;c)dt &\leqslant \int_0^y \frac{(x-t)^2}{(x-y)^2}\mathcal{K}_{n,\theta}^{\beta,\rho}(x,t;c)dt \\ &\leqslant \theta(x-y)^{-2}\hat{L}_n^{(\beta,\rho,c)}\big((e_1-x)^2;x\big) \leqslant \frac{((1+s)\rho+c)}{n\rho}\frac{\theta x(1+x)}{(x-y)^2}.\end{aligned}$$

The proof of (ii) is similar to (i), so we omit the details. □

Theorem 4.2 *Let $\varphi \in DBV[0,\infty)$, $\beta = \beta(n) \to 0$ as $n \to \infty$ and $\lim_{n\to\infty} n\beta(n) = s \in \mathbb{R}$. Then for every $x \in (0,\infty)$ and sufficient large n, we have*

$$\begin{aligned}\Big|\hat{L}_{n,\theta}^{(\beta,\rho,c)}(\varphi;x) - \varphi(x)\Big| \leq& \frac{\theta(\rho s+c)}{(\theta+1)n\rho}\big|\varphi'(x+) + \theta\varphi'(x-)\big|x \\ &+ \frac{\theta}{\theta+1}\big|\varphi'(x+) - \varphi'(x-)\big|\sqrt{\frac{\theta((1+s)\rho+c)x(1+x)}{n\rho}} \\ &+ \frac{\theta((1+s)\rho+c)x(1+x)}{n\rho x}\sum_{k=1}^{\lceil\sqrt{n}\rceil}\overset{x}{\underset{x-\frac{x}{k}}{V}}\varphi_{x'} + \frac{x}{\sqrt{n}}\overset{x}{\underset{x-\frac{x}{\sqrt{n}}}{V}}\varphi_{x'} \\ &+ \frac{\theta((1+s)\rho+c)x(1+x)}{n\rho x}\big|\varphi(2x) - \varphi(x) - x\varphi(x+)\big| \\ &+ \frac{\theta((1+s)\rho+c)x(1+x)}{n\rho}\sum_{k=1}^{\lceil\sqrt{n}\rceil}\overset{x+\frac{x}{k}}{\underset{x}{V}}\varphi_x' + \frac{x}{\sqrt{n}}\overset{x+\frac{x}{\sqrt{n}}}{\underset{x}{V}}(\varphi_x') + M(\gamma,c,r,x) \\ &+ \frac{\theta((1+s)\rho+c)x(1+x)}{n\rho x}|\varphi(x)| + \sqrt{\frac{\theta N(\rho,s,c)x(1+x)}{n}}|\varphi(x+)|.\end{aligned}$$

Proof Using the (1.2) of the operators $\hat{L}_{n,\theta}^{(\beta,\rho,c)}$, we have

$$\begin{aligned}\hat{L}_{n,\theta}^{(\beta,\rho,c)}(\varphi;x) - \varphi(x) &= \int_0^\infty \mathcal{K}_{n,\theta}^{\beta,\rho}(x,t;c)(\varphi(t)-\varphi(x))dt \\ &= \int_0^\infty \left(\int_x^t \mathcal{K}_{n,\theta}^{\beta,\rho}(x,t;c)\varphi'(u)du\right)dt\end{aligned} \tag{4.2}$$

From the relation (4.1) we can write

$$\begin{aligned}\varphi'(u) =& \varphi_x'(u) + \frac{1}{\theta+1}(\varphi'(x+) + \theta\varphi'(x-)) \\ &+ \frac{1}{2}\big(\varphi'(x+) + \theta\varphi'(x-)\big)\Big(\operatorname{sgn}(u-x) + \frac{\theta-1}{\theta+1}\Big) \\ &\times \delta_x(u)\big[\varphi'(u) - \big(\varphi'(x+) + \varphi'(x-)\big)\big],\end{aligned} \tag{4.3}$$

where

$$\delta_x(u) = \begin{cases} 1, & x = u \\ 0, & x \neq u. \end{cases}$$

It is easy to say that

$$\int_0^\infty \mathcal{K}_{n,\theta}^{\beta,\rho}(x,t;c) \int_x^t [\varphi'(u) - \frac{1}{2}(\varphi'(x+) + \varphi'(x-)]\delta_x(u)dudt = 0. \tag{4.4}$$

Now

$$\begin{aligned} P_{n,\theta}^{\beta,\rho}(x) &= \int_0^\infty \mathcal{K}_{n,\theta}^{\beta,\rho}(x,t;c) \int_x^t \frac{1}{\theta+1}(\varphi'(x+) + \theta\varphi'(x-))dudt. \\ &= \frac{1}{\theta+1}(\varphi'(x+) + \theta\varphi'(x-)) \int_0^\infty \mathcal{K}_{n,\theta}^{\beta,\rho}(x,t;c)(t-x)dt \\ &= \frac{1}{\theta+1}(\varphi'(x+) + \theta\varphi'(x-))\hat{L}_{n,\theta}^{(\beta,\rho,c)}(t-x;x), \end{aligned} \tag{4.5}$$

and

$$\begin{aligned} Q_{n,\theta}^{\beta,\rho}(x) &= \int_0^\infty \mathcal{K}_{n,\theta}^{\beta,\rho}(x,t;c) \int_x^t \frac{1}{2}(\varphi'(x+) + \theta\varphi'(x-))\Big(\operatorname{sgn}(u-x) + \frac{\theta-1}{\theta+1}\Big)dudt \\ &= \frac{1}{2}(\varphi'(x+) + \theta\varphi'(x-))\Bigg(-\int_0^x \mathcal{K}_{n,\theta}^{\beta,\rho}(x,t;c) \int_x^t \Big(\operatorname{sgn}(u-x) + \frac{\theta-1}{\theta+1}\Big)dudt \\ &\quad + \int_x^\infty \mathcal{K}_{n,\theta}^{\beta,\rho}(x,t;c) \int_x^t \Big(\operatorname{sgn}(u-x) + \frac{\theta-1}{\theta+1}\Big)\Bigg) \\ &\leq \frac{\theta}{\theta+1}(\varphi'(x+) + \theta\varphi'(x-)) \int_0^\infty \mathcal{K}_{n,\theta}^{\beta,\rho}(x,t;c)|t-x|dt \\ &\leq \frac{\theta}{\theta+1}(\varphi'(x+) + \theta\varphi'(x-))\Big(\hat{L}_{n,\theta}^{(\beta,\rho,c)}\big((e_1-x)^2;x\big)\Big)^{\frac{1}{2}}. \end{aligned} \tag{4.6}$$

By using Lemma 2.2 and Lemma 4.1, from (4.2)-(4.6), we have

$$\begin{aligned} \hat{L}_{n,\theta}^{(\beta)}(\varphi;x) - \varphi(x) \leq &\Big|A_{n,\theta}^{(\beta,\rho,c)}(\varphi';x) + B_{n,\theta}^{(\beta,\rho,c)}(\varphi';x)\Big| \\ &+ \frac{\theta(\rho s + c)}{(\theta+1)n\rho}\big|\varphi'(x+) + \theta\varphi'(x-)\big|x \\ &+ \frac{\theta}{\theta+1}\big|\varphi'(x+) - \varphi'(x-)\big|\sqrt{\frac{\theta((1+s)\rho + c)x(1+x)}{n\rho}}, \end{aligned} \tag{4.7}$$

where

$$A_{n,\theta}^{(\beta,\rho,c)}(\varphi';x) = \int_0^x \left(\int_x^t \varphi'_x(u)du \right) \mathcal{K}_{n,\theta}^{\beta,\rho}(x,t;c)dt,$$

and

$$B_{n,\theta}^{(\beta,\rho,c)}(\varphi';x) = \int_x^\infty \left(\int_x^t \varphi'_x(u)du \right) \mathcal{K}_{n,\theta}^{\beta,\rho}(x,t;c)dt.$$

To estimate $A_{n,\theta}^{(\beta,\rho,c)}(\varphi';x)$, using integration by parts and applying Lemma 4.1 with $y = x - \frac{x}{\sqrt{n}}$, we obtain

$$\begin{aligned}
A_{n,\theta}^{(\beta,\rho,c)}(\varphi';x) &= \left| \int_0^x \left(\int_x^t \varphi'_x(u)du \right) d_t \zeta_{n,\theta}^{\beta,\rho}(x,t;c) \right| \\
&= \left| \int_0^x \zeta_{n,\theta}^{\beta,\rho}(x,t;c)\varphi'_x(t)dt \right| \\
&\le \int_0^y |\varphi'_x(t)| \left| \zeta_{n,\theta}^{\beta,\rho}(x,t;c) \right| dt + \int_0^y |\varphi'_x(t)| \left| \zeta_{n,\theta}^{\beta,\rho}(x,t;c) \right| dt \qquad (4.8) \\
&\le \frac{\theta((1+s)\rho+c)x(1+x)}{n\rho} \int_0^y \bigvee_t^x \varphi'_x (x-t)^2 dt + \int_y^x \bigvee_t^x \varphi'_x dt \\
&\le \frac{\theta((1+s)\rho+c)x(1+x)}{n\rho} \int_0^{x-\frac{x}{\sqrt{n}}} \bigvee_t^x \varphi'_x (x-t)^2 dt + \frac{x}{\sqrt{n}} \bigvee_{x-\frac{x}{\sqrt{n}}}^x \varphi'_x.
\end{aligned}$$

Substituting $u = \frac{x}{x-t}$, we get

$$\begin{aligned}
\frac{\theta((1+s)\rho+c)x(1+x)}{n\rho} \int_0^{x-\frac{x}{\sqrt{n}}} \bigvee_t^x \varphi'_x (x-t)^2 dt &= \frac{\theta((1+s)\rho+c)x(1+x)}{n\rho x} \int_1^{\sqrt{n}} \bigvee_{x-\frac{x}{u}}^x \varphi'_x du \\
&\le \frac{\theta((1+s)\rho+c)x(1+x)}{n\rho x} \sum_{k=1}^{\lfloor\sqrt{n}\rfloor} \int_k^{k+1} \bigvee_{x-\frac{x}{k}}^x \varphi'_x du \qquad (4.9) \\
&\le \frac{\theta((1+s)\rho+c)x(1+x)}{n\rho x} \sum_{k=1}^{\lfloor\sqrt{n}\rfloor} \bigvee_{x-\frac{x}{k}}^x \varphi'_x.
\end{aligned}$$

From (4.8) and (4.9), we get

$$A_{n,\theta}^{(\beta,\rho,c)}(\varphi';x) = \frac{\theta((1+s)\rho+c)x(1+x)}{n\rho x} \sum_{k=1}^{\lfloor\sqrt{n}\rfloor} \bigvee_{x-\frac{x}{k}}^x \varphi'_x + \frac{x}{\sqrt{n}} \bigvee_{x-\frac{x}{\sqrt{n}}}^x \varphi'_x. \qquad (4.10)$$

We can write

$$B_{n,\theta}^{(\beta,\rho,c)}(\varphi';x) \leq \left| \int_x^{2x} \left(\int_x^t \varphi'_x(u)du \right) d_t(1-\zeta_{n,\theta}^{\beta,\rho}(x,t;c)) \right| + \left| \int_{2x}^{\infty} \left(\int_x^t \varphi'_x(u)du \right) d_t \mathcal{K}_{n,\theta}^{\beta,\rho}(x,t;c) \right|.$$

From the second part of the Lemma 4.1, we get

$$\mathcal{K}_{n,\theta}^{\beta,\rho}(x,t;c) = d_t((1-\zeta_{n,\theta}^{\beta,\rho}(x,t;c)) for t > x.$$

Hence

$$B_{n,\theta}^{(\beta,\rho,c)}(\varphi';x) = B_{n,\theta,1}^{(\beta,\rho,c)}(\varphi';x) + B_{n,\theta,2}^{(\beta,\rho,c)}(\varphi';x),$$

where

$$B_{n,\theta,1}^{(\beta,\rho,c)}(\varphi';x) = \left| \int_x^{2x} \left(\int_x^t \varphi'_x(u)du \right) d_t(1-\zeta_{n,\theta}^{\beta,\rho}(x,t;c)) \right|,$$

and

$$B_{n,\theta,2}^{\beta,\rho,c}(\varphi';x) = \left| \int_{2x}^{\infty} \left(\int_x^t \varphi'_x(u)du \right) d_t \mathcal{K}_{n,\theta}^{\beta,\rho}(x,t;c) \right|.$$

Using integration by parts, applying Lemma 4.1, $1-\zeta_{n,\theta}^{\beta,\rho}(x,t;c) \leq 1$ and taking $t = x + \frac{x}{u}$ successively,

$$\begin{aligned}
B_{n,\theta,1}^{(\beta,\rho,c)}(\varphi';x) &= \left| \int_x^{2x} \varphi'_x(u)du(1-\zeta_{n,\theta}^{\beta,\rho}(x,2x;c)) - \int_x^{2x} \varphi'_x(t)(1-\zeta_{n,\theta}^{\beta,\rho}(x,t;c))dt \right| \\
&\leq \left| \int_x^{2x} (\varphi'(u)-\varphi'(x+))du \right| \left| 1-\zeta_{n,\theta}^{\beta,\rho}(x,2x;c) \right| + \left| \int_x^{2x} \varphi'_x(t)(1-\zeta_{n,\theta}^{\beta,\rho}(x,t;c))dt \right| \\
&\leq \frac{\theta((1+s)\rho+c)x(1+x)}{n\rho x} |\varphi(2x)-\varphi(x)-x\varphi(x+)| \\
&\quad + \frac{\theta((1+s)\rho+c)x(1+x)}{n\rho} \int_{x+\frac{x}{\sqrt{n}}}^{2x} \frac{\overset{t}{V}_x \varphi'_x}{(t-x)^2} dt + \int_x^{x+\frac{x}{\sqrt{n}}} \overset{t}{V}_x \varphi'_x dt \\
&\leq \frac{\theta((1+s)\rho+c)x(1+x)}{n\rho x} |\varphi(2x)-\varphi(x)-x\varphi(x+)| \\
&\quad + \frac{\theta((1+s)\rho+c)x(1+x)}{n\rho} \sum_{k=1}^{\lceil\sqrt{n}\rceil} \overset{x+\frac{x}{k}}{V}_x \varphi'_x + \frac{x}{\sqrt{n}} \overset{x+\frac{x}{\sqrt{n}}}{V}_x (\varphi'_x).
\end{aligned} \tag{4.11}$$

Using Lemma 4.1, we have

$$\begin{aligned}
B_{n,\theta,2}^{(\beta,\rho,c)}(\varphi';x) =& \left| \int_{2x}^{\infty} \left(\int_{x}^{t} (\varphi'(u) - \varphi'(x+))du \right) \mathcal{K}_{n,\theta}^{\beta,\rho}(x,t;c)dt \right| \\
&\leq \int_{2x}^{\infty} |\varphi(t) - \varphi(x)| \mathcal{K}_{n,\theta}^{\beta,\rho}(x,t;c)dt + \int_{2x}^{\infty} |t-x||\varphi(x+)| \mathcal{K}_{n,\theta}^{\beta,\rho}(x,t;c)dt \\
&\leq \left| \int_{2x}^{\infty} \varphi(t) \mathcal{K}_{n,\theta}^{\beta,\rho}(x,t;c)dt \right| + |\varphi(x)| \left| \int_{2x}^{\infty} \mathcal{K}_{n,\theta}^{\beta,\rho}(x,t;c)dt \right| \\
&\quad + |\varphi(x+)| \left(\int_{2x}^{\infty} (e_1 - x)^2 \mathcal{K}_{n,\theta}^{\beta,\rho}(x,t;c)dt \right)^{\frac{1}{2}} \\
&\leq M \int_{2x}^{\infty} t^{\gamma} \mathcal{K}_{n,\theta}^{\beta,\rho}(x,t;c)dt + |\varphi(x)| \left| \int_{2x}^{\infty} \mathcal{K}_{n,\theta}^{\beta,\rho}(x,t;c)dt \right| \\
&\quad + \sqrt{\frac{\theta((1+s)\rho + c)x(1+x)}{n\rho}} |\varphi(x+)|.
\end{aligned} \tag{4.12}$$

For $t \geq 2x$, we get $t \leq 2(t-x)$ and $x \leq t-x$, applying Hölder's inequality, we have

$$\begin{aligned}
B_{n,\theta,2}^{(\beta,\rho,c)}(\varphi';x) \leq& M 2^{\gamma} \left(\int_{2x}^{\infty} (e_1 - x)^{2r} \mathcal{K}_{n,\theta}^{\beta,\rho}(x,t;c)dt \right)^{\frac{\gamma}{2k}} \\
&+ \frac{\theta((1+s)\rho + c)x(1+x)}{n\rho x} |\varphi(x)| + \sqrt{\frac{\theta((1+s)\rho + c)x(1+x)}{n\rho}} |\varphi(x+)| \\
=& M(\gamma, c, k, x) + \frac{\theta((1+s)\rho + c)x(1+x)}{n\rho x} |\varphi(x)| \\
&+ \sqrt{\frac{\theta((1+s)\rho + c)x(1+x)}{n\rho}} |\varphi(x+)|.
\end{aligned} \tag{4.13}$$

From (4.11) and (4.13), we get

$$
\begin{aligned}
B_{n,\theta}^{(\beta,\rho,c)}(\varphi';x) =& \frac{\theta((1+s)\rho+c)x(1+x)}{n\rho x}|\varphi(2x)-\varphi(x)-x\varphi(x+)| \\
&+ \frac{\theta((1+s)\rho+c)x(1+x)}{n\rho}\sum_{r=1}^{\left[\sqrt{n}\right]} \overset{x+\frac{x}{r}}{\underset{x}{V}} \varphi'_x + \frac{x}{\sqrt{n}} \overset{x+\frac{x}{\sqrt{n}}}{\underset{x}{V}} (\varphi'_x) \\
&+ M(\gamma,c,r,x) + \frac{\theta((1+s)\rho+c)x(1+x)}{n\rho x}|\varphi(x)| \\
&+ \sqrt{\frac{\theta((1+s)\rho+c)x(1+x)}{n\rho}}|\varphi(x+)|.
\end{aligned} \tag{4.14}
$$

From (4.7), (4.10) and (4.14), we get our desired result. □

References

1. Abel, U., Gupta, V., Ivan, M.: The complete asymptotic expansion for Baskakov-Szász-Durrmeyer operators. Ann. Tiberiu Popoviciu Semin. Funct.Eqn. Approx. Convexity **1**, 3–15 (2003)
2. Acu, A.M., Gupta, V.: Direct results for certain summation-integral type Baskakov-Szász operators. Results. Math. **72**, 1161–1180 (2017)
3. Baskakov, V.A.: An instance of a sequence of linear positive operators in the space of continuous functions. Dokl. Akad. Nauk. **113**, 249–251 (1957)
4. Deo, N.: Faster rate of convergence on Srivastava-Gupta operators. Appl. Math. Comput. **218**(21), 10486–10491 (2012)
5. Deo, N.: A note on equivalence theorem for Beta operators. Mediterr. J. Math. **4**(2), 245–250 (2007)
6. Deo, N., Dhamija, M., Miclăuş, D.: New modified Baskakov operators based on the inverse Pólya-Eggenberger distribution. Filomat **33**(11), 3537–3550 (2019)
7. Deo N., Dhamija M.: Generalized positive linear operator based on PED and IPED. Iran. J. Sci. Technol. Trans. A Sci. **43**, 507–513 (2018)
8. DeVore, R.A., Lorentz, G.G.: Constructive approximation. Springer, Berlin (1993)
9. Dhamija, M., Deo, N.: Approximation by generalized positive linear-Kantorovich operators. Filomat **31**(14), 4353–4368 (2017)
10. Dhamija, M., Deo, N.: Jain-Durrmeyer operators associated with the inverse Pólya-Eggenberger distribution. Appl. Math. Comput. **286**, 15–22 (2016)
11. Dhamija, M., Pratap, R., Deo, N.: Approximation by Kantorovich form of modified Szász-Mirakyan operators. Appl. Math. Comput. **317**, 109–120 (2018)
12. Gupta, V., Rassias, T.M.: Lupaş-Durrmeyer operators based on Pólya distribution. Banach J. Math. Anal. **8**(2), 145–155 (2014)
13. Gupta, V., Srivastava, G.S.: Simultaneous approximation by Baskakov-Szász type operaators. Bull. Math. Sci. Math. Roumaine (N S) **37**(85), 73–85 (1993)
14. Gupta, V., Vasishtha, V., Gupta, M.K.: Rate of convergence of summation-integral type operators with derivatives of bounded variation. J. Inequal. Pure Appl. Math. **4**(2), 1–8 (2003)
15. Kajla, A., Acu, A.M., Agrawal, P.N.: Baskakov-Szász type operators based on inverse Pólya-Eggenberger distribution. Ann. Funct. Anal. **8**(1), 106–123 (2017)
16. Pratap, R., Deo, N.: Rate of convergence of Gupta-Srivastava operators with certain parameters. J. Class. Anal. **14**(2), 137–153 (2019)
17. Păltănea, R.: Modified Szász-Mirakjan operators of integral form. Carpathian J. Math. **24**(3), 378–385 (2008)
18. Srivastava, H.M., Gupta, V.: A certain family of summation-integral type operators. Math. Comput. Model. **37**, 1307–1315 (2003)
19. Stancu, D.D.: Two classes of positive linear operators. Anal Univ Timişoara Ser. Ştin. Matem. **8**, 213–220 (1970)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/MCE.2021.3140048, IEEE Consumer Electronics Magazine

Blockchain-Based Medical Certificate Generation and Verification for IoT-based Healthcare Systems

Suyel Namasudra, *Member, IEEE*, Pratima Sharma, Ruben Gonzalez Crespo, *Senior Member, IEEE,* and Vimal Shanmuganathan, *Senior Member, IEEE*

***Abstract*—Nowadays, medical certificates are very important for many users as they want to avail health benefits like tax purposes, insurance claims, legal procedures, and many more. Generating, issuing, and maintaining medical certificates remain a significant problem before the invention of the computer, they were available as hard copies. The digitization of medical certificates and documents leads to potential security issues, such as forging of certificates risk the privacy of healthcare documents. Moreover, individuals still need to be physically present and wait at the issuing healthcare centers to get the certificates. Currently, the infrastructure of any healthcare industry connects the Internet of Things (IoT) devices and application software that communicates with the Information Technology (IT) systems. Blockchain technology with IoT can significantly affect the healthcare industry by improving efficiency, security, transparency, and can provide more business opportunities. Therefore, a privacy-preserving technique has been proposed in this paper for IoT-based healthcare systems using blockchain technology. The proposed architecture provides an interface between the users and healthcare centers to generate and maintain health documents. Furthermore, the proposed scheme ensures security by specifying rules with a smart contract. Results and discussion show that the proposed scheme is more efficient than the existing schemes.**

***Index Terms*—Electronic Health Record, Integrity, InterPlanetary File System, Gas, BlockSim.**

I. Introduction

ACCORDING to the World Health Organization (WHO), medical certificates and documents are significantly important to describe a person's medical condition, indicate eligibility of activity, make an insurance claim, obtain benefits from the employer, and for tax or legal procedures. Birth certificate, sick certificate, medical certificate, death certificate, and fit-to-fly certificate related to COVID-19, are some important documents required for personal, business, official, legal and other purposes. These medical documents have been available in digital form in various parts of the world, and these documents are widely used in the form of hard copies. These medical certificates or documents must be always reliable as these are directly related to the patients' health conditions. However, the traditional centralized generation systems are prone to forgery, and there is no mechanism that verifies the authenticity of the user and validates the integrity of the generated certificate. Thus, the IoT-based medical system has been introduced to significantly enhance security and precision, and overcome geographical constraints for the remote tracking and maintaining health documents. The IoT defines smart devices, such as sensors, surveillance, etc., that can sense, maintain, and share information with the other connecting devices. These services can be achieved locally or remotely by using cloud systems. The current healthcare practices for managing medical records are vulnerable to several cybersecurity threats, including malware and ransomware. The healthcare data can be compromised by deploying ransomware, which encrypts the data until a ransom is paid by hacking the repository. The WannaCry ransomware attack on National Health Service (NHS) infected millions of servers and computers, resulting in a loss of healthcare records [1]. Furthermore, the malware attacks compromise unprotected medical devices, breach user identification, and access medical records. Therefore, the healthcare domain suffers from substantial loss of health records, which can cause financial loss, as well as even death of patients.

Therefore, to address the challenges mentioned above, modern technologies are required to secure and preserve the privacy of healthcare records. Since the last decade, the security of healthcare certificates has been innovated through many platforms, software, and communication technologies. In conventional healthcare approaches, researchers use encryption and embedding algorithms to secure sensitive medical data using cryptography techniques [2]. These methods, however, fail to protect the records. In addition, database management systems [3] or cloud management systems [4] are used to maintain medical credentials in most of the existing systems. These methods are ineffective to handle many attacks like man-in-the-middle attacks,

S. Namasudra is with the Department of Computer Science and Engineering, National Institute of Technology Patna, Bihar, India (Email: suyelnamasudra@gmail.com).

P. Sharma is with the Department of Computer Science and Engineering, Delhi Technological University, Delhi, India.

R. G. Crespo and S. Namasudra are with the International University of La Rioja (UNIR), Spain.

V. Shanmuganathan is with the Department of Artificial Intelligence and Data Science, Ramco Institute of Technology, Rajapalayam, Tamilnadu, India.

2162-2248 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 09:16:29 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/MCE.2021.3140048, IEEE Consumer Electronics Magazine

masquerade attacks, modification attacks, and many more [5]. Also, distinctive cutting-edge innovations in the healthcare sector optimize performance and concentrate on privacy preservation of health-related documents. Nowadays, it is very easy for the attacker to duplicate a medical certificate or impersonate the user identity. The user must carry the certificates all the time to access them. Unfortunately, if they are missed, it is a lengthy procedure to retrieve them back. Frauds and the illegitimate generation and maintenance of medical certificates are other problems of healthcare systems. Thus, there is a need for a scalable, immutable, transparent, and secure solution to protect healthcare certificates.

Blockchain-based applications can easily solve the above-mentioned problems. Blockchain technology offers a robust distributed ledger, which was initially proposed by Nakamoto [6]. It is utilized in cryptocurrencies for validating or verification transactions, such as Ether and Bitcoin. A transaction stores the user's information in a blockchain network to connect to a ledger and transmit information or data to other nodes, i.e. customers or users. A blockchain network removes the centralized authority; therefore, it removes a single point of failure issue. Due to the intrinsic characteristics of blockchain, such as immutability, openness, credibility assessment, distribution, and decentralization, it is used in many fields like banking, e-voting, supply chain, IoT, and many more [7-9].

Many researchers have already developed many blockchain-based healthcare systems [10-13]. For example, Kumar et al. [14] have designed an IoT-based healthcare system to efficiently handle heterogeneous healthcare data using a proportionate data analytics method. This analytical method differentiates the healthcare data streams based on service responses using a linear regression method. But, the designed system lacks the details of managing medical certificates. Ali et al. [15] have proposed an IoT-enabled medical ecosystem with fog-assisted communication infrastructure to provide reliable healthcare services. It minimizes the cost using integer programming methods to secure electronic medical records in the healthcare system. Their scheme is very complicated due to various components, such as mathematical models, fog computing, IoT, healthcare system, etc. Mishra et al. [16] have developed a decentralized application using blockchain to share the students' credentials securely. They have also measured the performance of the distributed application using the execution cost and transaction cost as the parameters. However, the throughput of this scheme is high. Jennath et al. [17] have designed a blockchain-based healthcare system to address the security and privacy issues present in the e-health system. Their designed scheme applies the artificial intelligence models on the blockchain network to provide a transparent healthcare data-sharing platform. Similarly, Raut et al. [18] have proposed an enhanced bat algorithm to forecast COVID-19 cases and reduce the error of forecasting. Sharma et al. [19] proposed a blockchain-based IoT architecture to enhance the security of Electronic Health Records (EHR) using an identity-based encryption algorithm. This scheme utilizes the functionality of smart contracts to store, maintain, and secure healthcare records using the Swarm decentralized storage system at the backend. It also provides the access control mechanism and integrity checking method to provide a privacy-preserving environment. However, it lacks the functionality to verify the stored records from the healthcare centers or authorities. Rosa et al. [20] have presented the operational phases of a distributed application in multi-native domain networking. The authors have shown only the designing part of their proposed framework. The performance of the proposed work is also discussed by taking into account specific usage contexts. Wang et al. [21] have proposed a blockchain-based parallel healthcare system based on artificial intelligence, computational, and parallel techniques. It utilizes the consortium blockchain platform to provide healthcare data sharing and audit services. This scheme performs diagnosis and treatment simultaneously using healthcare records. Furthermore, Khatoon [22] has used a blockchain system for the healthcare sector. Here, an application has been designed to manage medical procedures like clinical trials and surgery. It implements the permissionless system using the Ethereum public blockchain network, and this scheme only considers clinical and surgery-related healthcare data. Daisuke et al. [23] have developed a methodology to preserve the medical records using the Hyperledger blockchain network, a permissioned blockchain. Their proposed architecture collects medical records using smartphones and stores them on the blockchain network. Also, Salem et al. [24] have proposed a computational intelligence and knowledge engineering-based smart learning system. This scheme utilizes the benefits of intelligent paradigms to increase the efficiency of the smart learning system, and it lacks the integrity checking mechanism. In [25], the authors have designed a hybrid framework by utilizing blockchain and deep learning technologies. This designed model uses authentication and lattice-based cryptography to provide a secure mode for sharing electronic health records. It also predicts future diseases based on current indicators using a deep-learning as a service model. Although many architectures have been proposed to use blockchain for managing health records [26-29], there is still a need to develop a blockchain-based application for the IoT-enabled healthcare system to provide a secure environment to generate and maintain medical certificates.

This paper proposes a blockchain-based technique based on Mobile Application (MA) for IoT-enabled healthcare systems to provide a privacy-preserving environment. The proposed system supports the prevention of improper access to medical certificates and their preservation. Furthermore, it also speeds up verifying a physical certificate and prevents unauthorized access to records of birth, death, and sick leave.

The key contributions of the paper are summarized below:

- This work proposes a blockchain-based application to generate, and maintain medical certificates on IoT devices. The proposed system utilizes the Ethereum public blockchain network with the Proof of Stack consortium

2162-2248 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 09:16:29 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/MCE.2021.3140048, IEEE Consumer Electronics Magazine

algorithm.

- The proposed blockchain structure stores and maintains the generated certificates on InterPlanetary File System (IPFS). After storage, the IPFS returns a unique hash of the certificate with a unique transaction ID. The users can use or show their unique ID received from MA, whenever they want to access the corresponding certificate.
- Results and discussion show the effectiveness of the proposed approach.

The rest of the paper is organized as follows. The proposed scheme is discussed in section II. Section III discusses results and discussion. Finally, the conclusions and future work are presented in section IV.

II. Proposed Architecture

This paper proposes a blockchain-based application to generate and maintain healthcare certificates and documents. Here, through various IoT devices, users may maintain their health certificates. These devices can be any device that the user has and do not need to be any specific device. The users' data collected from the smart devices are recorded and stored on the proposed architecture in the form of certificates. In addition, the proposed application helps to prevent fraud by issuing and maintaining users' medical certificates, such as birth, death, and sick leaves. The proposed application is executed on a distributed network, and initially, all healthcare centers and users register in the MA and receive public and private key pairs by executing the registration function of the proposed scheme. If any user requests the healthcare centers to generate the healthcare certificate, the authorities first verify users' details stored on the blockchain network using the MA interface. Then, the blockchain network generates a certificate using the details shared by the user during the registration process. Next, the generated certificate is stored on the IPFS decentralized storage system, and the returned hash is shared with the user. Finally, the health document-related details are managed as a transaction with a unique ID in a specific block of the blockchain network. The entire process is transparent due to the blockchain's structure as miners mine every transaction as a specialized node in the blockchain network.

Fig. 1 presents the blockchain-based architecture to generate and maintain medical certificates, and documents. The proposed architecture consists of five components: users, healthcare centers, blockchain network, MA, and IPFS. Initially, users or healthcare centers are registered in the blockchain network using the MA. The healthcare center sends a key generation request to the MA. The MA sends a request to the blockchain network and executes the key generation function to generate public and private key pairs for the healthcare center. The blockchain network stores the healthcare center's details and responds with public-private key pair. Then, the MA shares the generated key pair with the healthcare center. Similarly, the users are also registered in the blockchain network and receive public-private key pairs by submitting necessary details, such as user name, address, etc. Next, the users may request a medical certificate from the corresponding registered healthcare center, if required. The corresponding healthcare center verifies the details of the user using the MA. After verification, the certificate is generated by the healthcare center. Also, the hash of the generated certificate is mapped with the combined public keys of the user and the issuing healthcare center. The public keys of the healthcare center and the user are combined using a hash function to generate the combined key. The combined key, certificate hash, and transaction hash value are stored on the blockchain network. Using the generated transaction hash, the issuing healthcare center can check the validity of the certificate, as well as the authenticity of the user. Furthermore, the user digitally signs the generated certificate using the private key of the user and stores it on the IPFS storage system using the MA. The user decrypts the stored certificate using the public key of the user and utilizes it for further use. The user may verify the integrity of the certificate using the hash values stored on the blockchain network using the MA. Table I shows the details of notations used in this paper.

TABLE I
Notations with their Description

Notation	Description
$User$	Registered user in the network
$Healthcare_{center}$	Registered healthcare center in the blockchain
$Uname$	Name of the user registered in the network
$Address$	Address detail of the registered user
$Blockchain_{network}$	Blockchain network
$Process_{request}$	Function to process request send by the MA
PK_h	Public key of the healthcare center
SK_h	Private key of the healthcare center
PK_u	Public key of the user
SK_u	Private key of the user
$Generate()$	Function to generate public-private key pair
$Share()$	Function to share generated details with the user
$Medical_{certificate}$	Medical certificate generated by the network
$Certificate_{hash}$	Hash of the generated certificate
$Hash()$	Function to generate hash
CK	Combined key generated from the public keys of the user and the healthcare center
$Map()$	Function to map details of a generated certificate with the healthcare center and user
$Generate_{block}()$	Function to generate a block of the blockchain network
$TransID$	Transaction ID
$Digitally_{sign}()$	Function to digitally sign the generated certificate
$Store()$	Function to store the certificate on the decentralized storage system
$IPFS()$	Function to store medical certificate on the decentralized storage system
$Store_{hash}$	Stored hash value of the certificate

The proposed blockchain-based application consists of mainly four smart contract functions, i.e. $Register()$, $Generate_{block}()$, $Issue_Certificate()$, and $Verify_Certificate()$. Algorithm 1 represents the sequential

2162-2248 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 09:16:29 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/MCE.2021.3140048, IEEE Consumer Electronics Magazine

steps or processes for the registration of users in the proposed architecture. First, the healthcare center sends the registration request to the MA, i.e. proposed mobile application with the details like user name and address. Then, the MA verifies the user request, if the user is a healthcare authority/center or an individual end-user. After the verification, the MA sends a request to the blockchain network to generate the public and private key pair. The blockchain network processes the request and generates the public and private key pair, and shares it with the user.

Algorithm 1: Registration of users ($Register()$)
Input: User details
Output: Share generated key pair with the user

1. if $(User == Healthcare_{center})$
2. $\quad Healthcare_{center}(Uname, Address) \xrightarrow{request} MA$
3. $\quad MA \xrightarrow{request} Blockchain_{network}$
4. $\quad Blockchain_{network}(Process_{request}) \rightarrow Generate(PK_h, SK_h)$
5. $\quad Share(PK_h, SK_h) \rightarrow Healthcare_{center}$
6. else
7. $\quad User(Uname, Address) \xrightarrow{requst} MA$
8. $\quad MA \xrightarrow{request} Blockchain_{network}$
9. $\quad Blockchain_{network}(Process_{request}) \rightarrow Generate(PK_u, SK_u)$
10. $\quad Share(PK_u, SK_u) \rightarrow User$

Algorithm 2 shows the step-by-step processes to generate medical certificates for the user by using the blockchain functionality. At first, the user sends a request to the healthcare center with the mentioned details. Then, the healthcare center verifies the user's details. When verifying the request, the healthcare center generates a medical certificate and calculates the unique hash of the generated certificate using the SHA-256 hashing algorithm. Next, the healthcare center shares the generated certificate and its hash value with the user. The blockchain network combines the public keys of the user with the healthcare center to obtain a combined key. Then, the proposed network maps the certificate's details with the combined key and certificate's hash value to generate a block for the blockchain network. Finally, the generated block is added to the blockchain network and returns the ID of the transaction.

Algorithm 2: Block generation ($Generate_{block}()$)
Input: User details
Output: Return transaction ID of the generated block

1. $User \xrightarrow{request} Healthcare_{center}$
2. $Healthcare_{center} \xrightarrow{verify} User$
3. $Healthcare_{center} \xrightarrow{generate} Medical_{certificate}$
4. $Certificate_{hash} \leftarrow Hash(Medical_{certificate})$
5. $Healthcare_{center} \xrightarrow{share} User(Medical_{certificate}, Certificate_{hash})$
6. $Blockchain_{network} \xrightarrow{combine} CK(PK_h, PK_u)$
7. $Blockchain_{network} \rightarrow Map(Medical_{certificate}, CK, Certificate_{hash})$
8. $Blockchain_{network} \rightarrow Generate_{block}(Map)$
9. $Return\ TransID$

Algorithm 3 shows the processes of storage of a medical certificate on the proposed architecture. Here, at first, the user digitally signs the obtained certificate using the private key. Then, the user utilizes the MA interface to store the encrypted certificate on the decentralized storage system via the blockchain network. In response, the blockchain stores the medical certificate on the IPFS storage system. After the successful storage of the certificate, IPFS returns the unique ID. The blockchain network shares this unique ID with the user. The user may utilize the shared ID to access the medical certificate at any time.

Algorithm 3: Issue certificate ($Issue_Certificate()$)
Input: User details
Output: Hash of the certificate

1. $User \rightarrow Digitally_{sign}(Medical_{certificate}, SK_u)$
2. $Store(Medical_{certificate}) \rightarrow MA$
3. $MA \xrightarrow{request} Blockchain_{network}$
4. $Blockchain_{network} \xrightarrow{store} IPFS(Medical_{certificate})$
5. $IPFS \xrightarrow{return} hash$
6. $Share(hash) \rightarrow User$

Algorithm 4 shows the verification process to check the integrity of the stored medical certificates on the blockchain network. The user executes the verification process to check the integrity of the medical certificate using the hash value shared by the healthcare center during the certificate generation process. The user sends a verification request to the MA with the certificate's hash value. The MA fetches the mapping details of the certificate using the hash value and obtains the stored hash value from the blockchain network. Then, the MA verifies the integrity of the certificate by comparing the stored hash value with the hash shared by the user. If both hash values are matched, then, the integrity of the certificate is maintained; otherwise, returns a validity fail message to the user.

Algorithm 4: Verify certificate ($Verify_Certificate()$)
Input: Certificate details
Output: Integrity verification (Yes/No)

1. $User(Certificate_{hash}) \xrightarrow{integrity_check} MA$
2. $MA \xrightarrow{fetch} Store_{hash}$
3. if$(Certificate_{hash} == Store_{hash})$
4. $\quad Integrity\ maintained$
5. else
6. $\quad Integrity\ verification\ failed$

III. Results and Discussion

The proposed architecture is developed using the Ethereum blockchain network. Ethereum is a public network, which is used for developing smart contracts to implement various operations of the MA. The Ethereum network links to MetaMask, a browser plugin that creates Ethereum addresses for the users to send and sign transactions in the proposed architecture. The proposed application also handles illegal manipulation, discrepancies, and unauthorized access to the healthcare certificate. In the experiment, the BlockSim toolkit [30] provides the environment for simulating the blockchain network, deploying smart contracts, and executing tests or experiments for evaluating the performance parameters, such as latency, processing time, throughput, and computation time.

Many experiments are performed to measure the performance of the proposed scheme. The performance

2162-2248 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 09:16:29 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/MCE.2021.3140048, IEEE Consumer Electronics Magazine

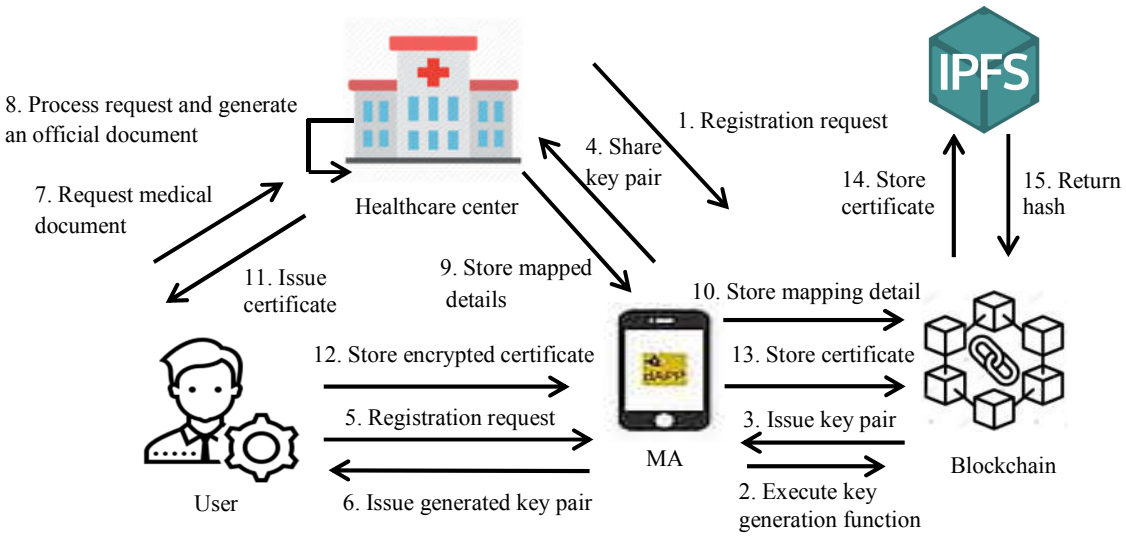


Fig. 1. Proposed architecture

TABLE II
SMART CONTRACT DEPLOYMENT COST

Function	Amount of Gas used	TxN fee (Gas price)	TxN hash	Block size (bytes)	TxN nonce	TxN index
$Register()$	30100 (92.55%)	0.000025	0x5b6db9b7f557853257ad65687764cf7f61234d569c591363bb5a151e145b	9995	3	4
$Generate_{block}()$	33600 (95.22%)	0.000032	0x2f4fd7d2b448422345bf22385764db55d589d582343bd5a262e125d	8125	5	7
$Issue_Certificate()$	38920 (98.24%)	0.000045	0x6a7dd8d8b568974378df78215764cf7f61352b567c245363bd5a251b122b	8822	4	6
$Verify_Certificate()$	36456 (97.24%)	0.000052	0x85ffe90fd223e3f43f34f43d245f1bb87f623f4d234b7c79cb04b22	5998	7	3

analysis of the proposed architecture is measured using the BlockSim toolkit. BlockSim provides the simulation framework to analyze the Ethereum-based architecture. It is an analytical platform for Ethereum that can explore a block in detail. It acts like an Ethereum gas tracker and can track the transactions, check the performance of the smart contract, and verify the state of the process. Each transaction collects some Gas as a cost of operation to run and execute on the Ethereum-based public blockchain. Gas is a unit of cost to perform a specific operation over a blockchain. Miners set the Gas price based on the supply and demand ratio. A lot of computational power is required to process smart contracts and transactions over a blockchain network.

Table II shows the operations costs of the proposed architecture by using the BlockSim, and the details of the transactions have been collected from the BlockSim toolkit. Here, Ethereum Gas pays Ether as a crypto-fuel to run the applications over the blockchain network. This execution cost depends on the internal storage cost of the smart contracts and manipulations on the blockchain. Transaction cost depends on the execution cost and contract deployment and transferring data into the blockchain. Generally, Gas has two parameters, i.e. limit and price. The Gas limit depends on the user's willingness to pay for performing a transaction and the minimum Gas limit is 21000. The Gas price depends on the user's willingness to spend on each Gas unit and is represented as 'Gwei'.

Table III shows the gas cost to generate the number of certificates using the proposed architecture, which uses the Metamask wallet. For the different number of certificates generation, the cost of executing the corresponding smart contract, such as $Generate_{block}()$, is analyzed based on the Gas limit, cost, and price. Here, the miners set the gas price depending on the supply and demand of the computational power of the blockchain network to process smart contracts. It is observed that the proposed architecture's operation cost increases by increasing the number of certificates. The Gas cost of these operations has been retrieved by deploying the smart contracts of related transactions over the blockchain network through MA. This experiment uses the BlockSim toolkit to execute the smart contract functions and analyzes the cost required for each operation.

2162-2248 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 09:16:29 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/MCE.2021.3140048, IEEE Consumer Electronics Magazine

TABLE III
GAS COST WITH DIFFERENT NUMBER OF CERTIFICATES

No. of certificates	Gas limit (units)	Gas cost	Gas price (CGWEI)
1	22645	147633	100
10	218730	1389224	1000
20	419660	2752440	2000
30	619290	4128660	3000
40	825720	5504880	4000
50	1032150	6881100	5000
100	2064300	13762200	10000

Fig. 2 presents the total Gas cost analysis of the proposed architecture by varying the certificate generation and verification transactions in the blockchain network. Gas consumption is measured in the unit of Eths. The total Gas consumption of the proposed architecture is calculated by varying the number of certificates ranging from 20 to 160. Here, the certificate generation and verification transactions are executed randomly using the BlockSim toolkit. It is observed that the proposed architecture requires more Gas consumption for the certificate generation transactions than the verification transactions. This is because it requires more processing power for calculating the hash of the generated certificate, mapping the user details with the certificate, and generating a new block. Thus, the Gas consumption is higher due to a large number of computations in the certificate generation transaction compared to the verification transaction.

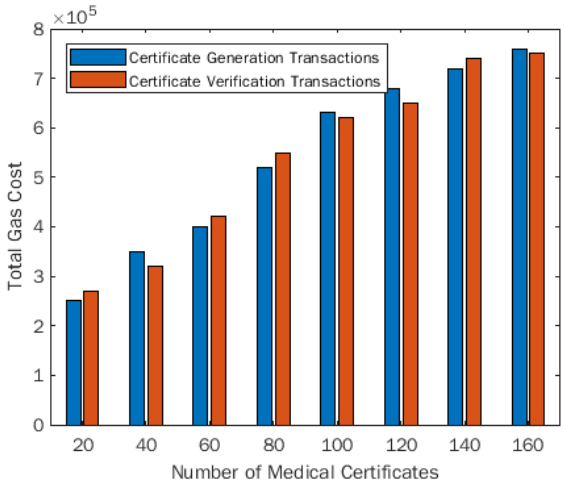


Fig. 2. Total gas cost of the proposed architecture for medical certificates

The processing time refers to the time required to initiate the transaction until the user gets the transaction execution ID. The processing time of the proposed architecture's transactions is analyzed for certificate generation and verification. The processing time is calculated between the interval of transactions requested and executed by the blockchain network for each set of transactions. It is analyzed that the certificate generation transactions require more processing time than the verification transactions due to the calculation involved to generate hash values for the certificates. Furthermore, certificate generation transactions involve blockchain nodes or miners to validate communications over the blockchain network by solving a back-end mathematical puzzle and placing these transactions into a new block. Then, the block is added to the blockchain network. Thus, the processing time also depends on the system configuration, the number of transactions, and miners' availability.

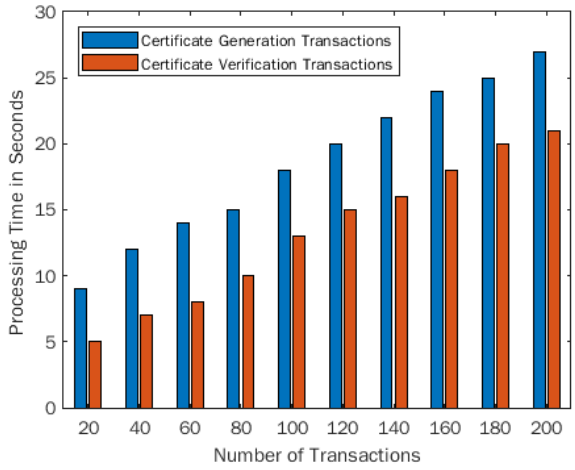


Fig. 3. Processing time of the proposed architecture

Table IV presents the smart contract cost analysis of the proposed architecture based on the transaction generation cost and execution cost in terms of Gas. The transaction cost denotes the cost of requesting the transaction in the blockchain network, and the execution cost refers to the total cost required to execute the corresponding smart contract transaction. The BlockSim toolkit is utilized to randomly calculate the transaction cost and execution cost for different smart contract operations.

TABLE IV
SMART CONTRACT COST ANALYSIS

Smart contract	Transaction cost (Gas)	Execution cost (Gas)
Healthcare center	2165431	1587652
User	863212	682101
Request	1154324	818553
Response	975438	832233
Share	854281	665321
Map	892256	723654
Generate	1252482	1356258
Store	652378	671289

Latency represents the delay between the transaction requests and transaction appearances on the blockchain network. In the proposed architecture, 200 transactions are considered to calculate the latency. All the transactions are defined in a separate set to evaluate the average time, which is shown in Fig. 3. The delay is analyzed by the BlockSim toolkit for each transaction set, and then, the average latency is calculated. As shown in Fig. 4, the proposed architecture's

2162-2248 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 09:16:29 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/MCE.2021.3140048, IEEE Consumer Electronics Magazine

TABLE V
COMPARISONS OF THE PROPOSED SCHEME WITH EXISTING SCHEMES

Reference	Blockchain type	Design/Implement	Blockchain tool	Cost analysis	Application
[14]	Not specified	Both	Not specified	No	Medical data stream analytics
[15]	Not specified	Both	Not specified	No	IoT-based cloud system for healthcare
[16]	Private	Both	Ethereum	No	Secure sharing of student credentials
[17]	Not specified	Designed	Not specified	No	Securing patient data using AI
[18]	Not specified	Designed	Not specified	No	COVID-19 forecast system
[19]	Private	Both	Ethereum	No	IoT-based blockchain healthcare system
[20]	Private	Both	Ethereum	No	Decentralized application
Proposed work	Public	Both	Ethereum	Yes	Medical certificates

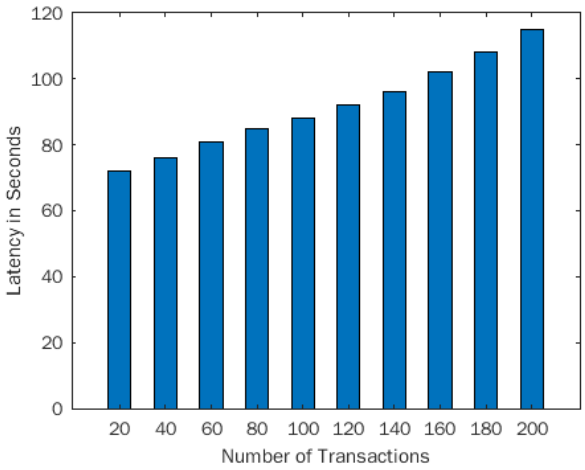


Fig. 4. Latency of the proposed architecture

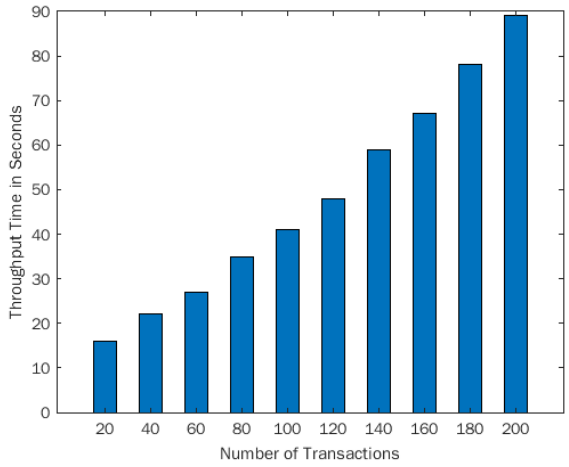


Fig. 5. Throughput of the proposed architecture

average latency increases by increasing the number of transactions.

The throughput, i.e. the total packets received per time frame, of the proposed application is represented in Fig. 5. Here, the blockchain network is analyzed for each set of transactions and calculated the total throughput of the proposed network. To calculate the throughput, a set of 20 transactions is utilized, and for each simulation, the number of transactions is increased in the set, and then, the average values are evaluated. The proposed architecture's throughput is increased with the increase in the number of transactions. Table V shows the comparison of the proposed work with the existing systems.

The computation time is the average time taken to execute a series of transactions. As shown in Fig. 6, the computation time increases with the number of transactions. The user initiates multiple transactions in the proposed application to analyze 20 transactions' average computation time intervals. As shown in Fig. 6, the average computation time is computed distinctly for each type of transaction. The computation is calculated separately for certificate generation and verification transactions. It is analyzed that the certificate generation transactions require more time than the verifications transactions.

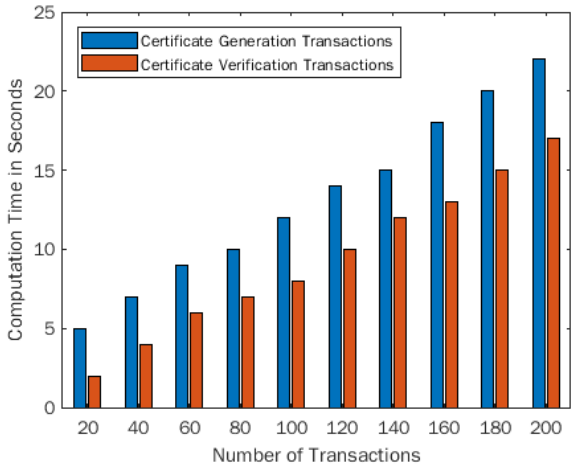


Fig. 6. Computation time of the proposed architecture

IV. CONCLUSIONS AND FUTURE SCOPE

This paper presents a blockchain-based architecture for the IoT-enabled healthcare system to provide privacy-preserving environment. The proposed architecture generates, and maintains medical certificates securely, maintains transparency, and enhances performance. The performance evaluation of the proposed architecture is measured based on various parameters, such as transaction latency, throughput, processing time, and computation time, for medical

2162-2248 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 09:16:29 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/MCE.2021.3140048, IEEE Consumer Electronics Magazine

certificates. The performance analysis of the proposed application shows that it is more efficient than the existing schemes. In the future, the proposed architecture can be improved by solving the scalability issue. Moreover, some features can also be added to update the generated certificates.

REFERENCES

[1] D. Kao, S. Hsiao, and R. Tso, "Analyzing wannacry ransomware considering the weapons and exploits," In *Proceedings of the 2019 21st International Conference on Advanced Communication Technology (ICACT),* PyeongChang Kwangwoon_Do, South Korea, 2019, pp. 1098-1107.

[2] H. Zhu, C. K. Wu, C. H. Koo, Y. T. Tsang, Y. Liu, H. R. Chi, and K. F. Tsang, " Smart healthcare in the era of Internet-of-Things," *IEEE Consumer Electronics Magazine,* vol. 8, no. 5, pp. 26-30, 2019.

[3] E. Daraghmi, Y. Daraghmi, and S. Yuan, "MedChain: A design of blockchain-based system for medical records access and permissions management," *IEEE Access*, vol. 7, pp. 164595-164613, 2019.

[4] S. Namasudra, "Fast and secure data accessing by using DNA computing for the cloud environment," *IEEE Transactions on Services Computing*, 2020. DOI: 10.1109/TSC.2020.3046471

[5] S. Raun, "Intelligent systems for smart health care: leveraging information for better well-being," *IEEE Consumer Electronics Magazine*, vol. 8, no. 2, pp. 71-71, 2019.

[6] S. Nakamoto. Bitcoin: A peer-to-peer electronic cash system, 2008. Available: http://www.bitcoin.org/bitcoin.pdf [Accessed on 15 September 2021]

[7] S. Namasudra, G. C. Deka, P. Johri, M. Hosseinpour, and A. H. Gandomi, "The revolution of blockchain: State-of-the-art and research challenges," *Archives of Computational Methods in Engineering*, vol. 28, no. 3, pp. 1497-1515, 2021.

[8] P. Sharma, R. Jindal, and M. D. Borah, "Blockchain technology for cloud storage: A systematic literature review," *ACM Computing Surveys*, vol. 53, no. 4, pp. 1-32, 2020. DOI: https://doi.org/10.1145/3403954

[9] P. Sharma, R. Jindal, and M. D. Borah, "A Review of Blockchain-Based Applications and Challenges," *Wireless Personal Communications*, 2021.

[10] S. Namasudra and G. C. Deka, "Applications of blockchain in healthcare", Springer, 2021. DOI: 10.1007/978-981-15-9547-9

[11] F. Jurado, O. Delgado, and Á. Ortigosa, "Tracking News Stories Using Blockchain to Guarantee their Traceability and Information Analysis," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 6, no. 3, 2020

[12] R. M. Alguliyev, R. M. Aliguliyev, and L. V. Sukhostat, "Efficient algorithm for big data clustering on single machine", *CAAI Transactions on Intelligence Technology*, vol. 5, no. 1, pp. 9-14, 2020.

[13] P. Sharma, M. D. Borah, and S. Namasudra, "Improving security of medical big data by using blockchain technology", *Computers and Electrical Engineering*, 2021. DOI: https://doi.org/10.1016/j.compeleceng.2021.107529

[14] P. M. Kumar, C. S. Hong, F. Afghah, G. Manogaran, K. Yu, Q. Hua, and J. Gao, "Clouds proportionate medical data stream analytics for internet of things-based healthcare systems", *IEEE Journal of Biomedical and Health Informatics*, 2021. DOI: 10.1109/JBHI.2021.3106387

[15] H. M. Ali, J. Liu, S. A. C. Bukhari, and H. T. Rauf, "Planning a secure and reliable IoT-enabled FOG-assisted computing infrastructure for healthcare", *Cluster Computing*, 2021. DOI: https://doi.org/10.1007/s10586-021-03389-y

[16] R. A. Mishra, A. Kalla, N. A. Singh, and M. Liyanage, "Implementation and analysis of blockchain based dapp for secure sharing of students' credentials," In *Proceedings of the IEEE 17th Annual Consumer Communications & Networking Conference (CCNC),* Las Vegas, NV, USA, 2020, pp. 1-2.

[17] H. S. Jennath, V. S. Anoop, and S. Asharaf, "Blockchain for healthcare: Securing patient data and enabling trusted artificial intelligence," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 6, no. 3, 2020.

[18] H. T. Raut, J. Gao, A. Almadhor, M. Arif, and M. T. Nafis, "Enhanced bat algorithm for COVID-19 short-term forecasting using optimized LSTM", *Soft Computing*, 2021. DOI: https://doi.org/10.1007/s00500-021-06075-8

[19] P. Sharma, N. R. Moparthi, S. Namasudra, S. Vimal, and C. H. Hsu, "Blockchain-based IoT architecture to secure healthcare system using identity-based encryption", *Expert Systems*, 2021. DOI: 10.1111/EXSY.12915

[20] R. V. Rosa, and C. E. Rothenberg, "Blockchain-based decentralized applications for multiple administrative domain networking," *IEEE Communications Standards Magazine*, vol. 2, no. 3, pp. 29-37, 2018.

[21] S. Wang, J. Wang, X. Wang, T. Qiu, Y. Yuan, L. Ouyang, Y. Guo, and F.Y. Wamg, "Blockchain-powered parallel healthcare systems based on the ACP approach", *IEEE Transactions on Computational Social Systems*, vol. 5, no. 4, pp. 942-950, 2018.

[22] A. Khatoon, "A blockchain-based smart contract system for healthcare management," *Advances in Blockchain and Distributed Ledger Technology (DLT) for Industry 4.0 Technologies,* vol. 9, no. 1, 2020.

[23] I. Daisuke, M. Kashiyama, and T. Ueno, "Tamper-resistant mobile health using blockchain technology," *JMIR Mhealth Uhealth*, vol. 5, no. 7, 2017.

[24] A. M. Salem, and A. Y. Nikitaeva, "Knowledge engineering paradigms for smart education and learning systems," In *Proceedings of the 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics* (MIPRO), Opatija, Croatia, 2019, pp. 1571-1574.

[25] P. Bhattacharya, S. Tanwar, U. Bodke, S. Tyagi and N. Kumar, "BinDaaS: Blockchain-based deep-learning as-a-service in healthcare 4.0 applications," *IEEE Transactions on Network Science and Engineering*, vol. 8, no. 2, pp. 1242-1255, 2021.

[26] S. Ndichu, S. Kim, and S. Ozawa, "Deobfuscation, unpacking, and decoding of obfuscated malicious JavaScript for machine learning models detection performance improvement", *CAAI Transactions on Intelligence Technology*, vol. 5, no. 3, pp. 184-192, 2020.

[27] P. Sharma, R. Jindal, and M. D. Borah, "Healthify: A blockchain-based distributed application for health care," In *Applications of Blockchain in Healthcare*, S. Namasudra, G. C. Deka (eds), Springer, vol. 83, 2021. https://doi.org/10.1007/978-981-15-9547-9_7

[28] Doreswamy, M. K. Hooshmand, and I. Gad, "Feature selection approach using ensemble learning for network anomaly detection", *CAAI Transactions on Intelligent Technology*, vol. 5, no. 4, pp. 283-293, 2020

[29] M. I. García Sáez, "Blockchain-enabled platforms: Challenges and recommendations," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 6, no. 3, 2020.

[30] M. Alharby, and A. V. Moorsel, "BlockSim: An extensible simulation tool for blockchain systems," *Frontiers in Blockchain*, vol. 3, pp. 1-16, 2020.

2162-2248 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 09:16:29 UTC from IEEE Xplore. Restrictions apply.

Neural Computing and Applications
https://doi.org/10.1007/s00521-021-06743-8

S.I.: LSNC & OUAI

CanarDeep: a hybrid deep neural model with mixed fusion for rumour detection in social data streams

Deepak Kumar Jain[1] · Akshi Kumar[2] · Akshat Shrivastava[3]

Received: 15 May 2021 / Accepted: 11 November 2021
© The Author(s), under exclusive licence to Springer-Verlag London Ltd., part of Springer Nature 2021

Abstract
The unrelenting trend of doctored narratives, content spamming, fake news and rumour dissemination on social media can lead to grave consequences that range from online intimidating and trolling to lynching and riots in real- life. It has therefore become vital to use computational techniques that can detect rumours, do fact-checking and inhibit its amplification. In this paper, we put forward a model for rumour detection in streaming data on social platforms. The proposed *CanarDeep* model is a hybrid deep neural model that combines the predictions of a hierarchical attention network (HAN) and a multi-layer perceptron (MLP) learned using context-based (text + meta-features) and user-based features, respectively. The concatenated context feature vector is generated using feature-level fusion strategy to train HAN. Eventually, a decision-level late fusion strategy using logical OR combines the individual classifier prediction and outputs the final label as rumour or non-rumour. The results demonstrate improved performance to the existing state-of-the-art approach on the benchmark PHEME dataset with a 4.45% gain in *F1*-score. The model can facilitate well-time intervention and curtail the risk of widespread rumours in streaming social media by raising an alert to the moderators.

Keywords Rumour · Social data streams · Deep learning · Data fusion

1 Introduction

The blatant social media technology has an obvious, potentially enormous dark side. The *tsunami* of polluted information threatens the underpinning trust between users and their experience online. The amplification of the most abominable views in social data streams, the virality of incorrect information at a lightning speed and the anonymization of our public discourse divulge the vulnerabilities associated. Moreover, the information disorders (misinformation, disinformation and mal-information) [1] vary in impact and mechanization which is not limited to a single social media platform. The primary function of the misleading and fabricated content is to make sense in a wrapped version of reality with intentions ranging from hurtful or harmless to the harmful one. The economics of social media favors these 'information statements' that are instrumentally relevant and often consumed within our online echo chambers and filter bubbles. Undoubtedly, doctored narratives, fake news, rumours, polarized content, deceptive propaganda, content spamming and fabricated hoaxes define the taxonomy of online content within the periphery of social media abuse [2–4].

A rumour is any statement that is not yet confirmed at the time of posting, irrespective of whether it's true or false [5]. It gets viral at great speed and is facilely believed, particularly during public crisis. As soon as users start engaging with the post by "liking", "replying", "commenting", "forwarding" and "sharing", technology proliferates diffusion at an unparalleled rate. The COVID-19 public health crisis made this vulnerability highly visible, with numerous rumours circulating through social media. To examine post legitimacy, many social media

✉ Akshi Kumar
akshi.kumar@nsut.ac.in

[1] Key Laboratory of Intelligent Air-Ground Cooperative Control for Universities in Chongqing, College of Automation, Chongqing University of Posts and Telecommunications, Chongqing, China

[2] Department of Information Technology, Netaji Subhas University of Technology, Delhi, India

[3] Department of Computer Science & Engineering, Delhi Technological University, Delhi, India

Published online: 08 January 2022

platforms have put in systematic efforts to moderate posts and improve online accountability. These include implementing the mandatory code of practice for content filtering using a combination of artificial intelligence and user feedback. But the guidelines are not transparent to the users, and the content moderators have to deal with information overload and exposure to harmful and undesirable content. Indeed the information credibility analysis on social media platforms is complex as it diversifies along three dimensions as post-level credibility, user-level credibility and topic-level credibility [6, 7]. Moreover, the problem with debunking rumours is that the exchange and distribution are at such an unprecedented rate that makes it difficult to scale up the combat operations to curb the viral spread. At the same time, the diffusion process of rumours often has an associated chronology. Often spread as breaking news with high virality and cascading effect, the rumours subside with time and eventually die out. It is reasonably practical to capture the changing characteristics of rumour spread as with time veracity and stance are comprehensible. But the damage done within the early hours of release is huge as it evokes high-arousal emotions. Thus, early detection of rumours (identifying if an online post is a rumour or non-rumour) and combating its viral spread is a pressing priority. A rumour detection system that identifies an online posting as a candidate rumour in its early stages can be effectively used for well-timed intervention by alerting the moderators to take corrective measures that impede the spreading of inappropriate content.

Pertinent studies report an enormous use of machine learning techniques to combat online content fabrication but have limited success due to lack of context awareness. Recent research trends report the use of deep learning models such as convolutional neural networks (CNNs) [8] and recurrent neural networks (RNNs) [9] for various natural language understanding tasks on streaming data of online media [10]. A superior hierarchical attention-based document classification model, called hierarchical attention network (HAN) has been put forward by Yang et al. Commonly built from bidirectional RNNs [11], a HAN consists of gated recurrent units (GRUs) or Long-short-term-memory model (LSTMs) with attention mechanisms. It is conceptual based on "hierarchies" where the output from lower levels in hierarchical structure becomes the input to upper level. The HAN architecture characterizes the hierarchically derived knowledge in a document with sentences and words as building blocks at the corresponding level of hierarchy. Falsehood on Twitter spreads like a wildfire and may be compiled as a single tweet, a tweet thread (multi-part tweet) or a conversational thread. Using HAN lets discrete contribution of compound fragments of a tweet (tweet-sentence-word) to its quintessence by implementing two attention mechanisms. The contextual connotation of these fragments is taken into account to construct the document representation. This research aims to cope with the language complexities by using HAN with deep contextualized language representations from pre-trained ELMo (Embeddings from Language Models) word embedding [12].

As an effectual technique to tweet-level analysis of rumours, an attention-based hybrid deep learning model for rumour detection from social data streams is put forward. The proposed *CanarDeep* model detects rumourous posts in real-time by using the learned features from both HAN and multi-layer perceptron (MLP). The model derives its nomenclature from the French word *"**Canard**"* which means *'unfounded, groundless or false report or story'* and ***"Deep"*** from the *'deep neural network'* built to detect rumours and combat its viral spread in online streams. A primary approach to decipher the truth value of a post is to look for some user-based and text-based evidence. Some meta-features such as re-post count and morpho-syntactic (exclamations) & typographic (capitalization, quotes) markers can also serve as non-trivial cues. Therefore, in the mixed-fusion hybrid deep neural model, rumour detection is done by the individual classification model, namely HAN and MLP using context-based features (textual + meta-features) and user-based features, respectively. The context-based features are initially fused to form a concatenated feature vector. At both word and sentence levels, HAN uses a bi-directional GRU with attention [13]. The attention mechanism enables to find the most important words and sentences in a document while taking the context into consideration. On the other hand, the MLP generates its output based on the user-based features using the back-propagation algorithm and adjusting the weights of the neurons accordingly. MLP is a simple yet highly sought-after model when it comes to binary classification using discrete numerical features [14]. Finally, a late fusion method (logical OR operation) is used to capture the results of the two classifier sub-networks and classify the output as a decision. The primary contributions of the work are:

- The hybrid of HAN and MLP is used to classify online streaming posts as rumours & non-rumours.
- HAN is trained using context-based features (the textual content & post-meta-data) and MLP is trained using the user-based features (discrete numeric features related to the user profile).
- Two types of data fusion techniques are used, that is, feature-level (early) fusion is done to concatenate context-based features whereas the output generated from individual classifiers is combined to produce a

final common decision using a logical OR for decision-level (late) fusion.
- Comparison with the state-of-the-art model [15] to validate the improvements and consistency in improvements across the events in the benchmark Twitter PHEME dataset [16].

The organization of the paper is as follows: related work within the area of rumour detection in social media is given in Sect. 2. Section 3 gives description of the mixed-fusion hybrid of HAN-MLP for rumour detection followed by the results in Sect. 4 and conclusion in Sect. 5.

2 Related work

Various primary [5, 15] and secondary [4, 17] literature studies have been reported on rumour detection in micro-blogs. The variety of techniques used to classify tweets as rumour are dominated by the use of machine learning and more recently deep learning. Diverse combinations of features which include content-based, post-based, user profile-based and rumour diffusion attributes have been used to train the learning models. A range of studies focuses on the use of text-only-based features from the tweet. Takahashi and Igata [18] used keywords and retweet ratio for rumour detection in Twitter data. Kumar et al. [19] compared machine learning and deep learning models trained using Tf-idf features and one-hot encoding, respectively, on the PHEME and reported Bi-LSTM as the best model. Bhattacharjee et al. [20] used the Glove word embedding to train a LSTM for rumour detection on the PHEME dataset. The use of RNN has also been reported with [21] and without [22] attention mechanism. Several studies have reported the combination of features. A hybrid of CNN and RNN was proposed by Nguyen et al. in 2017 [23] using text-based and temporal features. Guo et al. [24] suggested a word, post and sub-event level hierarchical network using attention-based Bi-LSTM encoder combined with social features. The results were evaluated on Sina Weibo and Twitter datasets. In 2020, Zubiaga et al. [16] used a conditional random field [CRF] model to detect rumours in the PHEME dataset. The authors used both content and context-based features to train the model. Vijeev et al. [25] used both user features and content features to train three supervised learning methods on PHEME. Our recently published research [26] describes an attention-based residual network (ARN) to detect rumour in the same Twitter benchmark dataset and reports an improvement in recall and *F1* score in comparison to Zubaiga's CRF model.

3 The proposed *CanarDeep* model

A comprehensive rumour classification system [3] consists of four core components, namely rumour detection, rumour tracking, stance classification and veracity classification as shown in Fig. 1. Automatic detection of rumours is typically a natural language processing (NLP) task to label online streaming posts as a rumour or a non-rumour. It intends to recognize potential rumours, especially the emerging ones.

The proposed *CanarDeep* model corresponds to the first component, i.e. the recognition of potential rumours. The publicly available PHEME dataset [15] for rumour detection which is compiled using event-based tweets for classifying tweets into rumourous or non-rumourous categories is used. It is a collection of 5802 tweets, out of which 1972 are labelled as rumours and the remaining 3830 as non-rumours. The data are initially pre-processed to transform it into a well-formed data for analytics. Preprocessing was done by converting the tokens to lowercase, removing the hyperlinks, unwanted characters, symbols, whitespace and stop-words from the text, performing spell check, lemmatization and stemming. *StandardScaler* from the *sklearn* library for Python to scale all the user profile-based features. Evaluation of a rumour strongly relies on its context which is formulated using the textual content and the post-meta-data. Characteristics of user profile also help in building trust in the content posted and therefore to train our model we use the context-based, the user-based features and the rumour/non-rumour label to train our model. A total of 11 context-based features and five user-based features are used as shown in Fig. 2.

The description of these features is as follows:

- ***Context-Based Features:*** These entail textual content of the post as well as the content-based cues called meta-features associated with the post.
  - *Word Vectors:* Word vectors map words or phrases from vocabulary to a corresponding vector of real numbers which are used to find word predictions, word similarities/semantics [27]. In *CanarDeep*, the word vector is built using the ELMo 5.5B word embedding.
  - *POS Tags:* These tags do the part-of-speech annotation. They tag each word in the tweet with the respective grammar tag such as nouns, adverbs, adjectives, etc.
  - *Capital Ratio:* Most of the time, the word which has been spelled in capital letters tends to have more impact than the word written in lower case.
  - *Word Count:* It depicts the *c*ount of words in a tweet.

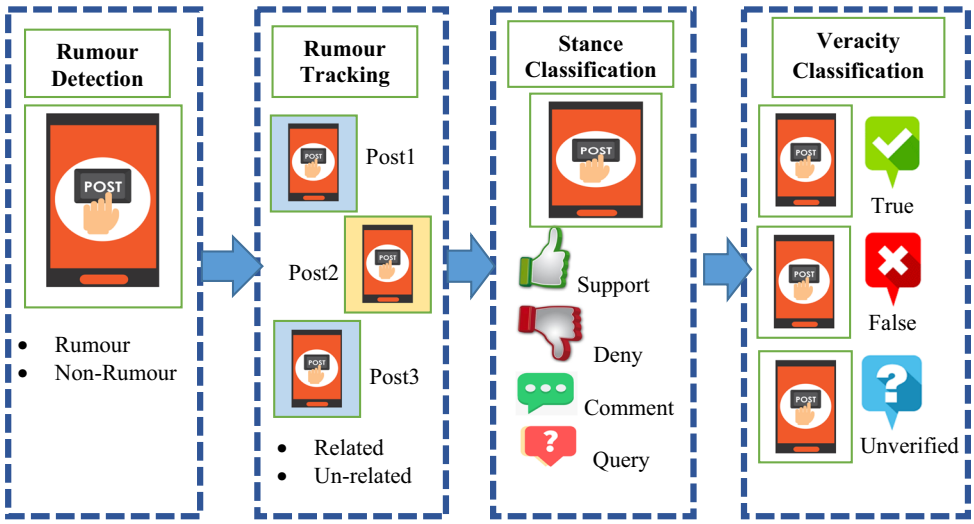


Fig. 1 Rumour classification system

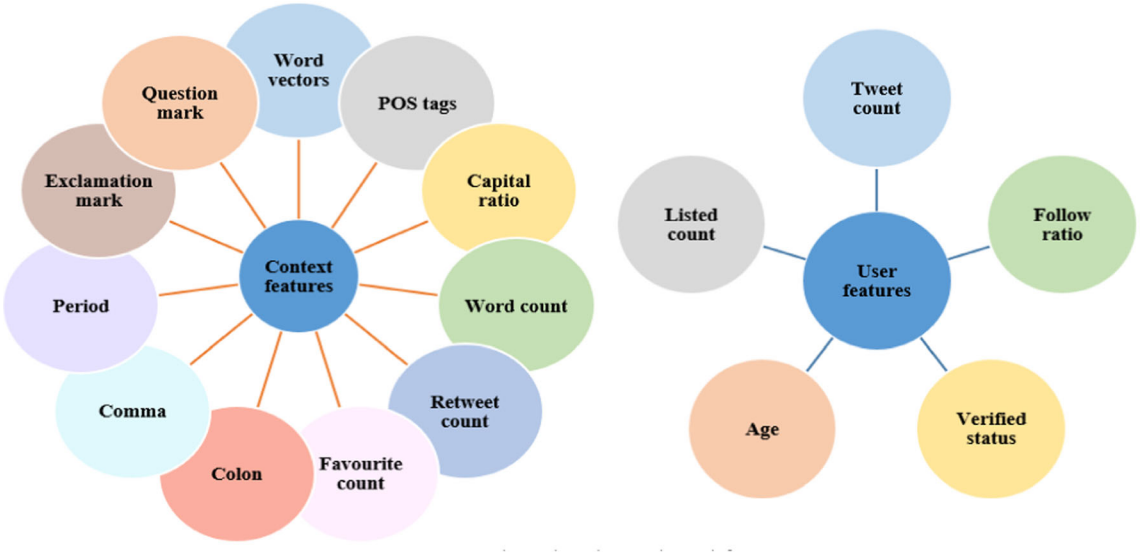


Fig. 2 Context-based and user-based features

- *Question Mark:* Question marks sometimes represent the uncertainty of a saying, disrespect, impatience or lack of tactfulness. Its presence or absence in a tweet is represented as a binary feature.
- *Exclamation Mark:* The exclamation marks in the tweets express surprise, astonishment or any strong emotion resulting in additional emphasis.
- *Period:* Punctuation might represent good writing and hence quality reporting. Its presence or absence in a tweet is represented as a binary feature.
- *Colon:* The use of a colon in tweets helps the user to add two independent clauses, thus allowing them to add two complete thoughts that stand alone as complete sentences. The presence of a colon may suggest careful reporting. Its presence or absence in a tweet is represented as a binary feature.
- *Comma:* The use of a comma in a tweet suggests quality reporting. Its presence or absence in a tweet is represented as a binary feature.
- *Favorite Count:* This depicts the number of users who have signaled a particular tweet as their favorite. The higher the count, the more are the chances of it not being a rumour because a higher favorite count shows that people believe in that tweet.
- *Retweet Count:* This depicts the number of users who have retweeted a particular tweet. Retweeting is defined as the sharing of a tweet by a user so that the user's followers can also read the tweet. If the retweet count of a tweet is high, there is a good chance that the tweet is not a rumour because the users trusted it enough to share it.

- ***User-based Features:*** These comprise of user profile-based discrete numeric features as follows:
  - *Tweet Count:* The total number of tweets a user had posted on Twitter.
  - *Listed Count:* The count of lists a user is a part of, i.e. the number of times they were added to a list by other users.
  - *Follow Ratio:* The reputation of a user is assessed on the basis of the count of their followers. But, sometimes, the count of followers does not reflect the true prominence of a user. For example, some users follow many others in order to be followed back. Keeping this scenario in mind, we take the follow ratio as a feature, which is the number of followers someone has divided by the number of people following them. It is basically the follower to following ratio.
  - *Age:* The age of a Twitter user pertains to the years they have been using Twitter. It is the time from the setting up of the account to current tweet time.
  - *Verified:* The account verification status of the user. A verified user account has traceability and accountability making such accounts least probable to spread rumours in comparison to an unverified account.

There are two primary fusion strategies for multiple input types, namely the model-free fusion and model-level fusion (medial). Early fusion (feature-level) and late fusion (decision-level) are types of model-free fusion. Early fusion combines the input types to form a single input vector to train a classifier. It basically refers to the fusing (concatenation) of features from multiple data sources to create a new feature set. The new feature vector has higher discriminative power in comparison to the individual input feature vectors. Whereas, in late fusion, individual classifiers are modelled using discrete input types and the predictions of the individual classifiers are combined to decide the final output [28, 29]. Model-level fusion pools the benefits of both of these strategies by concatenating high-level feature representations from diverse classifiers. Figure 3 depicts these types of data fusion.

In this research, two types of data fusion strategies are used. The early fusion strategy is used to combine the textual content with the meta-features of the post to generate a context-based feature vector. That is, the meta-data features which are categorical in nature are concatenated to the word vector generated by the embedding layer as a feature-level fusion strategy. Late fusion is applied to combine the decisions of multiple classifiers, namely HAN and MLP, trained using context-based and user-based features, respectively, to produce a final prediction. There are various abstract methods to accomplish decision-level fusion, such as using logical operators, votes or weighted majority. The proposed hybrid of HAN and MLP is trained using the user-based and context-based features with a mix of fusion strategies. Figure 4 represents the architecture of the proposed *CanarDeep* model.

3.1 Hierarchical Attention Network (HAN)

To learn the context-based features, the architecture of the HAN classifier has an embedding layer, encoders and attention layers. In this research, the context-based ELMo 5.5B model [12] is used to generate the word embedding to seed the classifier. The bi-directional GRU with attention layer is then used to capture the essential meaning of the document. The model includes two levels of attention—word-level and sentence-level. The differential contribution of various parts of the document to its essential meaning and changing context is considered when constructing the representation of the document. The encoders extract relevant context and the attention layers compute the degree of relevance of the sequence of tokens with respect to the document. The bi-directional GRU with attention layer is firstly used at word-level and repeated at the sentence level as follows:

- Word-level encoder and attention

  An embedding matrix maps the discrete categorical variables into a vector of continuous numbers. For a given sentence with words w_{it}, $t \in [0, T]$, a word to the vector embedding matrix, W_e, $x_{it} = W_e w_{it}$ is built using ELMo. The vectorized tokens are the inputs for the next layer, the word encoding layer. In this research, a bidirectional GRU is applied to get annotations of the words by summarizing information. The forward GRU $\overrightarrow{f}$ reads the sentence s_i from w_{i1} to w_{iT} as given in (1):

$$\vec{h}_{it} = \overrightarrow{GRU}(x_{it}), t \in [1, T] \tag{1}$$

  The backward GRU $\overleftarrow{f}$ reads sentences from w_{iT} to W_{i1}, as given in (2):

$$\vec{h}_{it} = \overrightarrow{GRU}(x_{it}), t \in [T, 1] \tag{2}$$

  The annotation of the word w_{it} is calculated by combining the forward and backward hidden states i.e. $h_{\mathrm{it}} = [\overrightarrow{h}_{\mathrm{it}}, \overleftarrow{h}_{\mathrm{it}}]$.

  The word-level attention layer extracts the words with major contributions to the essential meaning. These words form the sentence vector. That is, using a one-layer multi-layer perceptron (MLP) we calculate u_{it} to get the hidden representation of the h_{it} as given in (3):

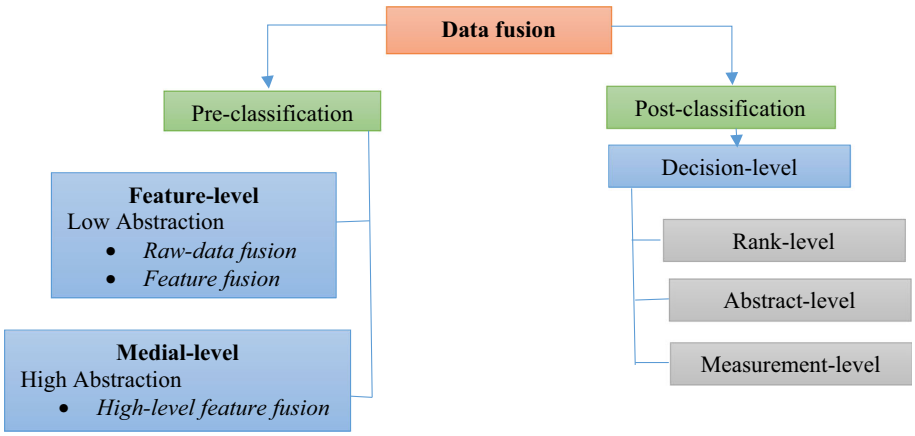


Fig. 3 Data fusion strategies

$$u_{it} = \tanh(W_w h_{it} + b_w) \tag{3}$$

The layer uses the tan*h* function to ensure that the network does not falter. The tan*h* function achieves this by correcting input values to be between − 1 and 1 and also maps zeros to near-zero. Next, the normalized importance, a_{it} is generated using a sigmoid or a softmax function that calculates the impact of the word as a similarity of u_{it} with a word-level context vector $u_w$as given in (4):

$$a_{it} = \frac{\text{expexp}\left(u_{it}^T u_w\right)}{\Sigma_t \text{expexp}\left(u_{it}^T u_w\right)} \tag{4}$$

After that, the weighted total of the word annotations is calculated using (5) thus generating the sentence vector s_i:

$$s_i = \sum_t a_{it} h_{it} \tag{5}$$

- Sentence-level encoder and attention

  Similar to the word level, now the entire network is re-run on sentence that is focusing on the sentence i. There is no embedding layer as we already get sentence vectors s_i from word level as input. The bidirectional GRU sentence encoder is given in (6) & (7):

$$\vec{h}_{it} = \overrightarrow{GRU}(s_i), i \in [1, L] \tag{6}$$

$$\vec{h}_{it} = \overrightarrow{GRU}(s_i), t \in [L, 1] \tag{7}$$

The annotation of sentence i is computed by concatenating the forward and backward hidden states i.e. $h_i = [\overrightarrow{h}_i, \overleftarrow{h}_i]$.

Similar to the word-level attention layer, the sentence-level attention layer extracts sentences that significantly convey the meaning of the document. Therefore, a one-layer MLP is used to calculate u_i and get the hidden representation of h_i as given in (8):

$$u_i = \tanh(W_s h_i + b_s) \tag{8}$$

Likewise, the normalized importance, a_i is also generated using a sigmoid or a softmax function. The similitude of u_i with a sentence-level context vector $u_{s\,i}$ is calculated to characterize the significance of the sentence as a as given in (9):

$$a_i = \frac{\text{expexp}\left(u_i^T u_s\right)}{\Sigma_i \text{expexp}\ \left(u_i^T u_s\right)} \tag{9}$$

The weighted total of the sentence annotations is then computed to generate the document vector. The document vector v_i summarizes all the data of the sentences present in the document as given in (10):

$$v_i = \sum_i a_i h_i \tag{10}$$

Trainable weights and biases are again randomly initialized and jointly learned during the training process as shown in (11) and (12).

$$u_i = \tanh(W_s h_i + b_s) \tag{11}$$

$$a_i = \frac{\exp\left(u_i^T u_s\right)}{\Sigma_i \exp\left(u_i^T u_s\right)} \tag{12}$$

The final output is a document vector v_i which can be used as features for document classification as depicted by (13)

$$v_i = \sum_i a_i h_i \tag{13}$$

Finally, the document vector v is used to generate the final class. The final layer which uses a sigmoid activation takes as input v to output the classification result of the HAN sub-network as given in (14)

$$p = \text{ sigmoid } (W_c v + b_c) \tag{14}$$

The training loss is taken as the negative of the log-likelihood of the correct labels as shown in (15)

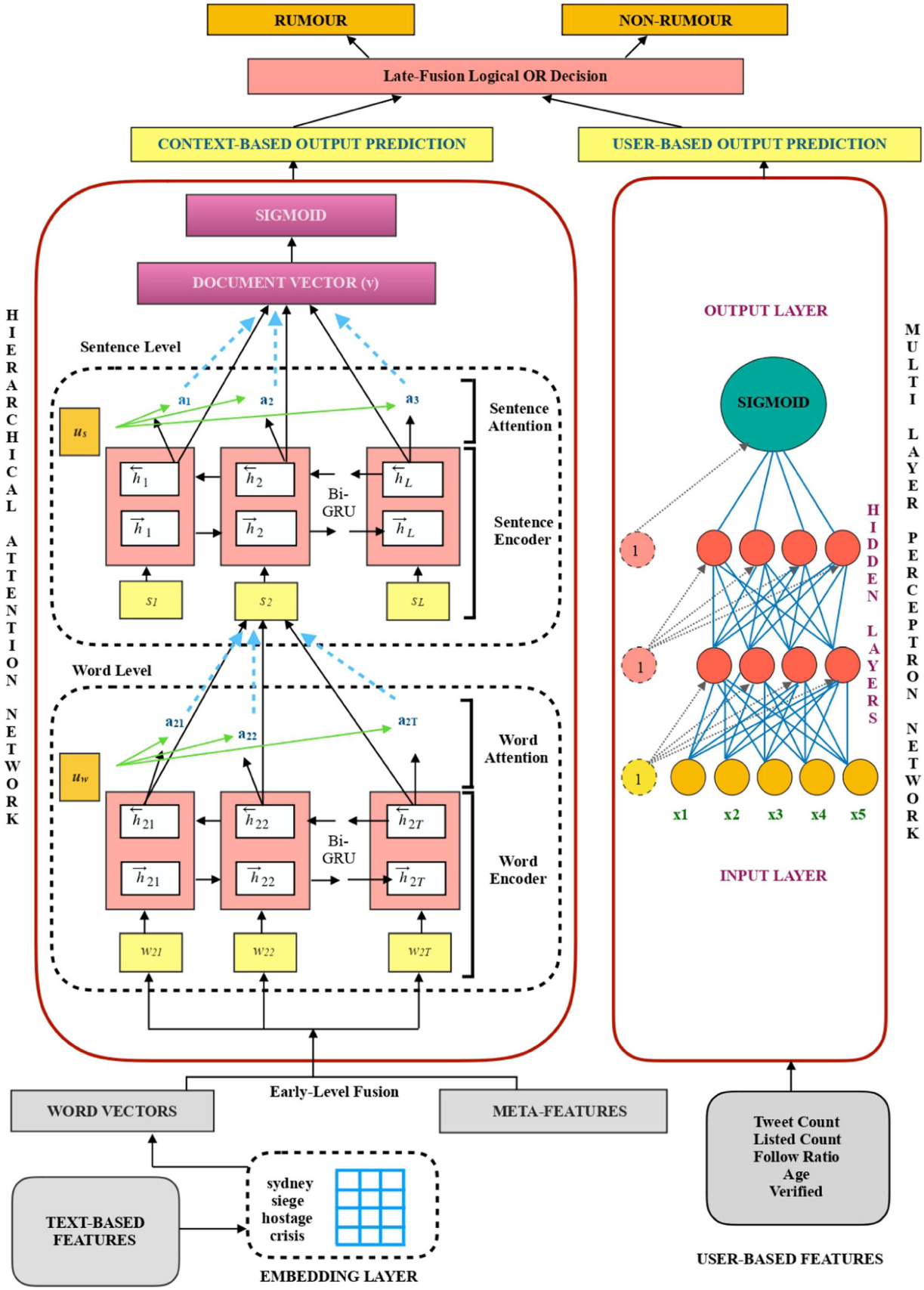


Fig. 4 The proposed *CanarDeep* model

$$L = -\sum_{d} \log pd_j \tag{15}$$

where j stands for the label of document d.

3.2 Multi-Layer Perceptron (MLP)

The MLP forms the second classifier of our hybrid model, and it takes the user-based features as input. It does so by using the back-propagation algorithm and adjusting the weights of the neurons after every backward pass, thus minimizing the error as much as possible and finally attaining convergence [30]. The MLP architecture used in this research is as follows: the input layer consists of five neurons, the two hidden layers consist of four neurons each, and the output layer has a single neuron with a sigmoid activation function which generates the final output class, i.e. rumour (positive class) or non-rumour (negative class).

Next, we fuse the outputs from the two classifiers, i.e. HAN and MLP using the logical OR operation to generate the final class for a particular post.

3.3 Decision-level fusion for classification

In the *CanarDeep* model, the output decisions from both the classifiers, i.e. HAN and MLP, are fused using the logical OR operation. If the decision from either classifier is that the given input is a rumour, then the input is classified as a rumour. The given input is classified as a non-rumour if and only if both the classifiers decide that the input is a non-rumour. Table 1 illustrates the logical OR operation.

The OR operation helps to debunk a rumour with maximum possibility. If both the classifiers detect the post as rumour then it is irrefutably a rumour. Textual content and its meta-features provide valuable markers to indicate a rumour, and therefore, even if only the context-based classifier is indicative of rumour, the output is marked as rumour. This is because rumours are driven based on psychology and behaviour of users which may alter with change in beliefs, confusion and anxiety or due to uncertainty. Hence, even a non-suspicious account can spread rumours. Likewise, if a user profile is identified suspicious using the user-based classifier, it is also marked as rumour. The primary notion is that intelligent bots and masqueraded profiles tend to use professional services for believable content writing tactics which can often be missed by the context-based classifier. Thus, the ORing decides to discard a post only if both the classifiers classify it as a non-rumour.

Table 1 Decision-level fusion using logical OR

HAN	MLP	Decision
+ (Rumour)	+ (Rumour)	Rumour
− (Non- Rumour)	− (Non- Rumour)	Non-Rumour
+ (Rumour)	− (Non-Rumour)	Rumour
− (Non-Rumour)	+ (Rumour)	Rumour

Table 2 Parameters used for HAN and MLP classifier

Parameter	Value
HAN	
Embedding Dimension	300
Bi-GRU Units	150
Hidden Units	300
Return Sequences	True
Trainable	True
Non-Linearity Function	ReLu
Loss Function	Binary Crossentropy
Optimizer	Adam
Dropout	0.5
Word Embedding	ELMo 5.5B
Batch Size	128
Epochs	7
Maximum Vocabulary Size	20,000
Maximum Sentence Length	50
Maximum Sentence Number	5
MLP	
Max Iterations	10
Solver	Adam
Learning Rate Initializer	0.01
Batch Size	200
No. of Hidden Layers	2
Units in each Hidden Layer	4
Tolerance	1e-4
Activation Function	ReLu

4 Results and discussion

Two events in the PHEME dataset suffer from the class imbalance problem. These are Charlie Hebdo (CH) and Ferguson (FS), where the data are skewed towards the non-rumour category in comparison to the rumour category, while the other three i.e. Germanwings (GW), Ottawa Shooting (OS) and Sydney Siege (SS) do not. To resolve this, the performance of the *CanarDeep* model is examined with respect to each of the individual events to analyse how

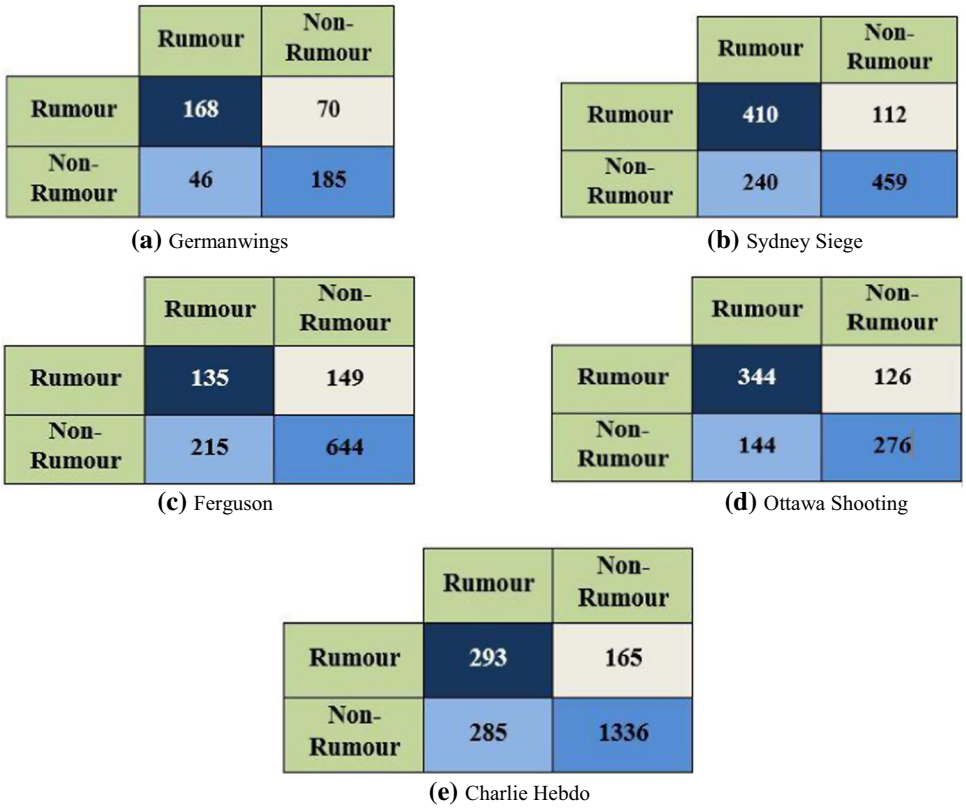


Fig. 5 Confusion matrices for individual events in PHEME

well the model performs across the dataset. Data skewness or class imbalance sabotages a classification task and using accuracy as a performance metric may lead to incorrect interpretation & evaluation [31, 32]. A relative-to-each-class measure (like ROC-AUC) is thus preferred over the absolute measure (accuracy). *F1*-Score, Precision and Recall have been used to evaluate the classifier performance.

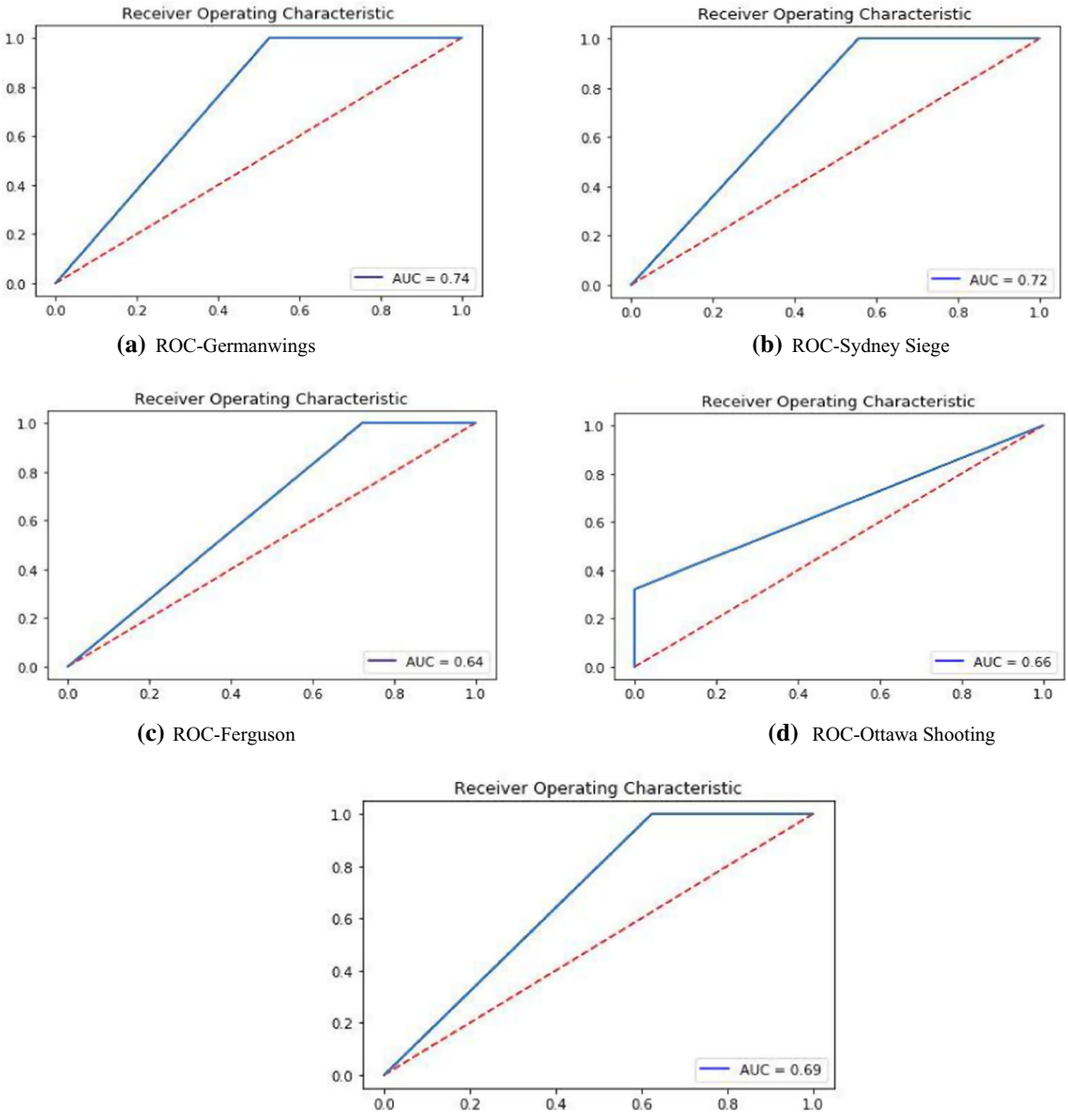
There are various parameters that have been used for both the sub-networks of our proposed model during the experiment. The values of these parameters are shown in Table 2.

A good understanding of how the proposed model performed on individual events can be obtained by taking a look at each event's confusion matrix. The confusion matrices [29, 30] are calculated by counting four values for each event, namely, True Positives (TP), False Positives (FP), False Negatives (FN) and True Negatives (TN). The confusion matrices for each PHEME dataset event are shown in Fig. 5a through 5e, wherewith horizontal axis represents the actual class and the vertical axis represents the predicted class.

The AUC scores of the five events present in the PHEME dataset range from 0.64 to 0.74. Figure 6a through e displays the ROC curve for individual events. The proposed *CanarDeep* classifier performs better for two events, namely, the Germanwings and Sydney Siege in comparison to the other three.

Our model performs fairly uniformly for all the five individual events in the dataset, as shown in Table 3. Precision values generated by the *CanarDeep* model outperform the CRF classifier [15] for three events. Recall values achieved by the *CanarDeep* model are also higher across four out of the five events with CH being an exception. The proposed model manages to achieve a precision-recall equilibrium, transcending the state-of-the-art.

The SS event has the highest number of tweets among the events and does not suffer from the class imbalance problem too. The highest improvement in the *F1* score is observed for this event with an improvement of 40.43%, as noted in Table 3. A significant improvement in the recall score is observed for four events accounting for *CanarDeep's* overall superior performance over the CRF classifier. Moreover, the proposed model does not let the event datasets suffering from class imbalance hamper its performance, another qualitative improvement over the CRF classifier. As compared to CRF which has *F1* score of 0.636 for CH and 0.465 for FS, *CanarDeep* achieves *F1* of 0.715 for CH and 0.606 for FS, respectively. The proposed model follows a similar trend with datasets that are free from the class imbalance problem. Superior *F1* scores are obtained for the GW, SS and OS events in comparison to



(a) ROC-Germanwings **(b)** ROC-Sydney Siege

(c) ROC-Ferguson **(d)** ROC-Ottawa Shooting

(e) ROC-Charlie Hebdo

Fig. 6 ROC curve for individual events in PHEME

the CRF classifier. Our previous work using attention-based residual network (ARN) [26] which reported an improvement over the state-of-the-art CRF in terms of recall and *F1* scores were also compared to *CanarDeep*. The effectiveness of the *CanarDeep* model is assessed using the whole dataset, i.e. all the five events combined. Table 4 depicts the performance results on the complete data.

Thus, training the hybrid HAN and MLP using user-based and context-based features with a mix of fusion strategies outperforms the existing state-of-the-art model for all the three metrics.

5 Conclusion

Rumours thrive in adversarial situations due to their high manipulative power and the inability of naïve users to differentiate between legit and false content. It is thus imperative to determine the potential rumour candidate at its emergence. As a key to classify online rumours, a deep learning model was proposed which coalesced information from two classifiers to label tweets as rumours in benchmark PHEME dataset. Two different input types, namely context-based and user-based were learned separately using the classifiers (HAN for context-based and MLP for user-based). The output predictions of these were then combined

Table 3 Classifier performance for individual events

Event	Metric	CRF [15]	*CanarDeep*
GW	P	0.743	0.753
	R	0.668	0.755
	F1	0.704	0.754
CH	P	0.545	0.732
	R	0.762	0.698
	F1	0.636	0.715
OS	P	0.841	0.695
	R	0.585	0.696
	F1	0.690	0.695
SS	P	0.764	0.721
	R	0.385	0.717
	F1	0.512	0.719
FS	P	0.566	0.613
	R	0.394	0.599
	F1	0.465	0.606

GW Germanwings, *CH* Charlie Hebdo, *OS* Ottawa Shooting, *SS*,Sydney Siege, *FS* Ferguson, *P* Precision, *R* Recall, *F1* F1 score

Table 4 Performance on complete dataset

Model	Metrics		
	P	*R*	*F1*
CRF [15]	0.667	0.556	0.607
ARN [26]	0.662	0.570	0.612
CanarDeep	0.685	0.592	0.634

using a logical OR decision-level operation to finally classify the tweets. The benefit of using early-level fusion to concatenate textual and meta-features as context-based features are that it does not isolate interactions between correlated features whereas the advantage of using decision-level fusion for final output is that the model needs not synchronize between different types of features. The robustness of the technique is validated for both individual events and the whole dataset. The experimental evaluation reveals superior performance in comparison with the existing state-of-the-art with a 4.45% gain in *F1*-score.

As rumours can harm and do irreparable damage, it is equally important to achieve high performance with interpretability and verifiability of decisions. We intend to use explainable artificial intelligence (XAI) to realize the action traceability and build robust, trustworthy and unbiased learning models to detect rumours in social data streams. Further future work also includes using characteristics of breaking news and long-standing rumours for training learning models. Also, the current research only uses the rumour and non-rumour categories to label the post but future studies need to develop models that identify fine-grain rumour categories such as dread rumours, wish rumours, wedge-driving rumours and reputation rumours.

Authors' contributions All the authors have equally contributed in the manuscript preparation.

Funding No Funding has been received.

Availability of data and materials Publicly accessible data have been used by the authors.

Code availability Can be made available on request.

Declarations

Conflict of interest The authors certify that there is no conflict of interest in the subject matter discussed in this manuscript.

Ethics approval The work conducted is not plagiarized. No one has been harmed in this work.

Consent to participate All the authors have given consent to submit the manuscript.

Consent for publication Authors provide their consent for the publication.

References

1. Bounegru L, Gray J, Venturini T, Mauri M (2018) A Field Guide to'Fake News' and Other Information Disorders. A Field Guide to" Fake News" and other information disorders: a collection of recipes for those who love to cook with digital methods. Public Data Lab, Amsterdam.
2. Li G, Dong M, Yang F, Zeng J, Yuan J, Jin C, Zheng B (2020) Misinformation-oriented expert finding in social networks. World Wide Web 23(2):693–714
3. Zubiaga A, Aker A, Bontcheva K, Liakata M, Procter R (2018) Detection and resolution of rumours in social media: a survey. ACM Comput Surveys (CSUR) 51(2):1–36
4. Chen C, Wen S, Zhang J, Xiang Y, Oliver J, Alelaiwi A, Hassan MM (2017) Investigating the deceptive information in Twitter spam. Futur Gener Comput Syst 72:319–326
5. Kumar A, Sangwan SR (2019) Rumor detection using machine learning techniques on social media. In: International conference on innovative computing and communications. Springer, Singapore, pp 213–221.
6. Alrubaian M, Al-Qurishi M, Alamri A, Al-Rakhami M, Hassan MM, Fortino G (2018) Credibility in online social networks: a survey. IEEE Access 7:2828–2855
7. Kumar A, Sharma H (2020) PROD: A potential rumour origin detection model using supervised machine learning. In: International Conference on Intelligent Computing and Smart Communication 2019. Springer, Singapore, pp 1269–1276.
8. Gu J, Wang Z, Kuen J, Ma L, Shahroudy A, Shuai B, Liu T, Wang X, Wang G, Cai J, Chen T (2018) Recent advances in convolutional neural networks. Pattern Recogn 77:354–377

9. Otter DW, Medina JR, Kalita JK (2020) A survey of the usages of deep learning for natural language processing. IEEE Trans Neural Netw Learning Syst 32(2):604–624
10. Young T, Hazarika D, Poria S, Cambria E (2018) Recent trends in deep learning based natural language processing. IEEE Comput Intell Mag 13(3):55–75
11. Yang Z, Yang D, Dyer C, He X, Smola A, Hovy E (2016) Hierarchical attention networks for document classification. In: Proceedings of the 2016 conference of the North American chapter of the association for computational linguistics: human language technologies, pp 1480–1489.
12. Peters ME, Neumann M, Iyyer M, Gardner M, Clark C, Lee K, Zettlemoyer L (2018) Deep contextualized word representations. arXiv preprint arXiv:1802.05365.
13. Kumar A (2021) Contextual semantics using hierarchical attention network for sentiment classification in social internet-of-things. Multimedia Tools Appl. https://doi.org/10.1007/s11042-021-11262-8
14. Ting FF, Sim KS (2017) Self-regulated multilayer perceptron neural network for breast cancer classification. In: 2017 International Conference on Robotics, Automation and Sciences (ICORAS). IEEE, New York, pp 1–5.
15. Zubiaga A, Liakata M, Procter R (2017) Exploiting context for rumour detection in social media. In: International Conference on Social Informatics. Springer, Cham, pp 109–123.
16. Zubiaga A, Wong Sak Hoi G, Liakata M, Procter R (2016) PHEME dataset of rumours and non-rumours [Internet]. figshare; 2016 [cited 2020 May 24]. Available from: https://figshare.com/articles/PHEME_dataset_of_rumours_and_non-rumours/4010619/1
17. Cao J, Guo J, Li X, Jin Z, Guo H, Li J (2018) Automatic rumor detection on microblogs: A survey. arXiv preprint arXiv:1807.03505.
18. Takahashi T, Igata N (2012) Rumor detection on twitter. In: The 6th International Conference on Soft Computing and Intelligent Systems, and The 13th International Symposium on Advanced Intelligence Systems. IEEE, New York, pp 452–457.
19. Kumar A, Singh V, Ali T, Pal S, Singh J (2020) Empirical evaluation of shallow and deep classifiers for rumor detection. In: Advances in computing and intelligent systems. Springer, Singapore, pp 239–252.
20. Bhattacharjee U, Srijith PK, Desarkar MS (2019) Term specific tf-idf boosting for detection of rumours in social networks. In: 2019 11th international conference on communication systems & networks (COMSNETS). IEEE, New York, pp 726–731.
21. Chen T, Li X, Yin H, Zhang J (2018) Call attention to rumors: Deep attention based recurrent neural networks for early rumor detection. In: Pacific-Asia conference on knowledge discovery and data mining. Springer, Cham, pp 40–52.
22. Ma J, Gao W, Mitra P, Kwon S, Jansen BJ, Wong KF, Cha M (2016). Detecting rumors from microblogs with recurrent neural networks. In: Proceedings of IJCAI. 1, 3, 5, 6
23. Nguyen TN, Li C, Niederée C (2017) On early-stage debunking rumors on twitter: Leveraging the wisdom of weak learners. In: International conference on social informatics. Springer, Cham, pp. 141–158.
24. Guo H, Cao J, Zhang Y, Guo J, Li J (2018) Rumor detection with hierarchical social attention network. In: Proceedings of the 27th ACM international conference on information and knowledge management, pp 943–951.
25. Vijee A, Mahapatra A, Shyamkrishna A, Murthy S (2018) A hybrid approach to rumour detection in microblogging platforms. In: 2018 International conference on advances in computing, communications and informatics (ICACCI). IEEE, New York , pp 337–342.
26. Kumar A, Shrivastava A (2020) Rumour detection in benchmark dataset using attention-based residual networks. Int J Adv Sci Technol 29(3):14682. Retrieved from http://sersc.org/journals/index.php/IJAST/article/view/31956
27. Kumar A (2021) Rumour stance classification using A hybrid of capsule network and multi-layer perceptron. Turkish J Comput Math Educ (TURCOMAT) 12(13):4110–4120
28. Kumar A, Sachdeva N (2021) Multimodal cyberbullying detection using capsule network with dynamic routing and deep convolutional neural network. Multimedia Syst. https://doi.org/10.1007/s00530-020-00747-5
29. Saito T, Rehmsmeier M (2015) The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. PloS one 10(3):e0118432.
30. Hand DJ (2012) Assessing the performance of classification methods. Int Stat Rev 80(3):400–414
31. Sangwan SR, Bhatia MPS (2020) D-BullyRumbler: a safety rumble strip to resolve online denigration bullying using a hybrid filter-wrapper approach. Multimedia Syst, pp 1–17.
32. Kumar A, Dikshit S, Albuquerque VHC (2021) Explainable Artificial Intelligence for Sarcasm Detection in Dialogues. Wireless Commun Mobile Comput***.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Journal of Control, Automation and Electrical Systems
https://doi.org/10.1007/s40313-021-00864-0

Closed Loop Control Using RSAPS Algorithm for 5-Level CHB Multilevel Inverter

Praveen Bansal[1] · **Alka Singh**[1]

Received: 30 July 2021 / Revised: 12 October 2021 / Accepted: 28 October 2021
© Brazilian Society for Automatics--SBA 2021

Abstract
This paper discusses the design and implementation of a Robust Shrine Affine Projection Sign (RSAPS) adaptive control algorithm for power quality improvement. Various limitations of conventional 2-level inverters can be overcome using a 5-level Cascaded H-Bridge multilevel inverter (CHB-MLI). Such a CHB-MLI is realized and controlled as a distribution static compensator (DSTATCOM) unit in this paper. The DC link of capacitors have been maintained to reference value by using proportional-integral (PI) controllers without using complex balancing circuits. Extensive simulation has been performed and a scaled -down experimental prototype model has been developed in the laboratory to test the system under steady state and dynamic load conditions using dSPACE-1104 digital signal processor. The simulation and experimental results show the source current harmonics, DC voltage regulation, power factor correction, weight updation etc. under varying load. The proposed algorithm gives satisfactory performance in terms of weight convergence, computational burden, mean square error (MSE) etc. under dynamic load conditions and in addition the THD obtained in source voltage and current is within acceptable IEEE-519 limits.

Keywords Adaptive algorithm · DSTATCOM · Multilevel inverter · Nonlinear load · Power quality

List of Symbols

V_S	Source voltage
i_S	Source current
i_L	Load current
V_{DC1}	DC link voltage across capacitor-1
V_{DC2}	DC link voltage across capacitor-2
V_{DC}	Total DC link voltage
L_f	Interfacing Inductor
k_p	Proportional gain
k_i	Integral gain
CHB-MLI	Cascaded H-Bridge multilevel Inverter
PS-PWM	Phase Shift Pulse width modulation
V_{sp}	In-phase component of voltage
V_{sq}	Quadrature component of voltage
MV	Medium voltage
HV	High voltage
PQ	Power quality
DSTATCOM	Distributed static compensator
THD	Total harmonic distortion

✉ Alka Singh
alkasingh.eed@gmail.com

Praveen Bansal
pbansal444@gmail.com

[1] Delhi Technological University, Delhi, India

1 Introduction

Power electronic devices and systems have transformed the human life and penetrated modern power systems. They possess major benefits in the form of higher efficiency, reliability and low maintenance, however, power quality problems in the form of harmonic injection by these devices and poor power factor operation cannot be overlooked. Some of the major challenges for power engineers and researchers especially in case of medium voltage (MV) distribution systems deal with mitigation of power quality (PQ) problems using suitable techniques. Low voltage (LV) to MV AC distribution systems incorporate different types of converters used for applications in variable speed loads, traction drive and power conversion equipment in industries and renewable energy integration. Increased application of power electronic based systems has manifested into numerous PQ problems such as harmonics, deterioration of voltage, voltage sag and swell,

Published online: 10 January 2022

increased heating of conductor, reactive power burden etc. (Hong & Chen, 2011; Mahela & Shaik, 2016; Teke et al., 2011).

One of the feasible and cost-effective solutions is to design and install suitable passive filters for regulating PQ within stipulated international standards. However, passive filters are tuned for particular fixed frequency and provide fixed reactive power support. Hence, they are only partially effective for varying loads and show resonance and aging problems too. In recent years, fast and adaptive compensating devices such as unified power quality conditioner (UPQC), shunt active power filter (SAPF), dynamic voltage restorer (DVR) and DSTATCOM have been installed worldwide to mitigate different PQ problems. In this paper, DSTATCOM is used to compensate the harmonics generated by the nonlinear load and achieve unity power factor operation at supply side (Karanki et al., 2012).

DSTATCOM is commonly realized using a two-level inverter and this configuration is very widely used in LV distribution system. However, for MV to HV level, MLI is a feasible solution due to several advantages of MLI over conventional inverter. These include lower dv/dt stress across insulated gate bipolar transistor (IGBTs) switches, low peak inverse voltage (PIV), better voltage harmonics profile, lesser electromagnetic interference (EMI), availability of redundant states, modular structure etc. (Gupta & Jain, 2014; Gupta et al., 2012). Conventionally, MLIs are classified as diode clamped MLI (DCMLI), flying capacitor MLI (FCMLI) and cascaded H-bridge MLI (CHB-MLI) (Rodrigues et al., 2002, 2009). Due to simple and modular structure of CHB-MLI, it is used as DSTATCOM unit and controlled in this paper.

According to the literature review, the choice of the DSTATCOM controller has a significant impact on its performance. Effective control design affects tracking performance, robustness, accuracy and adaptability under transient conditions. In literature, various time-domain control techniques have been discussed viz. Instantaneous Reactive Power Theory (Singh et al., 2015); Synchronous Reference Frame Theory (Singh et al., 2014); Power balance Theory (Panchbhai et al., 2017) etc. These algorithms require internal tuning of parameters and substantially high arithmetical operations along with additional filters. Moreover, various frequency-domain algorithms such as Kalman Filter (Chittora et al., 2016) Fourier transform (Badoni et al., 2016) have also been implemented for shunt compensator.

Over the last few decades, several adaptive signal algorithms have been applied in domain of active noise cancellation, signal enhancement, noise filtering, echo cancellation etc. Beniwal et al. (2016) has presented adaptive Least Mean Square algorithm for grid tied PV system. Srinivas et al. (2016) have implemented combined LMS and Least Mean Forth (LMF) algorithms for PQ enhancement. The authors have discussed steady state performance improvement using LMS technique and faster convergence using LMF. Alhaj Hussam et al. (2014) have presented comparison of LMS, LMF and combined LMS-LMF algorithm. In addition, some other advanced adaptive control algorithms are mentioned in literature (Ahmad, 2017; Giri et al., 2018; Kumar et al., 2019; Mangaraj et al., 2020; Mitra & Venayagamoorthy, 2010; Modi et al., 2020; Shi & Haiquan, 2018; Shukl & Singh, 2019).

Mitra and Venayagamoorthy (2010) have discussed artificial immune adaptive algorithm for control of DSTATCOM but it requires prior knowledge of the system behavior to update the weights. Adaptive weighted leaky zero attracting LMS control algorithm has been presented by Shi and Haiquan (2018) and it does not require knowledge of system behavior yet it improves the performance over delayed LMS Algorithm. Modi et al. have presented improved Widrow–Hoff algorithm to improve grid current by extracting the fundamental load current (Modi et al., 2020) and shown that its performance outperforms that of conventional LMS algorithm. Further, variable learning and gradient least mean square (VLGLMS) algorithm has been implemented by Giri et al. (2018) to compute active component and to generate firing pulses. The authors consider this algorithm to be insensitive to step size and noise due to sensors unlike conventional least mean square algorithm. Some adaptive filters in the form of adaptive Laguerre filter for control of DSTATCOM have been discussed by Shukl and Singh (2019). This filter claims to show good performance even during abnormal grid conditions. A novel normalized Laplacian Kernel algorithm Kalman filter (NLKAKF) is reported by Kumar et al. in paper (2019) which is complex but shows satisfactory performance under abnormal grid conditions. Ahmed et al. (2017) have discussed adaptive Volterra second order filter (AVSF) for DSTATCOM control. This filter is able to converge the weights rapidly and achieve fast convergence. An adaptive Levenberg–Marquardt back-propagation (LMBP) training algorithm-based Icosθ control technique is discussed by Mangaraj et al. (2020), this algorithm shows very less oscillations in the DC link voltage as compared to other conventional algorithms. Moreover, all these algorithms have been tested on conventional two level configuration.

Some artificial neural network (ANN)-based techniques (Arya & Singh, 2014; Chittora et al., 2018; Bhattacharya & Chakraborty, 2011; Qasim & Khadkikar, 2014) have also been discussed in literature to mitigate PQ problems. Qasim et al. have discussed application of ANN for SAPF (Qasim & Khadkikar, 2014); Arya et al. have discussed Neural network based conductance estimation in Arya and Singh (2014); adaptive neuro fuzzy inference system (Chittora et al., 2018) has been reported by Badoni et al. and ANN-based predictive and adaptive control (Bhattacharya & Chakraborty, 2011) is used for controller design by Bhattacharya et al. In recent years, CHB MLI configuration can be used to connect two

independent DC sources also. Then, the proposed system can be augmented to connect two photovoltaic arrays. Ray et al. (2017) has proposed power management of grid tied PV fed CHB-MLI under distorted grid conditions. The detailed optimal efficiency calculations of PV array have been presented by Ahmed and Alhialy (2019), Gorbunova and Anisimov (2020) and the mathematical modeling of PV array has been discussed by Qerimi et al. (2020).

The closed loop performance of DSTATCOM controlled via conventional and adaptive control algorithms have been thoroughly analyzed in terms of convergence rate, learning rate parameters, self-adjustment of weights etc. It is observed that to obtain a THD < 5% in source current, almost all the authors have connected higher load on the system, which may cause additional losses in the system. However, the proposed system works well under light load conditions also.

This paper presents a new adaptive filter algorithm based on Robust Shrinkage Affine Projection Sign (RSAPS) algorithm (Shao et al., 2010). This algorithm is designed and implemented for the control of cascaded five level MLI operated as a DSTATCOM. It offers fast convergence and robust performance under dynamics, reduced oscillations and low computational burden. Simulation and experimental study have also presented and convergence of weights is compared to conventional SRFT, LMS and Recursive Least mean square (RLS) algorithms.

The organization of papers as follows: Sect. 2 discusses the system description. The basic operation of DSTATCOM is presented in Sect. 3. The mathematical formulation of CHB-MLI and computation of switching losses has been presented in Sect. 4. Section 5 presents the detailed mathematical modeling of proposed RSAP algorithm. Simulation

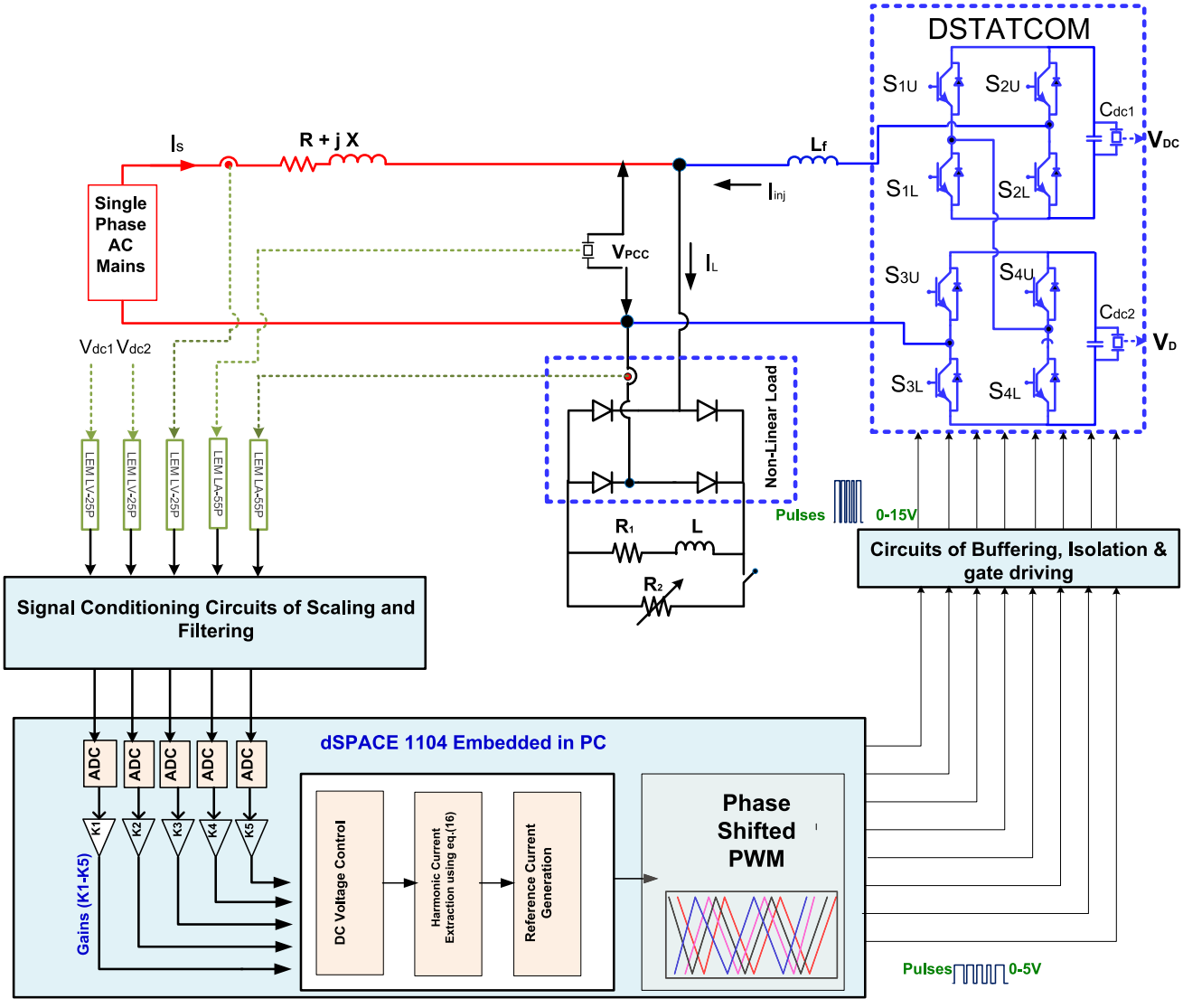


Fig. 1 Schematic diagram of RSAP Algorithm based 5-level CHB-MLI as DSTATCOM

and experimental results have been presented and analyzed in Sect. 6. Section 7 discusses the performance comparison of all algorithms in terms of weights, error, THD and finally the conclusion is presented in Sect. 8.

2 System Description

The schematic of a single phase CHB-MLI dependent DSTATCOM connected to a single-phase AC supply feeding a nonlinear load is shown in Fig. 1. DSTATCOM control is designed using sensed input variables viz. voltage of point of common coupling (PCC), source current (i_s), load current (i_L), DC link voltage across capacitor-1 (V_{DC1}) and DC link voltage across capacitor-2 (V_{DC2}). Effective operation of DSTATCOM depends upon the accurate estimation of fundamental load current which is achieved using RSAPS adaptive control algorithm. An interfacing inductor (L_f) is connected at AC output side of DSTATCOM for reactive power transfer and injection of compensating current (i_{inj}). Compensation current provided by DSTATCOM reduces harmonics in the supply current and improves supply-side power factor to unity.

3 Operation of DSTATCOM

Figure 2 depicts the layout of single-phase distribution system that connects DSTATCOM and the load at the point of common coupling (PCC). Equations (1) and (2) can be used to monitor DSTATCOM's active and reactive energy. The active power can be controlled by DSTATCOM by varying angle and the reactive power can be easily controlled by varying the voltage difference. In this paper, CHB-MLI (DSTATCOM) is controlled as a compensator and has the ability to inject/absorb reactive power like a traditional capacitor or inductor bank.

$$P_{ST} = \frac{V_{ST} V_{pcc}}{X_{ST}} \sin \delta_{ST} \tag{1}$$

$$Q_{ST} = \frac{V_{pcc}(V_{ST} \cos \delta_{ST} - V_{pcc})}{X_{ST}} \cong V_{pcc} \frac{(V_{ST} - V_{pcc})}{X_{st}} \tag{2}$$

The DSTATCOM supplies (+ Q_{ST}) reactive power to PCC when DSTATCOM output voltage V_{ST} is higher than the V_{pcc}. Similarly, it will absorb (− Q_{ST}) where V_{pcc} voltage leads the converter output voltage V_{ST}. However due to the non-idealities present in power semiconductor switches and other passive components, the converter has inherent internal losses therefore the voltage magnitude at the two capacitors continues to decrease unless switching and internal losses are compensated.

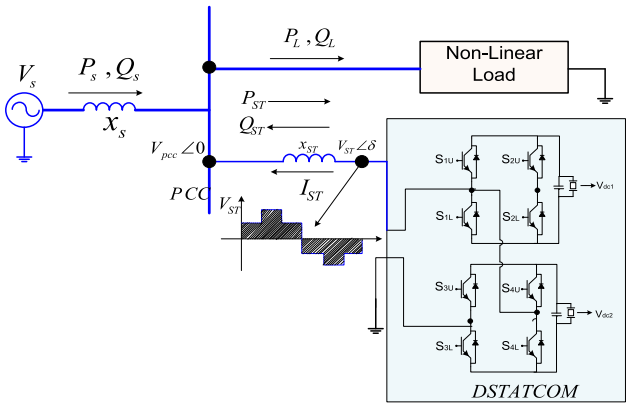


Fig. 2 Distribution system with five level cascaded MLI (DSTATCOM)

The digital implementation of proposed scheme is shown in Fig. 3. The DSTATCOM is controlled by RSAPS adaptive filtering technique. Appropriate firing pulses are generated for the control of CHB-MLI as a compensator.

4 Mathematical Formulation

The generalized switching function diagram for two-level (2L) and five-level (5L) inverters are shown in Fig. 4. For 5L MLI, the switching function is represented as S_{ak} and S_{bk} for left and right bridges. The suffix a and b represent the legs, where $k = 1, 2$ show the left and right bridges. The generalized function is given as

$$S_q = \left\{ \begin{array}{l} S_{ak} = \begin{pmatrix} +q \ if \ upper \ IGBT \ is \ ON \\ -q \ if \ lower \ IGBT \ is \ ON \end{pmatrix} \\ S_{bk} = \begin{pmatrix} +q \ if \ lower \ IGBT \ is \ ON \\ -q \ if \ upper \ IGBT \ is \ ON \end{pmatrix} \end{array} \right\} \tag{3}$$

Here, $q = 1$ represents the switching states. The output voltage 5-L MLI is

$$E_{5L}(t) = \sum_{j=1}^{n} E_j(t) \tag{4}$$

Further, the 5L output voltage can be represented in terms of switching function as given below

$$E_j(t)_{5L} = \sum_{j=1}^{n} (S_{ak} + S_{bk}) \frac{E_{dc}}{2} \tag{5}$$

Similarly, the switching function for 2L inverter and output voltage is given below

$$E_j(t)_{2L} = \sum_{j=1}^{n} (S_{pk} + S_{qk}) E_{dc} \tag{6}$$

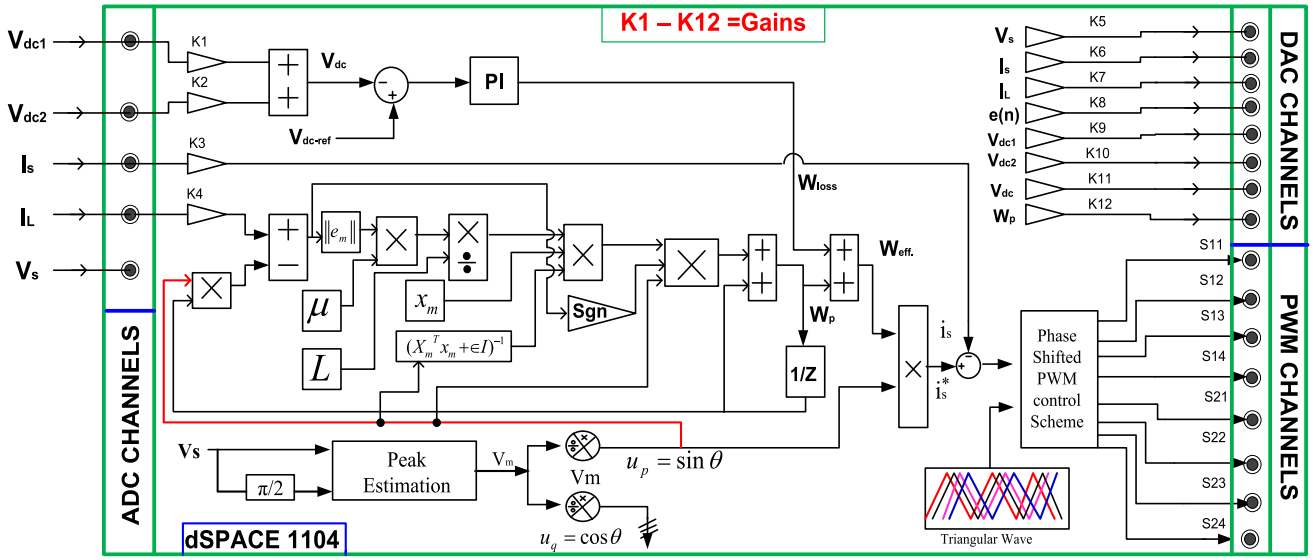


Fig. 3 Detailed block diagram of proposed controller

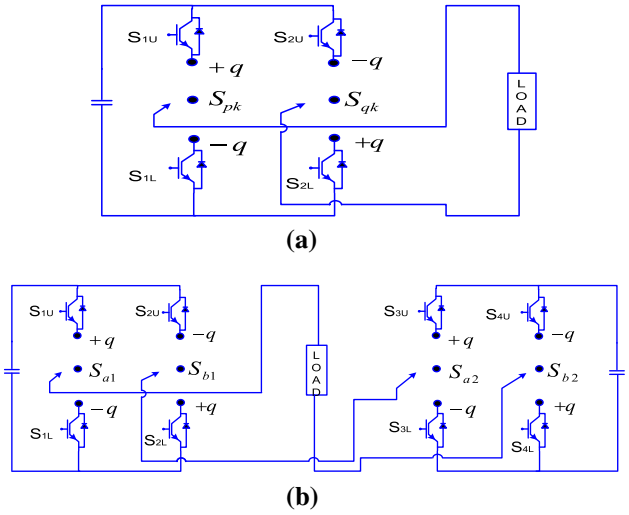


Fig. 4 Generalized model of **a** Conventional 2L inverter, **b** 5L Inverter

Table 1 Switching Levels of 2L Inverter

Mode	S_{pk}	S_{qk}	$E_j(t)_{2L}$	Voltage stress
1	+ q	− q	E_{dc}	E_{dc}
2	− q	+ q	$-E_{dc}$	E_{dc}

The switching states of 2L and 5L inverters are given in Tables 1 and 2.

The design of inverters and the cost estimation involves computation of switching loss P_{sL} and conduction loss P_{cL}.

Table 2 Switching Levels of 5L Inverter

Mode	S_{a1}	S_{a2}	S_{b1}	S_{b2}	$E_j(t)_{5L}$	Voltage stress
1	+ q	+ q	+ q	+ q	$2E_{dc}$	$\frac{E_{dc}}{2}$
2	+ q	+ q	+ q	− q	E_{dc}	$\frac{E_{dc}}{2}$
3	− q	+ q	+ q	− q	0	$\frac{E_{dc}}{2}$
4	+ q	− q	− q	− q	$-E_{dc}$	$\frac{E_{dc}}{2}$
5	− q	− q	− q	− q	$-2E_{dc}$	$\frac{E_{dc}}{2}$

The generalized expression of conduction loss (Ahmed & Alhialy, 2019) is given by

$$P_{c(loss)}(avg) = \frac{1}{\pi}\int_0^{\pi}\left[\left\{N_s(t)V_s + N_d(t)\,V_d\right\}i_c(t) + \left\{N_s(t)\,R_s i_c^{\beta+1}(t)\right\} + \left\{N_d(t) i_c^2(t)\right\}\right]\mathrm{d}(\omega t) \tag{7}$$

where each IGBT switch carries a current $i_c(t)$, conduction diodes and switches are represented as $N_d(t)$ and $N_s(t)$; corresponding ON state voltage drop is given as V_d and V_s; the ON state resistance is R_d & R_s for diode and switch, respectively and β denotes switch characteristics.

Similarly, the generalized expression for switching loss is expressed as (Gorbunova & Anisimov, 2020)

$$P_{sL} = \frac{E_j(t)\,i_c(t)(T_{ON} + T_{OFF})\,f_{sw}}{6} \tag{8}$$

where f_{sw} is the switching frequency; $E_j(t)$ is the blocking voltage; T_{ON} and T_{OFF} denote the switch ON and OFF time. Eq. (8) can further be represented as

$$P_{sL} = \delta E_j(t) f_{sw} \; where \, \delta = \frac{i_c(t)(T_{ON} + T_{OFF})}{6} \tag{9}$$

It is observed that the switching loss depends upon the blocking voltage and the switching frequency. As compared to conventional 2L inverter the blocking voltage in 5L MLI is reduced to half i.e., $\frac{E_{dc}}{2}$, therefore the switching loss expression for 5L MLI can be represented as (Gupta & Jain, 2014).

$$P_{sL(5Level)} = 4\delta \frac{E_{dc}}{2} f_{sw} \tag{10}$$

Similarly, for 2L inverter expression is given by

$$P_{sL(2Level)} = 4\delta E_{dc} f_{sw} \tag{11}$$

Thus, due to low blocking voltage in 5L inverter, the switching loss is lower as compared to same loss in conventional 2L inverter. Further, hysteresis PWM technique operating at high switching frequency is not required for 5L inverter. Instead, a phase shifted PWM of switching frequency of 5 kHz is used to control IGBTs in this paper (Tsunoda et al., 2014).

5 Modeling of Robust Shrine Affine Projection Sign Algorithm for Determining the Active Weight Element of Current Load

The block diagram of proposed algorithm is shown in Fig. 3. The current drawn by the nonlinear load is usually expressed as

$$i_L(t) = \sum_{h=1}^{\infty} i_{Lh} \sin(h\omega t + \varphi_h) \tag{12}$$

$$i_L(t) = i_{h1} \sin(\omega t + \varphi_1) + \sum_{h=2}^{\infty} i_{Lh} \sin(h\omega t + \varphi_h) \tag{13}$$

where h denotes the order of harmonics, φ_h is the harmonic phase angle and ω is the angular frequency of the fundamental harmonic, i_{L1} represents fundamental component and i_{Lh} represents the amplitude of h^{th} harmonic component. Furthermore, i_{L1} can be expressed as active and reactive components given as

$$i_{L1} = i_{L1} \cos\varphi_1 \sin\omega_1 t + i_{L1} \sin\varphi_1 \cos\omega_1 t \tag{14}$$

$$i_L(t) = i_{h1} \sin(\omega t + \varphi_1) + \sum_{h=2}^{\infty} i_{Lh} \sin(h\omega t + \varphi_h) \tag{15}$$

where $i_{mL1(p)} = i_{L1} \cos\varphi_1$ and $i_{mL1(q)} = i_{L1} \sin\varphi_1$, ω_1 is fundamental frequency component

$$i_{L1} = i_{L1(p)} + i_{L1(q)} \tag{16}$$

Here, the main objective of the DSTATCOM is to nullify the unwanted harmonics generated by the load current i.e., $i_L(t) = \sum_{h=2}^{\infty} i_{Lh} \sin(h\omega t + \varphi_h)$ and make the source current (i_s) nearly sinusoidal in phase in supply voltage. To achieve this, DSTATCOM is controlled to inject a compensating current (i_{inj}) and make source current free from harmonics. $i_{L1(q)}$ can be represented as

$$i^*_{inj} = -\sum_{h=2}^{\infty} i_{Lh} - i_{L1(q)} \tag{17}$$

Due to compensating current, the grid current is predominately sinusoidal hence can be represented as

$$i_s(t) = i_1 \sin(\omega t + \varphi_1) \tag{18}$$

Now, the RSAPS algorithm is designed by minimizing the cost function, J given as

$$\min_w . J(w) = 0.5 \left\| d_m - x_m^t w \right\| \tag{19}$$

Where $d_m = [d(m) \; d(m-1) \; .....d(m-L+1)]^T \in R^{Lx1}$ is the desired vector for system identification, $w \in R^{Mx1}$ is an adaptive weight vector and $x_m = [x_m \, x_{m-1} \, ...x_{m-L+1}] \in R^{MxL}$ is the input signal matrix. Equation (19) is made continuous function and $J(w)$ can be approximated by the minimization function

$$\widehat{J(w)} = 0.5 \left\| O_i^{0.5} (d_m - x_m^t w) \right\| \tag{20}$$

where $O_i^{0.5} = dig\left\{ \frac{1}{\sqrt{|e_i(1)|}}, \frac{1}{\sqrt{|e_i(2)|}} ..... \frac{1}{\sqrt{|e_i(L)|}} \right\}$ and e_i is the priori error signal which is commonly used for any adaptive identification algorithm and O_i is the diagonal matrix of error signal, its expression can be given by

$$e_i = d_m - x_m^t w_{i-1} \tag{21}$$

The weight updating equation of proposed RSAPSS algorithm is derived by minimizing the cost function $J(w)$, represented as

$$w_i = w_{i-1} + Y_m x_m (x_m^t x_m + \varepsilon I)^{-1} O_{i-2} e_i \tag{22}$$

where Y_m is obtained by settling the gradient $J(w)$ with respect to w equating to zero. The values of m, d_m and x_m are updated and process is carried forward until the convergence

is achieved. This approach makes Eq. (22) offline. In order to achieve real time weight updating technique, Eq. (22) is realized with i = m and hence the new weight updating equation is given by

$$w_m = w_{m-1} + Y_m x_m (x_m^t x_m + \varepsilon I)^{-1} \mathrm{sgn}(e_m) \tag{23}$$

where it is assumed that $O_{i-2} e_i \cong O_i e_i = \mathrm{sgn}(e_m)$. This approach of real time implementation is commonly used in l_1 norm of LMS algorithm but exact minimization of $\widehat{J(w)}$ is not achieved and hence Eq. (23) obtained is needed to be modified as given below

$$Y_k = \frac{\|e_k\|_1}{L} \tag{24}$$

where L is constant to be in range 1 to 5. Using Eqs. (23) and (24), the final weight updating equation is calculated as

$$w_i = w_{i-1} + \mu \frac{\|e_m\|}{L} (x_m^t x_m + \varepsilon I)^{-1} \mathrm{sgn}\, e_m \tag{25}$$

where μ is the step size parameter and a suitable value between 0 and 1 is selected.

Further the obtained RSAPS Eq. (25) is implemented for obtaining the fundamental active power component of load current as depicted in Fig. 3. For load current $i_L(m)$, desired response d(m), input vector x(m) and the weight vector w(m) is expressed as

$$x(m) = [\, x_1(m)\ x_2(m)\ x_3(m).....x_N(m)\,]^T \tag{26}$$

where $x_1(m) = [\sin w_1 m\Delta t\ \ \cos w_1 m\Delta t\ ]^T$ (18)

$$w(m) = [\, w_1(m)\ w_2(m)\ w_3(m).....w_N(m)\,]^T \tag{27}$$

and here, $w_1(m) = [w_{1p}(m)\ w_{1q}(m)]^T$ (28)

here w_{1p} and w_{1q} denote the in phase and quandrature weight component.

Using Eqs. (26) and (28), the updating expression of output error and weights can be expressed as

Where

$$e(m) = i_L(m) - \begin{bmatrix} w_{1p}(m) \\ w_{1q}(m) \\ w_{2p}(m) \\ w_{2q}(m) \\ - \\ - \\ - \\ - \\ w_{Np}(m) \\ w_{Nq}(m) \end{bmatrix}^T * \begin{bmatrix} \sin \omega_1 m\Delta t \\ \cos \omega_1 m\Delta t \\ \sin 2\omega_2 m\Delta t \\ \cos 2\omega_2 m\Delta t \\ - \\ - \\ - \\ - \\ \cos N\omega_1 m\Delta t \\ \sin N\omega_1 m\Delta t \end{bmatrix} \tag{29}$$

Load current is composed various fundamental and harmonic components can be expressed as

$$\begin{bmatrix} i_{L1}(m) \\ i_{L2}(m) \\ i_{L3}(m) \\ - \\ - \\ i_{LN}(m) \end{bmatrix} = \begin{bmatrix} i_{L1p}(m) + i_{L1q}(m) \\ i_{L2p}(m) + i_{L2q}(m) \\ i_{L3p}(m) + i_{L3q}(m) \\ - \\ - \\ i_{LNp}(m) + i_{LNq}(m) \end{bmatrix}$$

$$= \begin{bmatrix} \omega_{1p}(m).\sin \omega_{1m} m\Delta t + \omega_{1q}(m).\cos \omega_{1m} m\Delta t \\ \omega_{2p}(m).\sin 2\omega_{2m} m\Delta t + \omega_{2q}(m).\cos 2\omega_{2m} m\Delta t \\ - \\ - \\ \omega_{Np}(m).\sin N\omega_{Nm} m\Delta t + \omega_{Nq}(m).\cos N\omega_{Nm} m\Delta t \end{bmatrix} \tag{30}$$

Using Eq. (30), at mth sampling instant the DSTATCOM injects the compensating current estimated as

$$i_{inj}^*(m) = i_{L1(q)}(m) + \sum_{h=2}^{N} i_{Lh}(m) \tag{31}$$

$$i_{inj}^*(m) = \omega_{1q}(m) * \cos \omega_1 m\Delta t + \sum_{h=2}^{N} \left\{ \omega_{hp}(m) * \sinh \omega_1 m\Delta t + \omega_{hq}(m) * \cosh \omega_1 m\Delta t \right\} \tag{32}$$

5.1 Estimation of Switching Losses for CHB-MLI

In practice, due to sudden variation in the load, the system performance is affected due to unwanted transients. For proper DSTATCOM operation, it is important to monitor and regulate the DC link voltage. A classical conventional proportional- integral (PI) feedback controller can be simply

realized to minimize the error in DC link voltage. The error can be represented as

$$e_{dcerror} = V_{DCref} - V_{\mathrm{dc}} \tag{33}$$

To meet the switching loss requirement, the reference DC signal V_{DCref} is subtracted from the real DC voltage signal (measured by adding both DC link capacitor voltages). At the i$^{\mathrm{th}}$ sampling instant, the PI controller's output is as follows:

$$w_{loss} = k_p\, e_{\mathrm{dc}} + k_d \int_0^{t_d} e_{\mathrm{dc}} dt \tag{34}$$

Discretizing (34) results into

$$w_{loss}(i) = k_p e_{\mathrm{dc}}(i) + k_d \sum_{j=0}^{m} e_{\mathrm{dc}}(i)\tau_d \tag{35}$$

For $n = 1$,

$$w_{loss}(1) = k_p\, e_{\mathrm{dc}}(1) + k_i\, [e_{\mathrm{dc}}(0) + e_{\mathrm{dc}}(1)]\, \tau_d \tag{36}$$

For $n = 2$,

$$w_{loss}(2) = k_p\, e_{\mathrm{dc}}(2) + k_i\, [e_{\mathrm{dc}}(0) + e_{\mathrm{dc}}(1) + e_{\mathrm{dc}}(2)]\, \tau_d \tag{37}$$

Substituting the (33) in (34), then

$$w_{loss}(i) = w_{loss}(i-1) + k_d\, [e_{\mathrm{dc}}(i) - e_{\mathrm{dc}}(i) - e_{\mathrm{dc}}(i-1)] + k_i\, e_{\mathrm{dc}}(i)\, \tau_d \tag{38}$$

where k_p and k_i are the proportional and integral gains of the PI controller.

5.2 Estimation of Active Power Component of Load Current and Generation of Unit Vector Templates

By adding the estimated fundamental active component of load current and the loss component obtained from dc link voltages, the total active power requirement of the load is determined.

$$w_{\mathrm{eff.}} = w_{\mathrm{loss}} + w_P \tag{39}$$

The unit vector template is generated by sensing the supply voltage (V_{S}) and maximum amplitude (V_{m}) of voltage signals computed by in-phase and quadrature components. The quadrature voltage signal v_{sq} is obtained by using a time delay of $T/4$ i.e., 90° on the supply voltage. Here, T denotes the time period of complete cycle. The in-phase and quadrature signal templates (u_{p} and u_{q}) are mathematically expressed as

$$V_m = \sqrt{v_{sp}^2 + v_{sq}^2}; u_p = \frac{v_{sp}}{v_m}; u_q = \frac{v_{sq}}{v_m} \tag{40}$$

Furthermore, the controller is designed such that the grid provides the reference current i_s^* and the entire harmonic and reactive component of load current are provided by the DSTATCOM. The reference i_s^* is generated by multiplying the estimated fundamental active component of load with unit in-phase component ($\mathrm{u_p}$).

$$i_s* = w_{eff.} \times u_p \tag{41}$$

In order to produce eight PWM pulses for DSTATCOM, this reference current is compared to the phased shifted pulse width modulation (PWM) scheme (Qerimi et al., 2020).

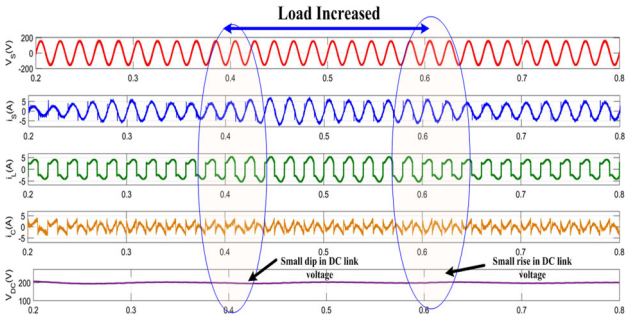


Fig. 5 System behavior under dynamic load change at $t = 0.4$ s and $t = 0.6$ s **a** Supply voltage $\mathrm{v_S}$ (V), **b** Source current i_{S} (A), **c** Load current i_{L} (A), **d** DTATCOM current i_{ST} (A) and, **d** total DC link voltage V_{DC} (V)

6 Simulation and Experimental Results

This section presents the simulation results and their analysis of the proposed system. The performance of the system is investigated under steady-state and dynamic load conditions and simulation models have been developed using MATLAB/Simulink software. The diode rectifier and R-L branch together acts as a nonlinear load which is powered by a single-phase AC supply. The response of several signals has been analyzed and presented in Figs. 5, 6, 7. Figure 5 shows plots of source voltage ($\mathrm{v_S}$), source current (i_{S}), load current ($\mathrm{i_L}$), DC link voltage across capacitor-1 (V_{DC1}), DC link voltage across capacitor-2 (V_{DC2}), total DC link voltage ($\mathrm{V_{DC}}$). In addition, some intermediate signals such as estimated fundamental weights ($\mathrm{w_{eff}}$) error signal (e) and output voltage of CHB-MLI have also been observed. The values of parameters used in simulation are given in Table 3.

Table 3 Simulation and experimental parameters

S. no	Parameters	Simulation	Experimental
1.	Source/Grid side voltage	110 V (rms) AC	110 V (rms) AC
2.	Source/Grid side inductance	Ls = 0.5 mH	Ls = 0.5 mH
3.	DC link voltage capacitor	$C_{DC} = 2800\ \mu F$	$C_{DC} = 3000\ \mu F$
4.	DC reference voltage	200 V	200 V
5.	Interfacing Inductor	3 mH	3 mH
6.	Nonlinear load	Single phase diode rectifier with $R = 80\ \Omega$ and $L = 100$ mH	Single phase diode rectifier with $R = 90\ \Omega$ and $L = 100$ mH
7.	Switching frequency of CHB-MLI	5 kHz	5 kHz
8.	Controller gains	$k_p = 1.5$ and $k_i = 0.33$	$k_p = 1.3$, $k_i = 0.96$

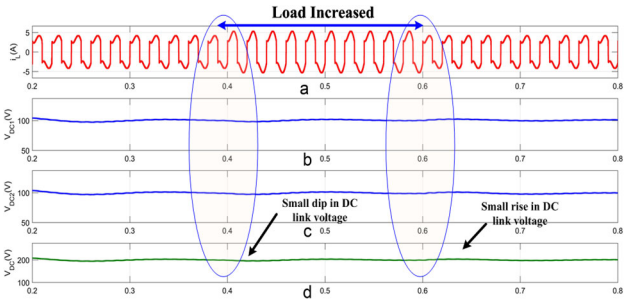


Fig. 6 System behavior under dynamic load change at $t = 0.4$ s and $t = 0.6$ s **a** Load Current i_L(A), **b** DC link voltage across capacitor-1 V_{DC1} (V), **c** DC link voltage across capacitor-2 V_{DC2} (V) and **d** Total DC link voltage V_{DC} (V)

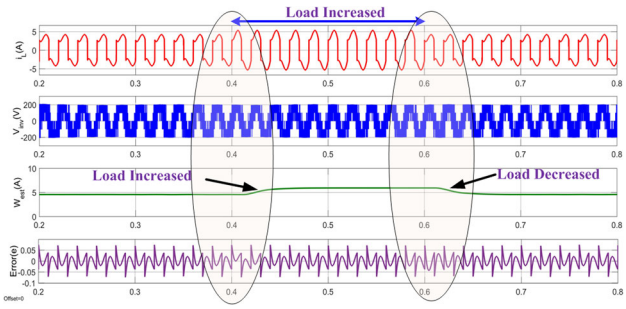


Fig. 7 Intermediate system behavior under dynamic load changes at $t = 0.4$ s and $t = 0.6$ s

In order to validate the simulated study of proposed adaptive RSAPS algorithm, a scaled down prototype model for single phase system is modeled, developed, and tested in the laboratory using a dSPACE 1104 microcontroller. A single phase 110 V (rms), 50 Hz supply feeds the nonlinear load. A bridge diode rectifier based nonlinear load is selected for experiments. The hardware setup includes inverter circuit, gate drive circuit, measuring and conditioning circuits. A CHB-MLI circuit is designed using SEMIKRON SKM75GB12T4 IGBT array having a rating of 1200 V, 75A. On DC side of inverter two ALCON make electrolytic capacitors (3000μF) are connected. Interfacing inductor of value 3mH is used to interconnect CHB-MLI to the grid. Three voltage sensors (LV-25P) are used in the proposed system to sense the voltage of the PCC (V_{pcc}) and the DC link voltage of capacitor-1 (V_{DC1}), and the DC link voltage of capacitor-2 (V_{DC2}), respectively. The load current (i_L) and source current (i_S) are also sensed using two current sensors (LA-25NP). By using correct scaling and gain correction factors, these voltage and current signals are amplified and brought back to their original value. Furthermore, the proposed control algorithm estimates reference current, as a result of which the dSPACE 1104 microcontroller produces a PWM voltage which is amplified to + 15 V using SEMIKRON make gate drive circuits (SKYPER 32 pro) to trigger DSTATCOM IGBTs. The simulation results are verified by implementing RSAPS algorithm experimentally. Table 3 lists the parameter values chosen for conducting experiments.

6.1 Design Aspects

The design of inductor and capacitor are discussed here.

6.1.1 Design of DC Reference value

The DC link reference voltage is usually selected higher than the supply voltage. (Bansal & Singh, 2021)

$$V_{dc-ref} = \frac{\sqrt{2}V_s}{\sqrt{3}m_f} \tag{42}$$

Here, $V_s = 110$ V is the source voltage, $m_f = 0.8$ is the modulation index. The dc reference value is calculated as 112.25 V for each H-bridge. A combined value of 200 V is taken as the total dc reference voltage for this system.

6.1.2 Design of (Interfacing) Inductor

The value of L_f is calculated as (Bansal & Singh, 2021)

$$L_f = \frac{V_{dc,ref}}{12\ (m-1)\ f_{sw}\ \Delta I_{cr-pp}} \tag{43}$$

$\Delta I_{cr-pp} = 0.2$ peak to peak ripple current. The approximate value of inductor is found to be 4 mH using Eq. 43, hence an inductor of 3mH is selected for experimentation as well as in simulation studies.

6.1.3 Design of DC link Capacitor

The proper designing of DC-link capacitors is needed which is calculated using the following relation (Bansal & Singh, 2021)

$$C_{\mathrm{dc}} = \frac{K_{\mathrm{L}}\ V_{\mathrm{s}}\ I_{\mathrm{c}}\ T}{2\,(V^2_{\mathrm{dc\ -\ ref}} - V^2_{\mathrm{dc}})} \tag{44}$$

where V_{s} is taken as 110 V AC, $\mathrm{I_c} = 4\mathrm{A}$, $T = 0.02$ s, $V_{\mathrm{dc\text{-}ref}} = 200$ V, $V_{\mathrm{dc}} = 190$ V, considering a voltage drop of 10 V, $\mathrm{C_{dc}} = 2707.69\ \mu\mathrm{F}$ is calculated and for this system it is taken as 3000 μF in experimentation and 2800 μF for simulation studies.

6.2 Dynamic State Simulation Results

The simulated waveform of source voltage, source current, load current, compensating current, and total DC link voltage is shown in Fig. 5. The load current is varied at time $\mathrm{t_1} = 0.4$ s and $\mathrm{t_2} = 0.6$ s and an additional nonlinear load is added and removed. The DC link voltage varies at the instant of load variation which can be observed in the waveform. A minor variation in total (and individual) DC link voltages has been observed in comparison to the reference value (200 V). However, due to the PI controller's fast action, the DC link reaches the reference value within a few cycles.

The waveforms of i_{L} (A), V_{DC1} (V), V_{DC2} (V), and V_{DC} are shown in Fig. 6. The total DC link voltage in the proposed system is the summation of the DC link voltages across individual capacitors, i.e., $V_{\mathrm{DC}} = V_{\mathrm{DC1}} + \mathrm{V_{DC2}}$.

Small variations in both the DC link voltages have been observed during sudden load variations at $t_1 = 0.4$ s and $t_2 = 0.6$ s. The tuned PI controller is easily able to regulate the total DC link voltage and keeps it at the reference value in a few cycles.

6.3 Intermediate State Simulation Results

The intermediate waveforms of i_{L}, w_{est}, error (e), and inverter output voltage (V_{inv}) are shown in Fig. 7. At $t_1 = 0.4$ s and $t_2 = 0.6$ s, the load is varied, and the computed fundamental weight component of the load current changes accordingly. This demonstrates that the proposed RSAPS algorithm is capable of accurately estimating the fundamental active power component within a few cycles and the adaptive algorithm can estimate the new weight value under varied load conditions.

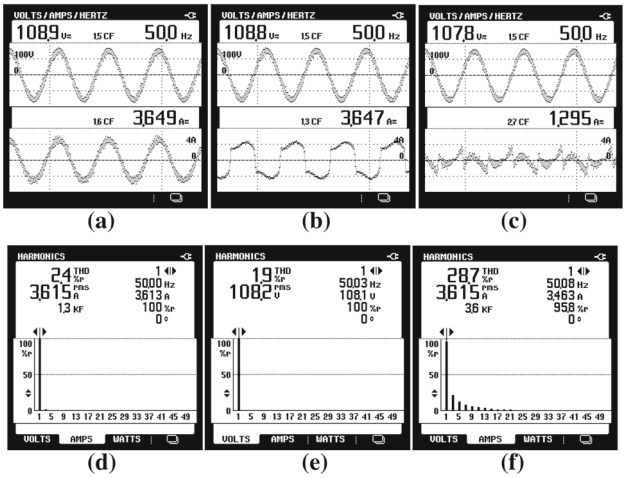


Fig. 8 Steady state performance of system **a** V_{S} and i_{S}, **b** V_{S} and i_{L}, **c** V_{S} and i_{C} along with harmonic spectrum, **d** %THD of i_{S}, **e** %THD of V_{S} and **f** % THD of $\mathrm{i_L}$

6.4 Steady State Experimental Results

Steady state experimental results of the proposed system are presented in Fig. 8a–e. It depicts the voltage of PCC (V_{pcc}), source current (i_{s}), load current (i_{L}) and compensating current (i_{C}). Source current follows load current with a THD of 28.7% before compensation. The THD of source current is reduced to 2.4% after compensation. The injected DSTATCOM current mitigates the harmonics generated by the nonlinear load. The %THD of PCC voltage is 1.9% as shown in Fig. 8.

The steady state active and reactive power flow in the proposed system is shown in Fig. 8a–c. As shown in Fig. 8, the active and reactive power demand of the load is 0.364 kW and 0.392 kVAR, respectively. In order to satisfy the load and switching loss requirements of DSTATCOM, 0.393 kW of active power is supplied from the source. The load's reactive power requirement is 0.146kVAR, which is met by the DSTATCOM, resulting in a system power factor improvement of 0.93 to 0.99, or nearly unity, after compensation. Furthermore, the supply current is nearly sinusoidal and in phase with the source voltage. It is clearly seen from recorded waveforms that the proposed system works satisfactory and the obtained %THD of V_{pcc} (V) and i_{S} (A) is within stipulated IEEE-519 standards. The steady state power requirements

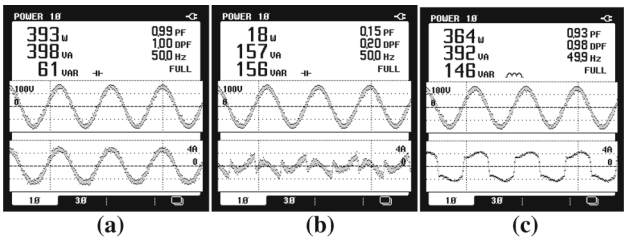


Fig. 9 Waveforms of active and reactive power flow of the proposed system in steady state

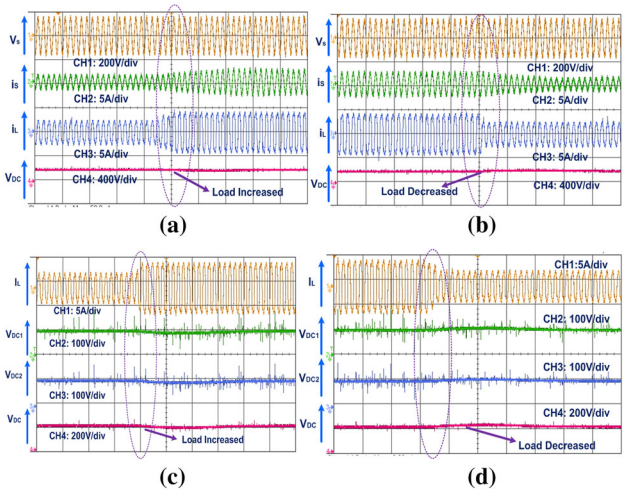


Fig. 10 Dynamic behavior of the proposed system

in the system is shown in Fig. 9a–c. The load has real and reactive power demand of 364 W and 146VAR and the DSTATCOM meets the reactive power demand of the load by feeding 156VAR. The grid supplies 393 W to meet the real power of the load along with the switching losses.

6.4.1 Dynamic State Experimental Results

The performance of the proposed system under dynamic load conditions is shown in Fig. 10. The waveforms of PCC voltage (V_{pcc}), source current (i_S), load current (i_L), and total DC link voltage (V_{DC}) are shown in Fig. 10a–d. As shown in this figure, a sudden change in load increases the load current. However, the fast action of the PI controller regulates the total DC link voltage quickly to the reference value of 200 V ± 2 V. Similarly, the performance of system under decrease in load is depicted in Fig. 10b.

Figure 10c, d shows the waveforms of i_L, V_{DC1}, V_{DC2} and V_{DC} during sudden perturbation in load. It is seen from the waveforms that the both DC link voltages of capacitors have reached ~ 100 V within few cycles. Hence the total DC

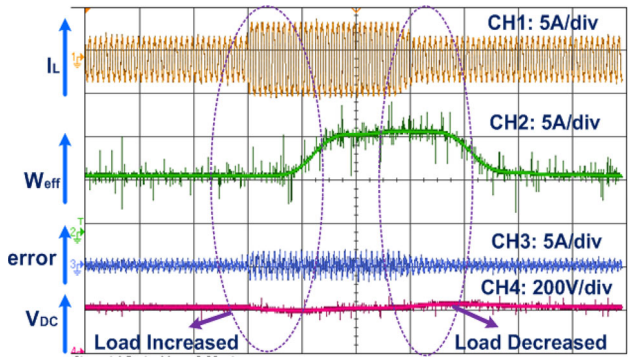


Fig. 11 Intermediate experimental results under dynamic load conditions

link voltage remains stable and well-regulated even during dynamic load conditions.

6.4.2 Intermediate State Experimental Results

The intermediate results of the system are shown in Fig. 11 under simultaneous increase and decrease in load. Figure 11 shows load current (i_L), estimated eight (w_{est}), error (e) and total DC link voltage (V_{DC}) on oscilloscope. As seen in figure that on increasing or decreasing the load, the DC link voltage slightly varies showing a voltage dip and rise, respectively. The tuned PI controller on the DC link voltage easily restores the voltage to the 200 V reference value. This result depicts satisfactory performance of RSAPS algorithm. The fundamental estimated component is effectively computed by RSAPS algorithm and all parameters reach steady state within few cycles.

The simulation and experimental results are analyzed thoroughly under steady state and dynamic load conditions. The active power and reactive power requirement of the load is supplied from source and compensator, respectively. In addition, the control is developed such that power losses occurring during switching are met from the grid. No complex balancing circuits are required to balance DC link capacitors and within 1–2 cycles, the DC link voltages in each capacitance self regulate to steady state value of 100 V due to rapid action of PI controller. On analyzing the waveforms, it is clear that the parameters such as source current (i_S), load current (i_L), compensator current (i_C) and DC link voltages of each capacitances respond very quickly during step change in load. Furthermore, the intermediate results such as fundamental estimated weight (w_{eff}) and error (e) show least oscillations and quickly respond at step change in load. The simulation and experimental THD obtained for different control algorithm is tabulated in Table 4. It is observed that the THD obtained in proposed algorithm is very less as compared to other conventional techniques.

Table 4 THD Comparison of Simulation and Experimental results

S. no	Algorithm	Simulation THD%	Experimental THD%
1.	SRFT	4.21%	4.5%
2.	LMS	3.34%	3.8%
3.	RLS	4.31%	4.5%
4.	RSAPS	2.54%	2.4%

7 Comparison of RSAPS Control Algorithm with Conventional Algorithm

The proposed RSAPS algorithm is compared with the traditional LMS and RLS algorithms in both steady and dynamic conditions. The detailed comparison is tabulated in Table 5. Adaptive control algorithm is compared in terms of estimated weight and its convergence, phase lock loop (PLL) requirement, sampling time (T_s), harmonic spectrum of source current (i_S), online training of weights (w_{est}), steady state error (e). It is observed from Fig. 12 that fast convergence is achieved using RSAPS algorithm i.e., within few cycles. The conventional LMS and RLS algorithms show a comparatively slower response.

Figure 13c depicts the RSAPS harmonic spectrum with a load current THD of 24.20% and a grid current THD of 2.54% as shown in Fig. 13a. According to the figures, the source current follows the load current with a THD of 24.20% before compensation. As shown in Fig. 13b, after compensation the THD of source current is reduced to 2.54% and THD of supply voltage is found to be 0.01% shown in Fig. 13b. In addition, the power factor of the system has been enhanced from 0.93 to 0.99. Hence the performance of proposed algorithm is satisfactory for harmonic compensation and power factor improvement.

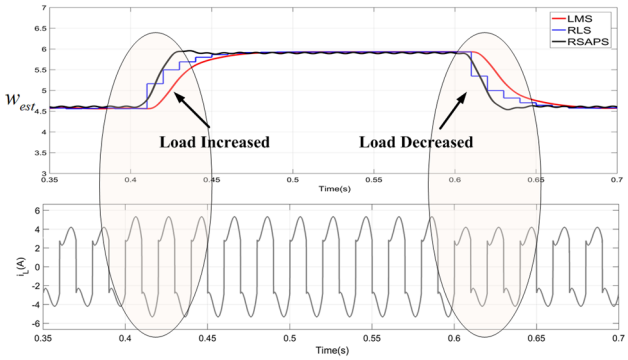


Fig. 12 Convergence of active power weights during load variation

Under variable load conditions, faster weight convergence assures superior performance and closed-loop stability of the proposed technique. These findings show that RSAPS algorithm outperforms LMS, RLS and SRFT based control algorithms, particularly in terms of transient performance and harmonic compensation. Table 5 summarizes the performance of all algorithms.

8 Conclusion

A single phase DSTATCOM has been developed using a 5-level cascaded H-bridge inverter. The proposed system has been simulated in the MATLAB/ Simulink environment and a scaled-down prototype model developed in the laboratory to test the system's performance under steady-state and dynamic loading. The RSAPS control algorithm has been

Table 5 Comparison performance of algorithms

S. no.	Parameters		RSAPS	RLS	LMS	SRFT
1	PLL requirement		No	No	No	Yes
2	Weight convergence	Steady State	Fast Convergence	Fast but slower than RSAPS	Slow convergence	Fast
		Dynamic State	Converge within 1∼2 cycles	More than 2 cycles required to converge	Sustained Oscillations	Sustained Oscillations
3	Harmonic Compensation	Grid Current (i_S) %THD	2.54%	4.31%	3.34%	4.21%
		Load current (i_L) %THD	24.20%	24.20%	24.20%	24.2%
4	Steady state error		Low	Moderate	High	Moderate
5	Training of weights		Online	Online	Online	NA
6	Complexity		Moderate	Moderate	Less	More
7	Sampling Time (Ts)		50 µs	60 µs	65 µs	70 µs

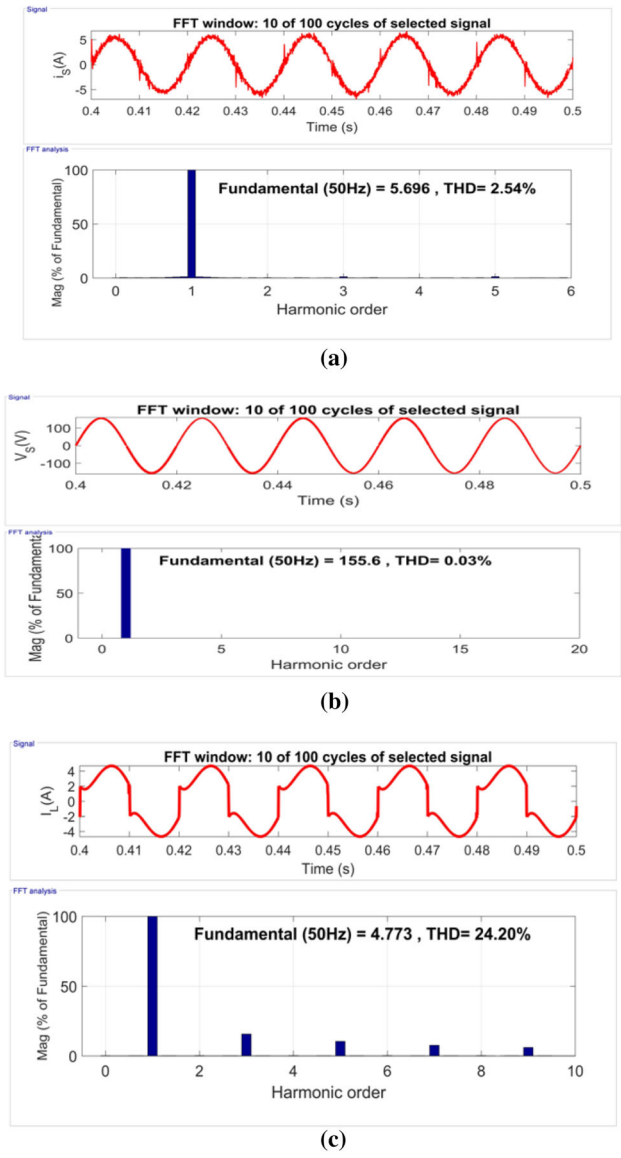


Fig. 13 Harmonic spectrum analysis of RSPAS algorithm **a** source current i_S (A), **b** source voltage V_S (V) and **c** load current i_L (A)

developed and implemented to resolve power quality issues such as harmonics and poor power factor. The proposed control algorithm shows the ability to adapt, self-organize and be able to extract the fundamental component of load current under load perturbations.

A reasonable match between experimental and simulated results suggests that adaptive control algorithms are feasible for the control of five level MLI. The RSAPS algorithm has shown good performance in terms of convergence speed, harmonic compensation, error minimization and moderate computational complexity. It has outperformed other conventional algorithms such as LMS and RLS and SRFT. Table 4 shows the comparative results for the adaptive control algorithms, namely LMS, RLS, SRFT and RSAPS. When the load current has a THD of 24.20%, the grid current distortion achieved using different controllers viz. RSAPS is 2.54% after compensation, 4.31% using LMS and 3.34% using RLS algorithm and 4.12% using SRFT. The THD of supply currents obtained is the lowest using the designed algorithm. Moreover, as compared to the conventional algorithms, the proposed algorithm achieves weight convergence within 1–2 cycles, and the dc link voltage variation is also very low, floating around ± 2 V. These findings show that the proposed RSAPS algorithm performs well in both open and closed loop conditions as well as under load variations. The proposed system may be used for medium and high voltage distribution systems to take advantage of use of multilevel inverters in terms of less PIV rating of switches, low switching frequency losses and availability of redundant states. In addition, another advantage of proposed system is the availability of two DC link sources which in future may be replaced by two PV arrays.

Declarations

Conflict of interests There are no conflicts of interest associated with this publication.

References

Ahmad, M. T., Kumar, N., & Singh, B. (2017). AVSF-based control algorithm of DSTATCOM for distribution system. *IET Generation, Transmission & Distribution, 11*(13), 3389–3396. https://doi.org/10.1049/iet-gtd.2017.0225

Ahmed, B. M., & Alhialy, N. F. F. (2019). Optimum efficiency of PV panel using genetic algorithms to touch proximate zero energy house (NZEH). *Civil Engineering Journal, 5*, 1832–1840.

Alhaj Hussam, M. M., Nor, N. M., & Asirvadam Vijanth, S., *et al.* 'Power system harmonics estimation using LMS, LMF and LMS/LMF'. In *Proc. ICIAS, 2014*.

Arya, S., & Singh, B. (2014). Neural network based conductance estimation control algorithm for shunt compensator. *IEEE Transactions on Industrial Informatics, 10*(1), 569–577.

Badoni, M., Singh, A., & Singh, B. (2016). Adaptive neuro fuzzy inference system least-mean-square-based control algorithm for DSTATCOM. *IEEE Transactions on Industrial Informatics, 12*(2), 483–492.

Bansal, P., & Singh, A. (2021). Nonlinear adaptive normalized least mean absolute third algorithm for the control of five-level distribution static compensator. *International Journal of Circuit Theory and Applications*. https://doi.org/10.1002/cta.3045

Beniwal, N., Hussain, I., Singh, B., *et al.*: 'Adaptive control scheme for three phase four wire grid tied SPV system with DSTATCOM capabilities. Proc. of2016 *National Power Systems Conf. (NPSC), Bhubaneswar, 2016, pp. 1–6*

Bhattacharya, A., & Chakraborty, C. (2011). A shunt active power filter with enhanced performance using ANN-based predictive and adaptive controllers. *IEEE Transactions on Industrial Electronics, 58*(2), 421–428. https://doi.org/10.1109/TIE.2010.2070770

Chittora, P., Singh, A., & Singh, M. Performance evaluation of a new Kalman filter based least mean square algorithm for power quality

improvement. In*International conference on power electronics, intelligent control and energy systems (ICPEICES), Delhi, India*, 2016 (pp. 1–5).

Chittora, P., Singh, A., & Singh, M. (2018) Application of hopfield neural network for harmonic current estimation and shunt compensation. *Electric Power Components and Systems, 46*(3), 290–301. https://doi.org/10.1080/15325008.2018.1445794

Giri, A. K., Arya, S. R., Maurya, R., & Mehar, R. (2018). Variable learning adaptive gradient based control algorithm for voltage source converter in distributed generation. *IET Renewable Power Generation, 12*, 1883–1892. https://doi.org/10.1049/iet-rpg.2018.5213

Gorbunova, A., & Anisimov, I. (2020). Assessment of the use of renewable energy sources for the charging infrastructure of electric vehicles. *Emerging Science Journal, 4*, 539–550. https://doi.org/10.28991/esj-2020-01251

Gupta, K. K., & Jain, S. (2014). Comprehensive review of a recently proposed multilevel inverter. *IET Power Electronics, 7*, 467–479. https://doi.org/10.1049/iet-pel.2012.0438

Gupta, K. K., & Jain, S. (2014). A novel multilevel inverter based on switched DC sources. *IEEE Transactions on Industrial Electronics, 61*(7), 3269–3278. https://doi.org/10.1109/TIE.2013.2282606

Gupta, N., Singh, S. P., & Bansal, R. C. (2012). A digital signal processor-based performance evaluation of three-phase four-wire active filter for harmonic elimination, reactive power compensation and balancing of nonlinear loads under non-ideal mains voltages. *Elec Power Compon Sys., 40*, 1119–1148.

Hong, Y. Y., & Chen, Y. Y. (2011). Placement of power quality monitors using enhanced genetic algorithm and wavelet transform. *IET Generation, Transmission & Distribution, 5*, 461–466.

Karanki, S. B., Geddada, N., Mishra, M. K., & Kumar, B. K. (2012). A DSTATCOM topology with reduced DC-Link voltage rating for load compensation with non-stiff source. *IEEE Transactions on Power Electronics, 27*(3), 1201–1211.

Kumar, N., Singh, B., Panigrahi, B. K., Chakraborty, C., Suryawanshi, H. M., & Verma, V. (2019). Integration of solar PV with low-voltage weak grid system: using normalized Laplacian kernel adaptive Kalman filter and learning based InC algorithm. *IEEE Transactions on Power Electronics, 34*(11), 10746–10758. https://doi.org/10.1109/TPEL.2019.2898319

Mahela, O. P., & Shaik, A. G. (2016). Topological aspects of power quality improvement techniques: A comprehensive overview. *Renew Sust Energy Rev., 58*, 1129–1142.

Mangaraj, M., Panda, A. K., Penthia, T., & Dash, A. R. (2020). Adaptive LMBP training-based *icosϕ* control technique for DSTATCOM. *IET Generation, Transmission and Distribution, 14*, 516–524. https://doi.org/10.1049/iet-gtd.2018.6295

Mitra, P., & Venayagamoorthy, G. K. (2010). An adaptive control strategy for DSTATCOM applications in an electric ship power system. *IEEE Transactions on Power Electronics, 25*(1), 95–104. https://doi.org/10.1109/TPEL.2009.2024152

Modi, G., Kumar, S., & Singh, B. (2020). Improved Widrow–Hoff based adaptive control of multiobjective PV-DSTATCOM system. *IEEE Transactions on Industry Applications, 56*(2), 1930–1939. https://doi.org/10.1109/TIA.2019.2960732

Panchbhai, A., Parmar, S., & Prajapati, N. Shunt active filter for harmonic and reactive power compensation using p-q theory. In *2017 International conference on power and embedded drive control (ICPEDC)*, Chennai, India, 2017 (pp. 260–264). https://doi.org/10.1109/ICPEDC.2017.8081097.

Qasim, M., & Khadkikar, V. (2014). Application of artificial neural networks for shunt active power filter control. *IEEE Transactions on Industrial Informatics, 10*(3), 1765–1774.

Qerimi, D., Dimitrieska, C., Vasilevska, S., & Alimehaj, A. (2020). Modeling of the solar thermal energy use in urban areas. *Civil Engineering Journal, 6*, 1349–1367. https://doi.org/10.28991/cej-2020-03091553

Ray, S., Gupta, N., & Gupta, R. A. (2017). Active and reactive power management of photovoltaic fed CHB inverter based active filter with improved control under normal/distorted supply. In *2017 12th IEEE conference on industrial electronics and applications (ICIEA)* (pp. 572–577). https://doi.org/10.1109/ICIEA.2017.8282909

Rodrigues, J., Jih-Sheng, L., & Peng, F. Z. (2002). Multilevel Inverters: A Survey of Topologies, Controls and Applications. *IEEE Transactions on Industrial Electronics, 49*, 724–738.

Rodriguez, J., Franquelo, L. G., Kouro, S., Leon, J. I., Portillo, R. C., Prats, M. A. M., & Perez, M. A. (2009). Multilevel converters: An enabling technology for high-power applications. *Proceedings of the IEEE, 97*, 1786–1817.

Shao, T., Zheng, Y. R., & Benesty, J. (2010). An affine projection sign algorithm robust against impulsive interferences. *IEEE Signal Processing Letters, 17*(4), 327–330.

Shi, L., & Haiquan, Z. (2018). Diffusion leaky zero attracting least mean square algorithm and its performance analysis. *IEEE Access*. https://doi.org/10.1109/ACCESS.2018.2871555

Shukl, P., & Singh, B. (2019). Grid integration of three-phase single-stage PV system using adaptive Laguerre filter based control algorithm under nonideal distribution system. *IEEE Transactions on Industry Applications, 55*(6), 6193–6202. https://doi.org/10.1109/TIA.2019.2931504

Singh, B., Chandra, A., & Al-Haddad, K. (2015). *Power quality: Problems and mitigation techniques*. John Wiley & Sons Ltd.

Singh, B., Jayaprakash, P., Kothari, D. P., Chandra, A., & Haddad, K. A. (2014). Comprehensive Study of DSTATCOM Configurations. *IEEE Transactions on Industrial Informatics, 10*(2), 854–870.

Srinivas, M., Hussain, I., & Singh, B. (2016). Combined LMS–LMF based control algorithm of DSTATCOM for power quality enhancement in distribution system. *IEEE Transactions on Industrial Electronics, 63*(7), 4160–4168.

Teke, A., Saribulut, L., & Tumay, M. (2011). A novel reference signal generation method for power-quality improvement of unified power-quality conditioner. *IEEE Transactions on Power Delivery, 26*, 2205–2214.

Tsunoda, A., Hinago, Y., & Koizumi, H. (2014). Level-and phase-shifted PWM for seven-level switched-capacitor inverter using series/parallel conversion. *IEEE Transactions on Industrial Electronics, 61*, 4011–4021.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

arXiv:2201.05803v1 [math.CV] 15 Jan 2022

Coefficient Estimates for Certain Analytic Functions

Surya Giri[1] and S. Sivaprasad Kumar[2]

[1, 2]Department of Applied Mathematics, Delhi Technological University, Delhi−110042
[1] suryagiri456@gmail.com
[2] spkumar@dce.ac.in

Abstract

Let $\lambda \in [0,1]$ and $p \in \mathbb{N}$, we introduce

$$\mathcal{S}^*_{\lambda,p}(\varphi) = \left\{ f \in \mathcal{A}_p : \frac{1}{p}\left(\frac{zf'(z)}{(1-\lambda)f(z)+\lambda z^p}\right) \prec \varphi(z) \right\},$$

where φ is a Ma-Minda function. We obtain sharp bound for a particular case of the generalized Zalcman conjecture and for some combinations involving the initial coefficients of f in $\mathcal{S}^*_{\lambda,p}(\varphi)$. Further sharp estimate for the fifth coefficient of functions belonging to $\mathcal{S}^*_{0,1}(\varphi) =: \mathcal{S}^*(\varphi)$ is deduced. Applications appear for several well-known classes for the particular choices of λ, p and $\varphi(z)$ along with a dominance result for the class $\mathcal{S}^*(\varphi)$.

Keywords- Analytic functions; Starlike functions; p-Valent functions; Fifth coefficient; Zalcman functional; Ma-Minda class.

2020 MSC: 30C45, 30C50, 30C55.

1 Introduction

Let $\mathcal{A}_p$ be the class of analytic functions of the form

$$f(z) = z^p + \sum_{n=p+1}^{\infty} a_n z^n, \quad p \in \mathbb{N}, \tag{1}$$

in the unit disk $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$ with $\mathcal{A}_1 =: \mathcal{A}$. Assume $\overline{\mathbb{D}} = \{z \in \mathbb{C} : |z| \leq 1\}$, $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$ and $\mathcal{S}$ denote the subclass of $\mathcal{A}$ containing univalent functions. Recall that the analytic function f is subordinate to another analytic function F, if there exists a Schwarz function ω such that $f(z) = F(\omega(z))$ for all $z \in \mathbb{D}$ and simply write $f \prec F$. By $\mathcal{B}_0$, we represent the class of Schwarz functions having the form

$$\omega(z) = \sum_{n=1}^{\infty} c_n z^n. \tag{2}$$

Let $\mathcal{P}$ be the Carathéodory class, contains analytic functions p defined in $\mathbb{D}$ which satisfy $\operatorname{Re} p(z) > 0$ with

$$p(z) = 1 + \sum_{n=1}^{\infty} p_n z^n, \quad z \in \mathbb{D}. \tag{3}$$

Clearly, the function

$$L(z) = \frac{1+z}{1-z}, \quad z \in \mathbb{D} \tag{4}$$

is a member of $\mathcal{P}$ as it maps the unit disk on to the right half-plane. We say function $g(z) = \sum_{n=0}^{\infty} a_n z^n$ is dominated by function $G(z) = \sum_{n=0}^{\infty} A_n z^n$ if $|a_n| \leq A_n$, for each integer $n \geq 0$ and write $g(z) \ll G(z)$, where series for function g and G are convergent in $\mathbb{D}$.

Ma and Minda [21] introduced the class

$$\mathcal{S}^*(\varphi) = \left\{ f \in \mathcal{A} : \frac{zf'(z)}{f(z)} \prec \varphi(z) \right\}, \tag{5}$$

where φ is an analytic univalent function having positive real part with $\varphi'(0) > 0$ and mapping $\mathbb{D}$ onto a domain starlike with respect to 1, which is symmetric about the real axis. Suppose, for $z \in \mathbb{D}$, φ has the following series expansion

$$\varphi(z) = 1 + B_1 z + B_2 z^2 + B_3 z^3 + \cdots, \quad B_1 > 0. \tag{6}$$

Since $\varphi(\mathbb{D})$ is symmetric with respect to the real axis, we have $\overline{\varphi(\bar{z})} = \varphi(z)$, which gives all B_i's, $i \in \mathbb{N}$ are real. For the family $\mathcal{S}^*(\varphi)$ sharp growth, distortion theorems and estimates for the coefficient functional $\left|a_3 - r a_2^2\right|$ are known, where $r \in \mathbb{R}$ [21]. In fact, the class $\mathcal{S}^*(\varphi)$ unifies several subfamilies of starlike functions. For instance, if we take $\varphi(z) = (1+z)/(1-z)$ and $(1+(1-2\alpha)z)/(1-z)$, $0 \leq \alpha < 1$, $\mathcal{S}^*(\varphi)$ reduces to the known classes $\mathcal{S}^*$ of starlike functions and $\mathcal{S}^*(\alpha)$ of starlike functions of order α respectively. For $\varphi(z) = 1 + \left(2/\pi^2\right)\left(\log\left((1+\sqrt{z})/(1-\sqrt{z})\right)\right)^2$, we obtain the class of parabolic starlike functions $\mathcal{S}_P$, which is introduced and studied by Rønning [30]. In the past, the class $\mathcal{S}^*(\varphi)$ is studied for various alternatives to $\varphi(z)$, some of them are listed below in table 1.

Table 1: Subclasses of starlike functions obtained for different Choices of $\varphi(z)$

Class	$\varphi(z)$	**Reference**
$\mathcal{S}^*_C$	$1 + 4z/3 + 2z^2/3$	[32] K. Sharma, N. K. Jain and V. Ravichandran
$\mathcal{S}^*_{SG}$	$2/(1+e^{-z})$	[13] P. Goel and S. Kumar
$\mathcal{S}^*_{\wp}$	$1 + ze^z$	[17] S. S. Kumar and K. Gangania
$\mathcal{S}^*_L$	$\sqrt{1+z}$	[34] J. Sokół
$\mathcal{S}^*_{Ne}$	$1 + z - z^3/3$	[36] L. A. Wani and A. Swaminathan
$\mathcal{S}^*_{q_b}$	$\sqrt{1+bz}, \quad b \in (0,1)$	[35] J. Sokół
Δ^*	$z + \sqrt{1+z^2}$	[26] R. K. Raina and J Sokól
$\mathcal{S}^*_e$	e^z	[22] R. Mendiratta, S. Nagpal, and V. Ravichandran
$\mathcal{S}^*_{hpl}(t)$	$1/(1-z)^t, \quad t \in (0,1]$	[15] S. Kanas, V. S. Masih and A. Ebadian
$\mathcal{S}^*_{\varphi_o}$	$1 + \frac{z}{k}\left(\frac{k+z}{k-z}\right), \ k = 1+\sqrt{2}$	[18] S. Kumar and V. Ravichandran
$\mathcal{S}^*_{RL}$	$\sqrt{2} - (\sqrt{2}-1)\sqrt{\frac{1-z}{1+2(\sqrt{2}-1)z}}$	[23] R. Mendiratta, S. Nagpal and V. Ravichandran

Recently, R. M. Ali et al. [2] studied some subclasses of $\mathcal{A}_p$, which generalize many classes including $\mathcal{S}^*(\varphi)$ and obtained sharp bound of $\left|a_{p+2} - r a_{p+1}^2\right|$ and $|a_{p+3}|$, where $r \in \mathbb{R}$. Also, Raina and Sokól [27] introduced and studied the coefficient and radius results for the following class when $\lambda \in [0,1]$:

$$\mathcal{F}_\lambda = \left\{ f \in \mathcal{A} : \frac{zf'(z)}{(1-\lambda)f + \lambda z} \prec z + \sqrt{1+z^2} \right\}.$$

Motivated by these works, here below, we define a new subclass of $\mathcal{A}_p$.

Definition 1. *Let $\varphi(z)$ be given by (6) and $\lambda \in [0,1]$, $p \in \mathbb{N}$. Then the class $\mathcal{S}^*_{\lambda,p}(\varphi)$ is the collection of $f \in \mathcal{A}_p$ satisfying*

$$\frac{1}{p}\left(\frac{zf'(z)}{(1-\lambda)f(z) + \lambda z^p}\right) \prec \varphi(z).$$

Note that as a special case, this class includes several well known classes. For instance, when $p = 1$ and $\varphi(z) = z + \sqrt{1+z^2}$, we have $\mathcal{S}^*_{\lambda,1}(\varphi) = \mathcal{F}_\lambda$. For $\lambda = 0$ and $p = 1$, $\mathcal{S}^*_{\lambda,p}$ reduces to the class $\mathcal{S}^*(\varphi)$ and for $\lambda = p = 1$, we obtain the newly defined class

$$\mathcal{S}^*_{1,1}(\varphi) =: \mathcal{R}(\varphi) = \{ f \in \mathcal{A} : f'(z) \prec \varphi(z) \}.$$

For $\varphi(z) = (1+z)/(1-z)$, $\mathcal{R}(\varphi)$ reduces to the class $\mathcal{R}$. By taking $p = 1$, we have $\mathcal{S}^*_{\lambda,1}(\varphi) =: \mathcal{S}^*_\lambda(\varphi)$, which is more general form of the class $\mathcal{F}_\lambda$. Finding the upper bound of coefficients comes under the

coefficient problem. Since the growth, covering and distortion theorems for functions $f \in \mathcal{S}$ can be proved using the fact $|a_2| \leq 2$, so the coefficient bound plays a major role in identifying the geometric nature of the function and this in fact, is the foundation for our study of the classes introduced here. Note that sharp bound for the initial coefficients $|a_2|$, $|a_3|$ and $|a_4|$ of functions in $\mathcal{S}^*(\varphi)$ are already known, see [2, 3, 37], however, for $|a_5|$, it is still open. Sokól [33] conjectured that $|a_n| \leq 1/(2(n-1))$ for $f \in \mathcal{S}_L^*$ and Mendiratta et al. [23] conjectured that $|a_n| \leq (5-3\sqrt{2})/(2(n-1))$ for $f \in \mathcal{S}_{RL}^*$. V. Ravichandran and S. Verma [28] settled these conjectures for $n = 5$. For different choices of $\varphi(z)$ the bound for $|a_5|$ is known [10, 13, 17], but not for the general case. Recently, O. S. Kwon et al. [19] established fifth coefficient bound for functions belonging to Bazilevič class. Motivated by this paper, we obtain sharp bound for $|a_5|$ of functions in $S^*(\varphi)$, under special circumstances, covering most of the cases.

In 1999, Ma [20] proposed a conjecture for $f \in \mathcal{S}$ that

$$|J_{j,k}| := |a_j a_k - a_{j+k-1}| \leq (j-1)(k-1), \quad (j,k \in \mathbb{N} \setminus \{1\}). \tag{7}$$

It is also called generalized Zalcamn conjecture because it is a generalized form of Zalcamn conjecture $\left|a_n^2 - 2a_{n-1}\right| \leq (n-1)^2$. Ma proved the inequality (7) for univalent functions with real coefficients and starlike functions. The Importance of Zalcman's conjecture can be seen as it also implies Bieberbach conjecture. For the interesting history and relevant work, see [20, 29]. Many authors have computed upper bounds for the functional $|J_{2,3}| = |a_2 a_3 - a_4|$, a specific case of the generalized Zalcman coefficient functional $J_{j,k}$, for various classes, see [4, 9, 10]. In section 3, we obtain sharp bound of $|J_{2,3}(f)|$ for the above specified classes.

We require the following lemmas to prove our results:

Lemma 1. *[24, Lemma I] If the functions $1 + \sum_{n=1}^{\infty} b_n z^n$ and $1 + \sum_{n=1}^{\infty} c_n z^n$ are in $\mathcal{P}$, then the same holds for the function*

$$1 + \frac{1}{2} \sum_{n=1}^{\infty} b_n c_n z^n.$$

Lemma 2. *[24, Lemma II] Let $h(z) = 1 + u_1 z + u_2 z^2 + \cdots$ and $1 + G(z) = 1 + d_1 z + d_2 z^2 + \cdots$ be functions in $\mathcal{P}$, and set*

$$\gamma_n = \frac{1}{2^n} \left[1 + \frac{1}{2} \sum_{k=1}^{n} \binom{n}{k} u_k \right], \quad \gamma_0 = 1.$$

If A_n is defined by

$$\sum_{n=1}^{\infty} (-1)^{n+1} \gamma_{n-1} G^n(z) = \sum_{n=1}^{\infty} A_n z^n,$$

then $|A_n| \leq 2$.

It is worth recalling the Möbius function Ψ_ζ which maps the unit disk onto the unit disk and given by

$$\Psi_\zeta(z) = \frac{z - \zeta}{1 - \overline{\zeta} z}, \quad \zeta \in \mathbb{D}. \tag{8}$$

Lemma 3. *[8, Lemma 2.4] If $p \in \mathcal{P}$, then for some $\zeta_i \in \overline{\mathbb{D}}, i \in \{1,2,3\}$,*

$$p_1 = 2\zeta_1, \tag{9}$$

$$p_2 = 2\zeta_1^2 + 2(1 - |\zeta_1|^2)\zeta_2, \tag{10}$$

$$p_3 = 2\zeta_1^3 + 4(1 - |\zeta_1|^2)\zeta_1\zeta_2 - 2(1 - |\zeta_1|^2)\overline{\zeta_1}\zeta_2^2 + 2(1 - |\zeta_1|^2)(1 - |\zeta_2|^2)\zeta_3. \tag{11}$$

For $\zeta_1, \zeta_2 \in \mathbb{D}$ and $\zeta_3 \in \mathbb{T}$, there is a unique function $p = L \circ \omega \in \mathcal{P}$ with p_1, p_2 and p_3 as in (9)-(11), where

$$\omega(z) = z\Psi_{-\zeta_1}(z\Psi_{-\zeta_2}(\zeta_3 z)), \quad z \in \mathbb{D}, \tag{12}$$

that is

$$p(z) = \frac{1 + (\overline{\zeta_2}\zeta_3 + \overline{\zeta_1}\zeta_2 + \zeta_1)z + (\overline{\zeta_1}\zeta_3 + \zeta_1\overline{\zeta_2}\zeta_3 + \zeta_2)z^2 + \zeta_3 z^3}{1 + (\overline{\zeta_2}\zeta_3 + \overline{\zeta_1}\zeta_2 - \zeta_1)z + (\overline{\zeta_1}\zeta_3 - \zeta_1\overline{\zeta_2}\zeta_3 - \zeta_2)z^2 - \zeta_3 z^3}, \quad z \in \mathbb{D}.$$

Conversely, if $\zeta_1, \zeta_2 \in \mathbb{D}$ and $\zeta_3 \in \overline{\mathbb{D}}$ are given, then we can construct a (unique) function $p \in \mathcal{P}$ of the form (3) so that $p_i, i \in \{1, 2, 3\}$, satisfy the identities in (9)-(11). For this, we define

$$\omega(z) = \omega_{\zeta_1,\zeta_2,\zeta_3}(z) = z\Psi_{-\zeta_1}(z\Psi_{-\zeta_2}(\zeta_3 z)), \quad z \in \mathbb{D}, \tag{13}$$

where Ψ_ζ is the function given as in (8). Then $\omega \in \mathcal{B}_0$. Moreover, if we define $p(z) = (1+\omega(z))/(1-\omega(z)), z \in \mathbb{D}$, then p is represented by (3), where p_1, p_2 and p_3 satisfy the identities in (9)-(11) (see the proof of [8, Lemma 2.4]).

Lemma 4. *[25, Lemma 2] If $\omega \in \mathcal{B}_0$, then for any real number q_1 and q_2, the following sharp estimate holds:*

$$|c_3 + q_1 c_1 c_2 + q_2 c_1^3| \leq L(q_1, q_2),$$

where $L(q_1, q_2)$ is as given in the original Lemma cited above.

2 Estimation of the Fifth Coefficient

We begin with the following lemma:

Lemma 5. *If $-1 < \sigma < 1$, then $F(z) = (1 + 2\sigma z + z^2)/(1 - z^2)$ belongs to $\mathcal{P}$.*

Proof. Let us consider

$$\omega(z) = \frac{F(z) - 1}{F(z) + 1} = \frac{z(z + \sigma)}{1 + \sigma z}.$$

From (8)

$$\omega(z) = z\Psi_{-\sigma}(z), \quad z \in \mathbb{D}.$$

Since $\Psi_\sigma(z)$ is a conformal automorphism of $\mathbb{D}$, $|w(z)| < 1$ and $w(0) = 0$. Which means that w is a Schwarz function and $F \in \mathcal{P}$. ■

To prove our next result, we require the following assumption, which in general is satisfied by many functions:

Assumption. *Let $\varphi(z)$ be as defined in (6), whose coefficients, satisfy the following conditions:*

$$\mathbf{C1} : 0 < B_1 (2 - B_1) - 2B_2 < 4B_1,$$

$$\mathbf{C2} : 0 < 3B_1^4 + 2B_1^3 + B_1^2 (10B_2 - 9) + 18B_2^2 - 9B_1 B_3 < 3 \left(B_1^2 + 2B_1 + 2B_2\right) \left(2B_1^2 - 3B_1 + 3B_2\right),$$

$$\begin{aligned}\mathbf{C3} : 0 < &- (B_1^2 + 2B_1 + 2B_2)\Big(9B_1^6 - 120B_1^5 + B_1^4 (48B_2 + 163) - B_1^3 (530B_2 - 216) - 18B_1(35B_2^2 \\ &- 2B_2 (9 + 10B_3) - 18B_4) - 162 \left(B_2^3 - 2B_2^2 - 2B_3^2 + 2B_2 (2B_3 + B_4)\right) + B_1^2(468B_2 + 79B_2^2 \\ &- 18 (5B_3 + 9B_4 + 18))\Big) < 16\Big(9(B_1^7 + B_1^6) + B1^5 (63B_2 - 52) + B_1^4 (32B_2 - 27) + B_1^3(164B_2^2 \\ &- 189B_2 + 9 (B_3 + 9)) + 9B_1^2 B_2(2B_2 - B_3) + 9B_1 \left(22B_2^3 - 18B_2^2 - 9B_3^2\right) + 162B_2^2 B_3\Big),\end{aligned}$$

$$\mathbf{C4} : 0 < 4B_1^2 + 6(B_2 - B_1) < 3B_1^2 + 6(B_2 - B_1).$$

Theorem 6. *Let $\varphi(z)$ be as defined in (6), whose coefficients satisfy the above conditions $C1$ to $C4$. If $f \in \mathcal{S}^*(\varphi)$, then*

$$|a_5| \leq \frac{|B_1|}{4}.$$

The inequality is sharp.

Proof. Suppose $f \in \mathcal{S}^*(\varphi)$, then

$$\frac{zf'(z)}{f(z)} = \varphi(\omega(z)),$$

where $\omega \in \mathcal{B}_0$. If we choose $\omega(z) = (p(z)-1)/(p(z)+1)$, where $p \in \mathcal{P}$, given by (3), then by comparing the coefficients obtained by series expansion of $f(z)$ together with $p(z)$ and $\varphi(z)$, yields:

$$a_5 = \frac{B_1}{8} I, \tag{14}$$

where

$$I = p_4 + I_1 p_1^4 + I_2 p_1^2 p_2 + I_3 p_1 p_3 + I_4 p_2^2, \tag{15}$$

with

$$I_1 = \frac{B_1^4 - 6B_1^3 + 11B_1^2 + 6B_1^2 B_2 - 6B_1 + 3B_2^2 - 22B_1 B_2 + 18B_2 - 18B_3 + 8B_1 B_3 + 6B_4}{48B_1},$$

$$I_2 = \frac{3B_1^3 - 11B_1^2 + 9B_1 - 18B_2 + 11B_1 B_2 + 9B_3}{12B_1}, \quad I_3 = \frac{2B_1^2 - 3B_1 + 3B_2}{3B_1}$$

and

$$I_4 = \frac{B_1^2 - 2B_1 + 2B_2}{4B_1}.$$

Let $q(z) = 1 + b_1 z + b_2 z^2 + \cdots$ be in $\mathcal{P}$. Then Lemma 1 yields

$$1 + \frac{1}{2}(p(z) - 1) * (q(z) - 1) = 1 + \frac{1}{2}\sum_{n=1}^{\infty} b_n p_n z^n \in \mathcal{P}.$$

If we assume $h(z) = 1 + \sum_{n=1}^{\infty} u_n z^n \in \mathcal{P}$ and take $1 + G(z) :== 1 + \frac{1}{2}\sum_{n=1}^{\infty} b_n p_n z^n$, then Lemma 2 yields

$$|A_4| \leq 2,$$

where

$$A_4 = \frac{1}{2}\gamma_0 b_4 p_4 - \frac{1}{4}\gamma_1 b_2^2 p_2^2 - \frac{1}{2}\gamma_1 b_1 b_3 p_1 p_3 + \frac{3}{8}\gamma_2 b_1^2 b_2 p_1^2 p_2 - \frac{1}{16}\gamma_3 b_1^4 p_1^4 \tag{16}$$

and for $i \in \{0, 1, 2, 3\}$, γ_i's are given by $\gamma_0 = 1$,

$$\gamma_1 = \frac{1}{2}\left(1 + \frac{1}{2}u_1\right), \quad \gamma_2 = \frac{1}{4}\left(1 + u_1 + \frac{1}{2}u_2\right), \quad \gamma_3 = \frac{1}{8}\left(1 + \frac{3}{2}u_1 + \frac{3}{2}u_2 + \frac{1}{2}u_3\right). \tag{17}$$

So from (15) and (16), we can observe that if there exist $q, h \in \mathcal{P}$ such that

$$b_4 = 2, \quad I_1 = -\frac{1}{128}\left(1 + \frac{3}{2}u_1 + \frac{3}{2}u_2 + \frac{1}{2}u_3\right) b_1^4, \quad I_2 = \frac{3}{32}\left(1 + u_1 + \frac{u_2}{2}\right) b_1^2 b_2,$$
$$I_3 = -\frac{1}{4}(1 + \frac{u_1}{2}) b_1 b_3, \quad I_4 = -\frac{1}{8}\left(1 + \frac{u_1}{2}\right) b_2^2,$$

then we have

$$I = A_4. \tag{18}$$

The bound for $|A_4|$ can be obtained from Lemma 2, consequently, we can estimate the bound for $|I|$ and thus we arrive at the desired bound by using (14). To prove the theorem, we construct the functions q and h in such a way that we obtain (18).

From Lemma 3, suppose that the functions q and h are constructed by taking $\zeta_1, \zeta_2 \in \mathbb{D}$, $\zeta_3 \in \overline{\mathbb{D}}$ and $\xi_1, \xi_2 \in \mathbb{D}$, $\xi_3 \in \overline{\mathbb{D}}$ respectively, as follows:

$$q = L \circ \omega_1 \quad \text{and} \quad h = L \circ \omega_2, \tag{19}$$

where

$$\omega_1(z) = z\Psi_{-\zeta_1}(z\Psi_{-\zeta_2}(\zeta_3 z)), \quad \omega_2(z) = z\Psi_{-\xi_1}(z\Psi_{-\xi_2}(\xi_3 z)) \tag{20}$$

and L is the function given by (4). So again from Lemma 3, the b_i's and u_i's, $i \in \{1, 2, 3\}$ are given by

$$b_1 = 2\zeta_1, \quad b_2 = 2\zeta_1^2 + 2(1 - |\zeta_1|^2)\zeta_2,$$
$$b_3 = 2\zeta_1^3 + 4(1 - |\zeta_1|^2)\zeta_1\zeta_2 - 2(1 - |\zeta_1|^2)\overline{\zeta_1}\zeta_2^2 + 2(1 - |\zeta_1|^2)(1 - |\zeta_2|^2)\zeta_3,$$

and

$$u_1 = 2\xi_1, \quad u_2 = 2\xi_1^2 + 2(1 - |\xi_1|^2)\xi_2,$$
$$u_3 = 2\xi_1^3 + 4(1 - |\xi_1|^2)\xi_1\xi_2 - 2(1 - |\xi_1|^2)\overline{\xi_1}\xi_2^2 + 2(1 - |\xi_1|^2)(1 - |\xi_2|^2)\xi_3.$$

There may be many solutions for the above set of equations. For our purpose, we impose some restrictions on the parameters. W take all $\xi_i \in \mathbb{R}$, then

$$\begin{aligned} u_1 &= 2\xi_1, \quad u_2 = 2\xi_1^2 + 2(1-\xi_1^2)\xi_2, \\ u_3 &= 2\xi_1^3 + 4(1-\xi_1^2)\xi_1\xi_2 - 2(1-\xi_1^2)\xi_1\xi_2^2 + 2(1-\xi_1^2)(1-\xi_2^2)\xi_3. \end{aligned} \tag{21}$$

Further, if we define

$$\begin{aligned} \xi_1 &= -\frac{B_1^2 + 2B_2}{2B_1}, \quad \xi_2 = \frac{B_1^3 - B_1^2B_2 + 18B_2^2 - 18B_1B_3}{3\left(B_1^2 + 2B_1 + 2B_2\right)\left(2B_1^2 - 3B_1 + 3B_2\right)}, \\ \xi_3 &= \Big(-9B_1^8 + 30B_1^7 - B_1^6\left(66B_2 - 5\right) + 2B_1^5\left(85B_2 - 63\right) + 4B_1^3(5B_2\left(11B_2 - 18B_3 - 9\right) \\ &\quad + 27B_3) + 4B_1^2\left(B_2^3 - 36B_2^2 - 81B_3^2 + 45B_2B_3 + 162\left(B_2 - 1\right)B_4\right) - 144B_1\left(5B_2 - 9\right)B_2B_3 \\ &\quad + 324B_2(-2B_3^2 + B_2\left(\left(B_2 - 2\right)B_2 + 2B_4\right)) + 18B_1^4\left(9B_4 + 5B_3 + 6\right) - 5B_1^4B_2(35B_2 \\ &\quad - 2)\Big) \Big/ \Big(8(3B_1^4 + 2B_1^3 + 18B_2^2 + B_1^2(10B_2 - 9) - 9B_1B_3)(B_1(3B_1^2 + B_1 + 11B_2 - 9) + 9B_3)\Big), \end{aligned} \tag{22}$$

then the conditions $C1$, $C2$ and $C3$ on the coefficients B_1, B_2, B_3 and B_4 yield

$$|\xi_1| < 1, \quad |\xi_2| < 1, \quad |\xi_3| < 1.$$

Using these ξ_i in (21), we can obtain u_i, which in turn by using (17) gives

$$\begin{aligned} \gamma_1 &= \frac{1}{4}\left(2 - B_1 - \frac{2B_2}{B_1}\right), \\ \gamma_2 &= \frac{\left(B_1^2 + 2B_2 - 2B_1\right)\left(3B_1^3 - 11B_1^2 + B_1\left(11B_2 + 9\right) + 9\left(B_3 - 2B_2\right)\right)}{24B_1\left(2B_1^2 + 3B_2 - 3B_1\right)}, \\ \gamma_3 &= -\frac{1}{64B_1\left(2B_1^2 + 3B_2 - 3B_1\right)^2}\Big(3\left(B_1^2 + 2B_2 - 2B_1\right)^2\left(B_1^4 - 6B_1^3 + B_1^2\left(6B_2 + 11\right)\right. \\ &\quad \left. + B_1\left(8B_3 - 22B_2 - 6\right) + 3\left(B_2^2 + 6B_2 - 6B_3 + 2B_4\right)\right)\Big). \end{aligned} \tag{23}$$

Let us consider

$$q(z) = \frac{1 + 2\sigma z + z^2}{1 - z^2},$$

with $\sigma = \sqrt{(4B_1^2 + 6(B_2 - B_1)/((3B_1^2 + 6(B_2 - B_1))}$, then

$$b_1 = b_3 = 2\sigma \quad \text{and} \quad b_2 = b_4 = 2. \tag{24}$$

If we choose B_1 and B_2 such that $0 < \sigma < 1$, which is equivalent to condition $C4$. Then by Lemma 5, we have $q \in \mathcal{P}$. On putting the values of $b_i's$ and $\gamma_i's$ obtained from (23) and (24) respectively in (16), we get (18), which together with (14) gives the desired bound for $|a_5|$.

Let the function $H : \mathbb{D} \longrightarrow \mathbb{C}$ be given by

$$H(z) = z\exp\int_0^z \frac{\varphi(t^4) - 1}{t}dt = z + \frac{B_1}{4}z^5 + \frac{1}{32}(B_1^2 + 4B_2)z^9 + \cdots,$$

where coefficients of $\varphi(z)$ satisfies the conditions $C1$ to $C4$. Then $H(0) = 0$, $H'(0) = 1$ and $zH'(z)/H(z) = \varphi(z^4)$ and hence the function $H \in \mathcal{S}^*(\varphi)$, proving the result to be sharp for the function H. ■

Choose $\varphi(z) = 1 + \sin z$, whose coefficients satisfy $C1$ to $C4$, to obtain the following result:

Corollary 6.1. *If $f(z) \in \mathcal{S}^*(\varphi)$, where $\varphi(z) = 1 + \sin z$, then*

$$|a_5| \leq \frac{1}{4}.$$

The bound is sharp.

For all below mentioned choices of φ conditions $C1$ to $C4$ are valid. Therefore, the bounds of $|a_5|$ for some of the known classes are obtained from our result as a special case.

Remark. *1.* [13, *Theorem* 4.1] *If* $\varphi(z) = 2/(1+e^{-z})$, *then* $f \in \mathcal{S}^*_{SG}$ *and* $|a_5| \leq 1/8$.

2. [28, *Theorem* 3.1] *If* $\varphi(z) = \sqrt{1+z}$, *then* $f \in \mathcal{S}^*_L$ *and* $|a_5| \leq 1/8$.

3. [10, *Theorem* 3.1] *If* $\varphi(z) = \sqrt{1+bz}$, *then* $f \in \mathcal{S}^*_{q_b}$, *and* $|a_5| \leq b/8$, *where* $b \in (0,1]$.

4. [28, *Theorem* 3.1] *If* $f \in \mathcal{S}^*_{RL}$, *then* $|a_5| \leq \left(5 - 3\sqrt{2}\right)/8$.

5. [1] *If* $\varphi(z) = ((1+z)/(1-z))^{\delta}$, *then the conditions of Theorem 6 are satisfied only for* $0 < \delta \leq \delta_0 \approx 0.350162$. *Therefore,* $|a_5| \leq \delta/2$ *for* $f \in S^*(\varphi)$ *where* $0 < \delta \leq \delta_0$.

3 Zalcman's Functional

The following result provides the bound for the combination of coefficients of functions in the class $\mathcal{S}^*_{\lambda,p}(\varphi)$. Throughout this section, we use the notations $L(q_1, q_2)$, D_i's, $i = 1$ to 12, $\omega_{-1}(z)$, $\omega_{-2}(z)$ and $\omega_0(z)$, carry their meaning as given in [25, Lemma 2].

Theorem 7. *If* $f \in \mathcal{S}^*_{\lambda,p}(\varphi)$, *then for real numbers* r, s *and* u, *the following holds*

$$\left|ra_{p+3} + sa_{p+2}a_{p+1} + ua^3_{p+1}\right| \leq \frac{pB_1 r}{(\lambda p+3)} L(q_1, q_2), \quad B_1 > 0,\ r \neq 0,$$

with

$$q_1 = \frac{pB_1^2\left(s\left(3+p\lambda\right) + r\left(1-\lambda\right)\left(3+2p\lambda\right)\right) + 2rB_2\left(p^2\lambda^2 + 3p\lambda + 2\right)}{B_1 r\left(1+p\lambda\right)\left(2+p\lambda\right)}, \tag{25}$$

and

$$\begin{aligned} q_2 =& \Big(B_1^3 p^2\left(r\left(\lambda-1\right)^2\left(p\lambda+1\right)^2 - \left(p\lambda+3\right)\left(s\left(\lambda-1\right)\left(p\lambda+1\right) - u\left(p\lambda+2\right)\right)\right) \\ & - B_1 B_2 p\left(p\lambda+1\right)^2\left(-s\left(p\lambda+3\right) + r\left(\lambda-1\right)\left(2p\lambda+3\right)\right) + B_3 r\left(p\lambda+2\right)\left(p^3\lambda^3 + 3p^2\lambda^2\right. \\ & \left. + 3p\lambda + 1\right)\Big) \Big/ \left(B_1 r\left(\lambda p+1\right)^3\left(\lambda p+2\right)\right). \end{aligned} \tag{26}$$

The result is sharp.

Proof. Since $f \in \mathcal{S}^*_{\lambda,p}(\varphi)$, there is a function $\omega \in \mathcal{B}_0$ of the form (2) satisfying

$$\frac{1}{p}\left(\frac{zf'(z)}{(1-\lambda)f(z) + \lambda z^p}\right) = \varphi(\omega(z)),$$

where and is given by (2). By comparing the coefficients of z, z^2 and z^3, we obtain

$$a_{p+1} = \frac{pB_1}{\lambda p+1} c_1, \quad a_{p+2} = \frac{p\left(\left(pB_1^2\left(1-\lambda\right) + B_2\left(\lambda p+1\right)\right)c_1^2 + B_1\left(\lambda p+1\right)c_2\right)}{\left(\lambda p+1\right)\left(\lambda p+2\right)}, \tag{27}$$

and

$$\begin{aligned} a_{p+3} =& p\Big(B_1^3 p^2 c_1^3\left(\lambda-1\right)^2 - pc_1c_2B_1^2\left(\lambda-1\right)\left(2p\lambda+3\right) + B_3c_1^3\left(p^2\lambda^2 + 3p\lambda + 2\right) + 2B_2c_1c_2(p^2\lambda^2 \\ & + 3p\lambda + 2) + B_1\left(B_2pc_1^3\left(1-\lambda\right)\left(2p\lambda+3\right) + c_3\left(p^2\lambda^2 + 3p\lambda + 2\right)\right)\Big) \Big/ (p^3\lambda^3 \\ & + 11p\lambda + 6p^2\lambda^2 + 6). \end{aligned}$$

Using the above values of a_{p+1}, a_{p+2} and a_{p+3} together with the real numbers r, s and u, a simple calculation yields

$$\left|ra_{p+3} + sa_{p+1}a_{p+2} + ua^3_{p+1}\right| = \frac{pB_1 r}{(\lambda p+3)}\left|c_3 + q_1c_1c_2 + q_2c_1^3\right|, \tag{28}$$

where q_1 and q_2 are given by (25) and (26). The desired result follows at once from (28) and Lemma 4 and the extremal functions depend on the values of q_1 and q_2. Now we give the extremal functions $H_i : \mathbb{D} \to \mathbb{C}$, satisfying:

$$\frac{1}{p}\left(\frac{zH_i'(z)}{(1-\lambda)H_i(z)+\lambda z^p}\right) = \begin{cases} \varphi(z^3) & \text{for } i=1,\ (q_1,q_2) \in D_1 \cup D_2 \cup \{(2,1)\}, \\ \varphi(z) & \text{for } i=2,\ (q_1,q_2) \in \cup_{k=3}^{7} D_k \cup \{(2,1)\}, \\ \varphi(\omega_{-1}(z)) & \text{for } i=3,\ (q_1,q_2) \in D_8 \cup D_9, \\ \varphi(\omega_0(z)) & \text{for } i=4,\ (q_1,q_2) \in D_{10} \cup D_{11} \setminus \{(2,1)\}, \\ \varphi(\omega_{-2}(z)) & \text{for } i=5,\ (q_1,q_2) \in D_{12}. \end{cases}$$

Which completes the proof. ■

If we put $s = u = 0$ and $r = 1$ in Theorem 7, we obtain the following result:

Corollary 7.1. *If $f \in \mathcal{S}^*_{\lambda,p}(\varphi)$, then*

$$|a_{p+3}| \leq \frac{pB_1}{\lambda p+3} L(q_1,q_2),$$

where

$$q_1 = \frac{B_1^2 p\,(1-\lambda)\,(2p\lambda+3) + 2B_2\left(p^2\lambda^2+3p\lambda+2\right)}{B_1\,(p\lambda+1)\,(p\lambda+2)},$$

$$q_2 = \frac{B_3\left(p^2\lambda^2+3p\lambda+2\right) + B_1 p\,(1-\lambda)\left(B_1^2 p\,(1-\lambda) + B_2\,(2p\lambda+3)\right)}{B_1\,(p\lambda+1)\,(p\lambda+2)}.$$

For $s = 1$, $r = -1$ and $u = 0$, Theorem 7 gives the following result:

Corollary 7.2. *If $f \in \mathcal{S}^*_{\lambda,p}(\varphi)$, then*

$$|a_{p+1}a_{p+2} - a_{p+3}| \leq \frac{pB_1}{\lambda p+3} L(q_1,q_2),$$

where

$$q_1 = \frac{B_1^2 p\lambda\,(p-2p\lambda-3) + 2B_2\left(p^2\lambda^2+3p\lambda+2\right)}{B_1\,(p\lambda+1)\,(p\lambda+2)},$$

$$q_2 = \frac{B_3\,(p\lambda+1)^2\,(p\lambda+2) + B_1 p\left(B_1^2 p\,(\lambda-1)\left(p\lambda^2+\lambda+2\right) + B_2\lambda\left(p^2\left(\lambda-2\lambda^2\right) - 5p\lambda + p - 3\right)\right)}{B_1\,(p\lambda+1)^2\,(p\lambda+2)}.$$

Similarly, by choosing different values for λ, s, p, r, u and $\varphi(z)$ in Theorem 7, we obtain several results pertaining to many known classes. Now we proceed to applications pertaining to the classes $\mathcal{S}^*_\lambda(\varphi)$, $\mathcal{F}_\lambda$, $\mathcal{S}^*(\varphi)$ and $\mathcal{R}(\varphi)$.

3.1 Applications to the classes $\mathcal{S}^*_\lambda(\varphi)$ and $\mathcal{F}_\lambda$

For $p = s = 1, r = -1$ and $u = 0$, Theorem 7 gives the bound of $|a_2a_3 - a_4|$, which is $|J_{2,3}|$, for the class $\mathcal{S}^*_\lambda(\varphi)$.

Corollary 7.3. *If $f \in \mathcal{S}^*_\lambda(\varphi)$, then*

$$|J_{2,3}(f)| \leq \frac{B_1}{(\lambda+3)} L(q_1,q_2),$$

where

$$q_1 = \frac{-2B_1^2\lambda + 2B_2\,(\lambda+2)}{B_1\,(\lambda+2)}, \quad q_2 = \frac{B_1^3\left(\lambda^3+\lambda-2\right) - 2B_1B_2\lambda\,(\lambda+1)^2 + B_3\,(\lambda+1)^2\,(\lambda+2)}{B_1\,(\lambda+1)^2\,(\lambda+2)}.$$

If we put $\varphi(z) = z + \sqrt{1+z^2}$ in corollary 7.3, then we obtain the bound of $|J_{2,3}|$ for $f \in \mathcal{F}_\lambda$.

Corollary 7.4. *If $f \in \mathcal{F}_\lambda$, then*

$$|J_{2,3}(f)| \leq \begin{cases} \dfrac{1}{\lambda+3} & \lambda \in [\lambda_0, 1], \\ \dfrac{2(1+\lambda^2)}{(\lambda+1)^2(\lambda+2)(\lambda+3)} & \lambda \in [0, \lambda_0), \end{cases}$$

where $\lambda_0 \approx 0.0639089$ is the root of the equation $27\lambda^4 + 216\lambda^3 + 439\lambda^2 + 284\lambda - 20 = 0$. Extremal functions f_1 and f_2 for $\lambda \in [\lambda_0, 1]$ and $\lambda \in [0, \lambda_0)$, respectively are given by

$$\frac{zf_1'(z)}{(1-\lambda)f_1(z)+\lambda z} = z^3 + \sqrt{1+z^6}, \quad \frac{zf_2'(z)}{(1-\lambda)f_2(z)+\lambda z} = z + \sqrt{1+z^2}.$$

3.2 Applications to the Class $\mathcal{S}^*(\varphi)$

By taking $\lambda = 0$, $p = s = 1$, $r = -1$ and $u = 0$ in Theorem 7, we obtain the bound of $|J_{2,3}(f)|$ for $f \in \mathcal{S}^*(\varphi)$.

Corollary 7.5. *If $f \in \mathcal{S}^*(\varphi)$, then*

$$|J_{2,3}(f)| \leq \frac{B_1}{3} L(q_1, q_2),$$

where

$$q_1 = \frac{2B_2}{B_1}, \quad q_2 = \frac{-B_1^3 + B_3}{B_1}.$$

Note that Corollary 7.5 yields bound of $|J_{2,3}|$ for different classes by choosing the appropriate choices of $\varphi(z)$. In Table 2, we provide bounds for $|J_{2,3}(f)|$ with extremal functions for some classes.

Table 2: Bound of $|J_{2,3}|$

Class	**Bound of $\lvert J_{2,3}\rvert$**	**Extremal functions**
$\mathcal{S}_C^*$	$64/81$	$z\exp\left((4z+z^2)/3\right)$
$\mathcal{S}_{Ne}^*$	$4/9$	$z\exp((9t-t^3)/9)$
$\mathcal{S}_L^*$	$1/6$	$z\exp\int_0^z(\sqrt{1+t^3}-1)/t)dt$
$\mathcal{S}_{SG}^*$	$1/6$	$z\exp\int_0^z (e^{t^3}-1)/(t(e^{t^3}+1))dt$
$\mathcal{S}_{\varphi_o}^*$	$1/(3\mathrm{k})$, $k = 1+\sqrt{2}$	$z\exp\int_0^z \left(t^2(k+t^3)/(k(k-t^3))\right)dt$
$\mathcal{S}_P$	$8/(3\pi^2)$	$z\exp\int_0^z(2/t\pi^2)\left(\log((1+\sqrt{t^3})/(1-\sqrt{t^3}))\right)^2 dt$

Corollary 7.6. *1. If $f \in \mathcal{S}_\wp^*$, then*

$$|J_{2,3}(f)| \leq \frac{2}{3}\sqrt{\frac{2}{5}},$$

and the equality holds for the function $f(z)$ given by $zf'(z)/f(z) = \varphi(\omega_{-1}(z))$, where $\omega_{-1}(z) = c_1^{-1}z + c_2^{-1}z^2 + c_3^{-1}z^3 + \cdots$ with $c_1^{-1} = \sqrt{2/5}$, $c_2^{-1} = -3/5$ and $c_3^{-1} = -(3/5)\sqrt{(2/5)}$.

2. If $f \in \mathcal{S}_e^*$, then

$$|J_{2,3}(f)| \leq \frac{8}{9\sqrt{7}},$$

and the equality holds for the function $f(z)$ given by $zf'(z)/f(z) = \varphi(\omega_{-1}(z))$, where $\omega_{-1}(z) = c_1^{-1}z + c_2^{-1}z^2 + c_3^{-1}z^3 + \cdots$ with $c_1^{-1} = 2/\sqrt{7}$, $c_2^{-1} = -3/7$ and $c_3^{-1} = -6/(7\sqrt{7})$.

Remark. *For different choices of $\varphi(z)$ in Corollary 7.5, we obtain several known results, which are listed below:*

1. *If $f \in \Delta^*$, then $|J_{2,3}(f)| \leq 4\sqrt{6}/27$ [10, Theorem 2.1].*

2. *If $f \in \mathcal{S}^*_{q_b}$, $b \in (0,1)$, then $|J_{2,3}(f)| \leq b/6$ [10, Theorem 3.1].*

3. *For $f \in \mathcal{S}^*$, then $|J_{2,3}(f)| \leq 2$ [4, Theorem 3.3].*

4. *For $\varphi(z) = (1+(1-2\alpha)z)/(1-z)$, where $\alpha \in (0,1]$, Corollary 7.5 reduces to a result due to Cho et al. [9, Theorem 2.1].*

For $s = u = \lambda = 0$ and $p = r = 1$, Theorem 7 gives the following bound of $|a_4|$ for the class $\mathcal{S}^*(\varphi)$, which we can also obtain from [2, Theorem 1].

Corollary 7.7. *If $f \in \mathcal{S}^*(\varphi)$, then $|a_4| \leq (B_1/3)\, L(q_1, q_2)$, where $q_1 = \left(3B_1^2 + 4B_2\right)/(2B_1)$, and $q_2 = \left(B_1^3 + 3B_1B_2 + 2B_3\right)/(2B_1)$.*

For different choices of $\varphi(z)$ in Corollary 7.7, we obtain the bound of fourth coefficient for different classes.

Corollary 7.8. 1. *If $f \in \mathcal{S}^*_C$, then $|a_4| \leq 68/81$. The result is sharp for the function $f(z)$ such that $zf'(z)/f(z) = 1 + (4/3)z + (2/3)z^2$.*

2. *If $f \in \mathcal{S}^*_{Ne}$, then $|a_4| \leq 1/2$. The result is sharp for the function $f(z)$ such that $zf'(z)/f(z) = \varphi(z^3)$, where $\varphi(z) = 1 + z - z^3/3$.*

3. *If $f \in \mathcal{S}^*_{\wp}$, then $|a_4| \leq 1/3$. The result is sharp for the function $f(z)$ such that $zf'(z)/f(z) = \varphi(z^3)$, where $\varphi(z) = 1 + ze^z$.*

4. *If $f \in \mathcal{S}^*$, then $|a_4| \leq 4$. The result is sharp for the function $f(z)$ such that $zf'(z)/f(z) = (1+z)/(1-z)$.*

Some of the already known results, which we obtain for various choices of $\varphi(z)$ in Corollary 7.7, are listed below.

Remark. 1. *If $f \in \mathcal{S}^*_{\wp}$, then $|a_4| \leq 5/6$ [17, Theorem 4.2].*

2. *If $f \in \mathcal{S}^*_e$, then $|a_4| \leq 17/36$ [22, Theorem 2.3].*

3. *If $f \in \mathcal{S}^*_{SG}$, then $|a_4| \leq 1/6$ [13, Theorem 4.1].*

4. *If $f \in \mathcal{S}^*(q)$, then $|a_4| \leq 5/12$ [10, Theorem 2.1].*

5. *If $f \in \mathcal{S}^*_{RL}$, then $|a_4| \leq (1/3)\left(5/2 - 3/\sqrt{2}\right)$ [23, Theorem 2.2].*

6. *If $f \in \mathcal{S}^*_L$, then $|a_4| \leq 1/6$ [33, Theorem 2].*

7. *If $f \in \mathcal{S}^*_{q_b}$ and $|a_4| \leq b/6$, where $b \in (0,1)$ [10, Theorem 3.1].*

8. *If $f \in \mathcal{S}^*_{hpl}(t), t \in (0,1]$. Then we have*

$$|a_4| \leq \begin{cases} t/3 & t \in (0, (\sqrt{769}-15)/34], \\ (17t^3 + 15t^2 + 4t)/36 & t \in (\sqrt{769}-15)/34, 1]. \end{cases}$$

The first inequality is sharp for the function $f(z)$ given by $zf'(z)/f(z) = \varphi(z^3)$ and for the second inequality extremal function $f(z)$ is given by $zf'(z)/f(z) = \varphi(z)$, where $\varphi(z) = 1/(1-z)^t, t \in (0,1]$ [11, Theorem 11].

3.3 Applications to the Class $\mathcal{R}(\varphi)$

For $\lambda = p = s = 1$, $r = -1$ and $u = 0$, Theorem 7 gives the following bound of $|J_{2,3}|$ for $f \in \mathcal{R}(\varphi)$.

Corollary 7.9. *If $f \in \mathcal{R}(\varphi)$, then $|J_{2,3}| \leq (B_1/4)\, L(q_1, q_2)$, where $q_1 = \left(-2B_1^2 + 6B_2\right)/(3B_1)$ and $q_2 = (-2B_1B_2 + 3B_3)/(3B_1)$.*

Example. *On taking $\varphi(z) = \sqrt{1+z}$ and $\varphi(z) = z + \sqrt{1+z^2}$ in Corollary 7.9, the following examples follow immediately.*

1. *Let $\varphi(z) = \sqrt{1+z}$ and $f \in \mathcal{R}(\varphi)$, then $|J_{2,3}(f)| \leq 1/8$. Extremal function is given by $f(z)$ such that $f'(z) = \varphi(z^3)$.*

2. *Let $\varphi(z) = z + \sqrt{1+z^2}$ and $f \in \mathcal{R}(\varphi)$, then $|J_{2,3}(f)| \leq 1/4$. Extremal function is given by $f(z)$ such that $f'(z) = \varphi(z^3)$.*

Remark. *For $\varphi(z) = (1+z)/(1-z)$ in Corollary 7.9, we obtain the result of K. Babalola [4, Theorem 3.1].*

For $s = u = 0$ and $\lambda = p = r = 1$, Theorem 7 gives the following sharp bound of $|a_4|$ for the class $\mathcal{R}(\varphi)$.

Corollary 7.10. *If $f \in \mathcal{R}(\varphi)$, then $|a_4| \leq (B_1/4)\, L(q_1, q_2)$ where $q_1 = 2B_2/B_1$ and $q_2 = B_3/B_1$.*

For different choices of $\varphi(z)$ in Corollary 7.10, we obtain the following example.

Example. 1. *Let $\varphi(z) = \sqrt{1+z}$ and $f \in \mathcal{R}(\varphi)$, then $|a_4| \leq 1/8$. Extremal function is given by $f(z)$ such that $f'(z) = \varphi(z^3)$.*

2. *Let $\varphi(z) = z + \sqrt{1+z^2}$ and $f \in \mathcal{R}(\varphi)$, then $|a_4| \leq 1/4$. Extremal function is given by $f(z)$ such that $f'(z) = \varphi(z^3)$.*

3. *Let $\varphi(z) = (1+z)/(1-z)$ and $f \in \mathcal{R}(\varphi)$, then $|J_{2,3}(f)| \leq 1/2$. Extremal function is given by $f(z)$ such that $f'(z) = \varphi(z^3)$.*

In the following theorem, we obtain the bound of $|a_{p+1}|$ and $|a_{p+2}|$ when $f \in \mathcal{S}^*_{\lambda,p}(\varphi)$.

Theorem 8. *If $f \in \mathcal{S}^*_{\lambda,p}(\varphi)$, then*

$$|a_{p+1}| \leq \frac{pB_1}{\lambda p + 1}, \quad |a_{p+2}| \leq \frac{p}{(\lambda p+1)(\lambda p+2)} \left| B_2(\lambda p + 1) + B_1^2(1-\lambda)p \right|.$$

These bounds are sharp.

Proof. From (27), we have

$$|a_{p+1}| = \frac{pB_1}{1+\lambda p} |c_1|.$$

Using the bound $|c_1| \leq 1$ (see [14, Vol I, p. 84-85]), we get the required bound of $|a_{p+1}|$. Again from (27), we obtain

$$|a_{p+2}| \leq \frac{p}{(\lambda p+1)(\lambda p+2)} \left(\left| \left(B_2 (\lambda p + 1) + B_1^2 (1-\lambda) p \right) \right| |c_1|^2 + |B_1 (1+\lambda p)| \, |c_2| \right),$$

Using the bound $|c_2| \leq 1 - |c_1|^2$ (see [14, Vol II, p. 78]) in the above inequality, required bound of $|a_{p+3}|$ follows. Equality sign holds for $f \in \mathcal{A}_p$, which satisfy

$$\frac{1}{p} \left(\frac{z f'(z)}{(1-\lambda) f(z) + \lambda z^p} \right) = \varphi(z).$$

This completes the proof. ■

Finally, we prove a result about dominance for the class $\mathcal{S}^*(\varphi)$. Brannan et al. [6] proved some results about dominance for the class $\mathcal{C}(\alpha)$ of convex functions of order α. Similar result is proved by Keopf [16] for the class $\mathcal{C}(\alpha)$, with $f \in \mathcal{A}_p$. Similarly, we prove the following result for the class $\mathcal{S}^*(\varphi)$.

Theorem 9. *Let $f \in \mathcal{S}^*(\varphi)$, then*

$$f'(z) \ll k'(z), \quad z \in \mathbb{D},$$

where $k(z) = z/(1-z)^2$ is the Koebe function.

Proof. Suppose $f \in \mathcal{S}^*(\varphi)$, so there exists $\omega \in \mathcal{B}_0$ such that

$$f'(z) = \left(\frac{f(z)}{z}\right) \varphi(\omega(z)).$$

Since $\mathcal{S}^*(\varphi) \subset \mathcal{S}^*$, so f is a starlike function. From [7, Theorem 3], we can write f in the following form

$$f(z) = \int_{|x|=1} \frac{z}{(1-xz)^2} d\mu,$$

where μ is a Borel probability measure on the unit circle. Thus we have

$$f'(z) = \left(\int_{|x|=1} \frac{1}{(1-xz)^2} d\mu\right) \varphi(\omega(z)).$$

We can rewrite the above equation

$$f'(z) = \left(\int_{|x|=1} \frac{1}{(1-x^2z^2)} \left(\frac{1+xz}{1-xz}\right) d\mu\right) \varphi(\omega(z)).$$

For fixed $x \in \partial\mathbb{D}$, $p_x := \left(\left(\frac{1+xz}{1-xz}\right) \varphi\left(\omega\left(z\right)\right)\right)^{1/2} \in \mathcal{P}$. By applying lemma [5] (see e.g. ([31, Theorem 2.21]), we have

$$p_x^2(z) \ll \left(\frac{1+z}{1-z}\right)^2.$$

Thus we get

$$\begin{aligned} f'(z) &= \int_{|x|=1} \frac{d\mu}{1-x^2z^2} \cdot p_x^2(z) \\ &= \sum_{k=0}^{\infty} z^{2k} \left(\int_{|x|=1} x^{2k} p_x^2(z) d\mu\right) \\ &\ll \sum_{k=0}^{\infty} z^{2k} \left(\frac{1+z}{1-z}\right)^2 = \frac{(1+z)}{(1-z)^3}, \end{aligned}$$

since μ has total mass one. ∎

Remark. *For $\varphi(z) = (1+z)/(1-z)$, Theorem 9 gives the well known bound $|a_n| \leq n$ for $f \in \mathcal{S}^*$.*

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgement

The work of the Surya Giri is supported by University Grant Commission, New-Delhi, India under UGC-Ref. No. 1112/(CSIR-UGC NET JUNE 2019).

References

[1] R. M. Ali and V. Singh, *On the fourth and fifth coefficients of strongly starlike functions*, Results Math. **29** (1996), no. 3-4, 197–202.

[2] R. M. Ali, V. Ravichandran and N. Seenivasagan, *Coefficient bounds for p-valent functions*, Applied Mathematics and Computation, Volume **187** (2007), Issue 1, Pages 35-46.

[3] S. Kumar, V. Ravichandran and S. Verma, Initial coefficients of starlike functions with real coefficients, Bull. Iranian Math. Soc. **43** (2017), no. 6, 1837–1854.

[4] K. O. Babalola, *On $H_3(1)$ Hankel determinant for some classes of univalent functions*, arXiv e-prints (2009): arXiv-0910.

[5] D. A. Brannan, *On coefficient problems for certain power series*, In Symposium on Complex Analysis, Canterbury, vol. 12 (1973), pp. 17-27.

[6] D. A. Brannan, J. G. Clunie, and W. E. Kirwan, *On the coefficient problem for functions of bounded boundary rotation*, Ann. Acad. Sci. Fenn. Ser. A. I. **1973**, no. 523, 18 pp.

[7] L. Brickman, D. J. Hallenbeck, T. H. MacGregor and D. R. Wilken, *Convex hulls and extreme points of families of starlike and convex mappings*, Trans. Amer. Math. Soc. 185 (1973), 413-428.

[8] N. E. Cho, B. Kowalczyk and A. Lecko, *Sharp bounds of some coefficient functionals over the class of functions convex in the direction of the imaginary axis*, Bull. Aust. Math. Soc. **100** (2019), no. 1, 86–96.

[9] N. E. Cho, B. Kowalczyk, O. S. Kwon, A. Lecko, and Y. J. Sim, *The bounds of some determinants for starlike functions of order alpha*, Bull. Malays. Math. Sci. Soc. **41** (2018), no. 1, 523–535.

[10] N. E. Cho, V. Kumar, O. S. Kwon, and Y. J. Sim, *Coefficient bounds for certain subclasses of starlike functions*, J. Inequal. Appl. **2019**, Paper No. 276, 13 pp.

[11] A. Ebadian, T. Bulboacă, N. E. Cho and E. A. Adegani, *Coefficient bounds and differential subordinations for analytic functions associated with starlike functions*, Rev. R. Acad. Cienc. Exactas Fís. Nat. Ser. A Mat. RACSAM **114** (2020), no. 3, Paper No. 128, 19 pp.

[12] K. Gangania, and S. Sivaprasad Kumar. *On Certain Generalizations of $\mathcal{S}^*(\psi)$*, Computational Methods and Function Theory (2021): 1-13.

[13] P. Goel and S. Kumar, *Certain class of starlike functions associated with modified sigmoid function. Bull*, Malays. Math. Sci. Soc. **43** (2020), no. 1, 957–991.

[14] A.W. Goodman, *Univalent Functions*, Mariner, Tampa (1983).

[15] S. Kanas, V. S. Masih and A. Ebadian, *Relations of a planar domains bounded by hyperbolas with families of holomorphic functions*, J. Inequal. Appl. **2019**, Paper No. 246, 14 pp.

[16] W. Koepf, *Coefficients of symmetric functions of bounded boundary rotation*, Proc. Amer. Math. Soc. **105** (1989), no. 2, 324–329.

[17] S. S. Kumar and K. Gangania, *A cardioid domain and starlike functions*, Anal. Math. Phys. **11** (2021), no. 2, 54.

[18] S. Kumar and V. Ravichandran, *A subclass of starlike functions associated with a rational function*, Southeast Asian Bull. Math. **40** (2016), no. 2, 199–212.

[19] O. S. Kwon, D. K. Thomas and Y. J. Sim. *The Fifth Coefficient of Bazilevič Functions*, Symmetry 12, no. 8 (2020): 1228.

[20] W. Ma, *Generalized Zalcman conjecture for starlike and typically real functions*, J. Math. Anal. Appl. **234** (1999), no. 1, 328–339.

[21] W. C. Ma and D. Minda, A unified treatment of some special classes of univalent functions. Proceedings of the Conference on Complex Analysis (Tianjin, 1992), 157–169, Conf. Proc. Lecture Notes Anal., I, Int. Press, Cambridge, MA, 1994.

[22] R. Mendiratta, S. Nagpal, and V. Ravichandran, *On a subclass of strongly starlike functions associated with exponential function*, Bull. Malays. Math. Sci. Soc. **38** (2015), no. 1, 365–386.

[23] R. Mendiratta, S. Nagpal and V. Ravichandran, *A subclass of starlike functions associated with left-half of the lemniscate of Bernoulli*, Internat. J. Math. **25** (2014), no. 9, 1450090, 17 pp.

[24] Z. Nehari and E. Netanyahu, *On the coefficients of meromorphic schlicht functions*, Proc. Amer. Math. Soc. 8 (1957), 15–23.

[25] D. V. Prokhorov and J. Szynal, *Inverse coefficients for (α, β)-convex functions*, Ann. Univ. Mariae Curie-Sklodowska Sect. A 35 (1981), 125–143.

[26] R. K. Raina and J Sokól, *Some properties related to a certain class of starlike functions*, C. R. Math. Acad. Sci. Paris **353** (2015), no. 11, 973–978.

[27] R. K. Raina and J. Sokół, *On a class of analytic functions governed by subordination*, Acta Univ. Sapientiae, Math. **11** (2019), no. 1, 144-155.

[28] V. Ravichandran and S. Verma, *Bound for the fifth coefficient of certain starlike functions*, C. R. Math. Acad. Sci. Paris **353** (2015), no. 6, 505–510.

[29] V. Ravichandran and S. Verma, *Generalized Zalcman conjecture for some classes of analytic functions*, J. Math. Anal. Appl. **450** (2017), no. 1, 592–605.

[30] F. Rønning, *Uniformly convex functions and a corresponding class of starlike functions*, Proc. Amer. Math. Soc. **118** (1993), no. 1, 189–196.

[31] G. Schober, *Univalent functions-selected topics, Lecture Notes in Math.*, Springer- Verlag, Berlin-Heidelberg-New York,vol. 478 (1975)

[32] K. Sharma, N. K. Jain and V. Ravichandran, *Starlike functions associated with a cardioid*, Afr. Mat. **27** (2016), no. 5-6, 923–939.

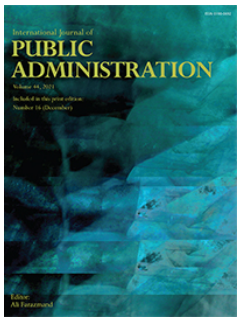
[33] J. Sokół, *Coefficient estimates in a class of strongly starlike functions*, Kyungpook Math. J. **49** (2009), no. 2, 349–353.

[34] J. Sokół and J. Stankiewicz, *Radius of convexity of some subclasses of strongly starlike functions*, Zeszyty Nauk. Politech. Rzeszowskiej Mat. No. 19 (1996), 101–105.

[35] J. Sokół, *On some subclass of strongly starlike functions*, Demonstratio Math. **31** (1998), no. 1, 81–86.

[36] L. A. Wani and A. Swaminathan, *Starlike and convex functions associated with a nephroid domain*, Bull. Malays. Math. Sci. Soc. **44** (2021), no. 1, 79–104.

[37] P. Zaprawa, On the Fekete-Szegö problem for classes of bi-univalent functions, Bull. Belg. Math. Soc. Simon Stevin **21** (2014), no. 1, 169–178.



International Journal of Public Administration

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/lpad20

Combating Workplace Toxicity with the Help of Workplace Spirituality: A Study of Indian Public Sector Oil and Gas Companies

Naval Garg, Ashi Rohtagi, Chaitanya Mittal & Deepali Sharma

To cite this article: Naval Garg, Ashi Rohtagi, Chaitanya Mittal & Deepali Sharma (2022): Combating Workplace Toxicity with the Help of Workplace Spirituality: A Study of Indian Public Sector Oil and Gas Companies, International Journal of Public Administration, DOI: 10.1080/01900692.2021.2011917

To link to this article: https://doi.org/10.1080/01900692.2021.2011917

Published online: 04 Jan 2022.

Submit your article to this journal

Article views: 81

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=lpad20

INTERNATIONAL JOURNAL OF PUBLIC ADMINISTRATION
https://doi.org/10.1080/01900692.2021.2011917

Routledge
Taylor & Francis Group

Check for updates

Combating Workplace Toxicity with the Help of Workplace Spirituality: A Study of Indian Public Sector Oil and Gas Companies

Naval Garg, Ashi Rohtagi, Chaitanya Mittal, and Deepali Sharma

Delhi Technological University, New Delhi, India

ABSTRACT

The study aimed to explore the effect of workplace spirituality on two dimensions of toxic work environments- incivility and ostracism in Indian public sector oil and gas companies. Six dimensions of workplace spirituality, namely Swadharma, Lokasangraha (Transcendence), Authenticity, Sense of community, Karma capital, and Gratitude, are explored here. The collected data is subjected to rigorous reliability, validity, multicollinearity, and common method bias tests. The association between six dimensions of workplace spirituality and two constructs of workplace toxicity is explored with the help of correlation and regression analysis. The study provides empirical evidence of a negative relationship between six Indian dimensions of workplace spirituality and two manifestations of workplace toxicity, i.e., incivility and ostracism in the workplace.

KEYWORDS

Workplace spirituality; ostracism; incivility; transcendence; gratitude

Introduction

Workplace toxicity has become a pivotal topic of discussion and researchers for the last few decades among Human Resource (HR) practitioners, academicians, and policymakers (Schilpzand et al., 2016). Cheung and Yip (2015) equated workplace toxicity with 'cancer,' which damages all stakeholders of an organization. The authors contemplated that it manifests as toxic employees, toxic leaders, toxic processes and procedures, and toxic work environment and culture. And such prevalent toxicity in every sphere of organization creates a toxic organization. Researches have revealed that toxicity in the workplace negatively influences individual-level, group-level, and organizational-level outcomes and results (Schilpzand et al., 2016).

At an individual level, workplace toxicity induces repulsive experiences. It may dent employee's motivation and morale through undesirable work practices like rejection or non-consideration of worthy ideas of an employee, social exclusion, conspiracy, backbiting, office politics, glass ceiling, passive-aggressiveness, ignoring or withholding vital information from an employee. Toxicity at the workplace may also lead to lower job satisfaction (Lim & Cortina, 2005), psychological stress, anxiety, and depression (Oyeleye et al., 2013), deterioration in physical health (Lim & Cortina, 2005), higher turnover ratio (Cortina et al., 2013), lowered employee productivity (Pouliakas & Theodossiou, 2013). At the group level, workplace toxicity damages cohesiveness, trust, team spirit, group norms, and productivity. And at the organizational level, it increases the organizational cost due to the loss of a positive company image, workplace conflict, aggression, and violence (Daniels et al., 2017; Herr et al., 2018). Researchers believe that employee-level impacts cascade to become group and organizational level effects. Thus, the focus of the present study is on exploring workplace toxicity at the individual level.

The workplace toxicity manifests in several forms, including ostracism, incivility, narcissism, bullying, aggressiveness, sexual harassment, glass ceiling, violence, passivity, and others (D. L. Ferris et al., 2015). This study examines two forms of workplace toxicity-ostracism and incivility for two reasons. Firstly, workplace incivility and ostracism are considered as a subtler and milder form of toxicity. Thus, organizations often ignore these two understated expressions of toxicity at the workplace. Although these are low-intensity deviant behaviour, their repercussions are all pervasive (Andersson & Pearson, 1999). Secondly, researchers argue that there is an upward trend in the incidence of minor forms of toxicity like workplace incivility (Estes & Wang, 2008). Frequent and repeated minor incidences are dangerous for an organization. Thus, we tend to focus only on workplace incivility and workplace ostracism.

CONTACT Naval Garg naval.garg@dtu.ac.in Assistant Professor, Delhi Technological University, New Delhi, India

© 2022 Taylor & Francis Group, LLC

It is also being argued that most of the studies in the domain of workplace toxicity have been conducted in western settings (see, Daniels et al., 2017; Tricahyadinata et al., 2020; Schilpzand et al., 2016). Thus, there is a need to explore the phenomena of workplace toxicity in non-western settings for the expansion of existing knowledge and to understand the different manifestations of the toxicity. Researchers also hailed that workplace toxicity has cultural and national dimensions. For inference, workplace ostracism is generally race or cult-based in western companies. However, In India, it is primarily based on caste, gender, or status. In this background, the present study examines workplace toxicity in Indian companies.

Further, practitioners and academicians advocate different interventions to counter the issue of toxicity in the workplace. These interventions include coaching problem employees, promoting strong communication, finding a healthy balance between authority and responsibility, etc. This study proposes workplace spirituality as one of the effective remedies to workplace toxicity. The proposed relationship between workplace spirituality and toxicity at the workplace is theoretically premised in the Conservation of Resource Theory (COR), Broaden-and-Build Theory, and the Efforts-Recovery Theory. Banking on these explanatory theories, it is being proposed that workplace spirituality develops cognitive, physical, psychological, and emotional resources that help an employee in the efficient taming of toxicity of the workplace. Further, since this study explores toxicity in Indian organizations, workplace spirituality is also conceptualized with the help of Indian dimensions of spirituality-Swadharma, Transcendence, Authenticity, Sense of community, Karma capital, and gratitude. Jurkiewicz and Giacalone (2004) recommend that "workplace spirituality has an outcropping of growing interest in Eastern philosophies that promote integration between self and environment." Thus, this study tends to explore the role of the Indian notion of workplace spirituality, which is quite comprehensive and holistic. We proposed to integrate the western understanding of workplace spirituality with Indian spiritual values through an exhaustive review of modern scientific studies and thousands of years old Indian epics and scriptures.

To achieve the desired objectives of the study, this paper is divided into four major parts. The first part gives a brief description of the Indian conceptualization of workplace spirituality. Secondly, several theories that explain the proposed association between workplace spirituality and toxicity at the workplace are addressed. In the third part of the paper, the proposed relationship is explored with the help of rigorous quantitative analysis. And finally, the theoretical and practical implications of the study are discussed, along with the limitations of the study.

Literature review

Workplace spirituality and its dimensions

The physical, psychological, and emotional aspect of work-life has been explored adequately by researchers; the spiritual side of the job has not been adequately addressed traditionally (Van der Walt, 2018). However, there is a renewed interest in exploring spirituality at the workplace in the twenty-first century (Jalan & Garg, 2021; Saxena et al., 2020). Researchers report several hurdles that restrict the smooth development of the field of spirituality at work. Spirituality has traditionally been defined with a religious connotation (Foster & Foster, 2019). Thus, call for a secular and scientific exploration of facts hinders studies in the domain of workplace spirituality. Badrinarayanan and Madhavaram (2008) claimed that the domain of spirituality at work had not flourished enough owing to lack of tools and techniques to understand individual spiritual needs. Another issue is the lack of a universally accepted definition of workplace spirituality (Garg et al., 2019; Markow & Klenke, 2005). "Defining spirituality in the workplace is like capturing an angel – it's ethereal and beautiful, but perplexing" (p. 63). Since spirituality is a culture-specific phenomenon, previous researchers have suggested different constructs of workplace spirituality.

Ashar and Lane-Maher (2004) defined workplace spirituality as an "innate and universal search for transcendent meaning in one's life . . . it involves a desire to do purposeful work that serves others and to be part of a principled community. It involves a yearning for connectedness and wholeness that can only be manifested when one is allowed to integrate his/her inner life with one's professional role in the service of a greater good" (p. 253). It is an acknowledgment of the inner life of employees at the workplace and subsequently enriching it through meaningful work profile in the context of community (Ashmos & Duchon, 2000). Giacalone and Jurkiewicz (2003) defined workplace spirituality as an organizational value-based framework that promotes workplace experience of transcendence and completeness through the work process and a sense of being connected to others. Workplace spirituality is seen as a medium to establish a connection with one's self (George et al., 2004). Garg (2020) defined Indian workplace spirituality as an experience of meaningful and meditative work done with honesty, integrity, gratefulness, and lack of fear of the unknown in a trustworthy group, and work is done for the welfare of the society. Since the present study is also based on the Indian conceptualization of workplace spirituality, the authors rely on this definition for this study.

Further, previous scholars have suggested different dimensions of workplace spirituality. Ashmos and Duchon (2000) advocated three constructs, namely meaningful work, inner life, and a sense of community. Different variables of workplace spirituality are a sense of belongingness (Harrington et al., 2001), transcendence through Work Process, and Sense of Community (Giacalone & Jurkiewicz, 2003), transcendence, meaning, purposeful work, sense of community, connectedness, and wholeness (Ashar & Lane-Maher, 2004). Pandey et al. (2009) provided four dimensions of spiritual climate- Swadharma, Lokasangraha (transcendence), Authenticity, and a sense of community. Garg (2017) proposed two more dimensions, namely, karma capital and kritgayata (gratitude), in addition to Pandey et al.'s (2016) conceptualization. Later on, lokasangraha has been referred to as transcendence and kritgayata as gratitude (Garg, 2020; Saxena et al., 2020). Accordingly, these six Indian dimensions of spirituality have been explored in the present study. These proportions of workplace spirituality represent a perfect blend of current empirical and experimental studies and ancient Indian epics like Ramayana, Shrimad Bhagwad Gita, and Upanishads.

Swadharma

Swadharma comprises of two terms: swa (self) and dharma (to uphold or to support). According to the Oxford dictionary, "dharma is seen as the cosmic law both upheld by the gods and expressed in right behaviour by humans, including adherence to the social order' (Oxford Dictionary, 2015). Bhawuk (2011) elaborates swadharma as a value that defines and explain one's behaviour and conduct in life. Modern literature defines swadharma in terms of Meaningful work" and "Meditative work" (Pandey et al., 2009). Meaningful work is ascertained through synchronization of one's purpose of life with organizational work profile and accomplishments (Pandey et al., 2009, 2016). One of the prominent reasons for employee turnover is the lack of engaging and meaningful work at the office (Saxena et al., 2020). Meditative work signifies the perfect alignment of body, mind, and spirit at work (Garg, 2020). Employees get engrossed in a meditative state of work (Pandey et al., 2019). In meditative work, the workers become one with the work (McCormick, 1994).

Transcendence (Lokasangraha)

Lokasangraha means 'coming together for the welfare of people of society (Sharma, 1999). It binds employees to work for the welfare of the general public, which leads to the awakening of one's consciousness and awareness (Garg et al., 2019). Indian spiritual tradition gives utmost importance to the service of humanity. The ultimate aim of one's life should be the upliftment of fellow beings (Shrimad BhagwadGita, Chapter 12, verse 4). The same epic claims that even God is bound to take birth again and again for the welfare of society (Shrimad Bhagwad Gita, Chapter 4, verse 7). It also involves concern for nature and the environment too (Pandey et al., 2009). R. Rastogi and Dave (2004) elaborate lokasangraha as 'service consciousness.' Modern researchers consider lokasangraha in terms of transcendence, which means 'connection to something greater than oneself' (Pandey et al., 2019; Sarkar & Garg, 2020).

Authenticity

Authenticity is a social phenomenon that supports genuineness and openness among employees. It advocates adherence to moral norms of truthfulness, non-violence, honesty, non-stealing, and celibacy (Garg, 2017c; Pandey et al., 2009). Besides being true to others, it also promotes synchronization of one's actions and behaviours with core and internalized values and beliefs (Garg, 2017b; Sarkar & Garg, 2020). Authenticity is hailed as a special virtue in Indian spiritual tradition. One of the revered Indian scripture, Mundaka Upanishad, elucidates "Satyamaevjayate" (Mantra 3.1.6), which means "Only truth triumphs." All major religions, including Buddhism, Jainism, Hinduism, Christianity, strongly support non-violence and honesty in both work-life and personal life. "Sense of authenticity reduced depressive-anxious feelings, physical response and cognition-thought of helplessness" (Ito & Kodama, 2005, p. 24).

Sense of community

A Sense of community is "a feeling that members have of belongingness, a feeling that members matter to each other as well as to the group, and a shared belief that members' needs will be sought after through their commitment to being together" (McMillan & Chavis, 1986). The realization of interdependence and interconnectedness amongst employees helps in the development of a sense of community at the workplace (Jurkiewicz & Giacalone, 2004). A company may nurture community feelings among employees through developing a shared vision, mission, and core organizational values (Naylor et al., 1996). At the workplace, it is expressed in teamwork, collaborative problem solving, group task commitment (Pandey et al., 2016).

Karma capital

Karma theory is a unique feature of Indian spiritual values. The importance of karma capital could be realized from the fact that Indian, one of the most revered

scripture, Shrimad Bhagwad Gita' devoted one complete chapter in the elaboration of the concept of karma capital. It supports a hundred percent of work efforts without any expectation of results (Garg, 2017b). An action (Karma) should be free from any impression of pleasure–sorrow, profit – loss, and win-loss (elicited from Shrimad Bhagwad Gita as in Garg & Saxena, 2020). "All actions that are done have the power to ordain for their doer's joy or sorrow in future depending upon the action whether it is good or bad" (A. Rastogi & Pati, 2014, p. 26). An expectation-free and selfless devotion to work are called as 'Nishkam Karma,' which is hailed for peace, enrichment, and satisfaction (Singh & Chopra, 2018).

Gratitude

Several psychologists like Algoe (2012), Fredrickson (2004), Watkins (2014), Garg (2020), and Garg and Gera (2019) have supported gratitude-based organizational culture for the efficient functioning of employees. And Indian spiritualism is considered to represent the pinnacle of gratuitous behaviour (Garg, 2017c). Indians respect and pray for every living and non-living aspect of life. On the one hand, non-living items like water, sun, mountain, fire, etc., are worship for their contributions to humanity. And on the other hand, living creatures like snakes, lions, cows, birds are considered equivalent to God (Garg, 2020).

Workplace toxicity

Workplace incivility

Andersson and Pearson (1999) defined workplace incivility as "low-intensity deviant behavior with ambiguous intent to harm the target, in violation of workplace norms for mutual respect" (p. 457). Schilpzand et al. (2016) stated three categories of incivility- experienced (employees who are a victim of workplace incivility), witnessed (employees who bystanders incivility occurs), and instigated (employees who direct workplace incivility). Instigated and witnessed incivility is difficult to measures, and hence the present focuses on the experienced type of workplace incivility. Although the majority of studies on incivility are confined to USA samples, the phenomenon is not confined to one geographic area. It has become universal and has negative consequences across the world (Schilpzand et al., 2016). Workplace incivility has several detrimental effects on employees, organizations, and society as a whole.

Workplace incivility is linked to absenteeism (Sliter et al., 2012), turnover intentions (Cortina & Magley, 2009), arriving consistently late to work (Sliter et al., 2012), job dissatisfaction (Cortina & Magley, 2009, reduce work effort, loss of office hours by talking with coworkers about the incivility one experiences (Cortina et al., 2017), distress and self-doubt (Kern & Grandey, 2009), inappropriate teacher-student learning (McClure & Titus, 2018), higher cost of production/service (Landry & Neubauer, 2016), and financial losses (Jawahar & Schreurs, 2018).

Workplace ostracism

Workplace Ostracism is defined as"exclusion, rejection, or ignoring of an individual (or group) by another individual (or group) that hinders one's ability to establish or maintain positive interpersonal relationships, work-related success, or favorable reputation within one's place of work" (Hitlan et al., 2006, p. 217). It is also defined as "an individual's perception of being ignored or excluded" (D. L. Ferris et al., 2008, p. 1348). Several constructs of workplace ostracism have been defined by previous researchers (Robinson et al., 2012). Linguistic ostracism indicates a typical situation in which a group of employees speaks in a language/dialect that ostracized employees cannot understand (Dotan-Eliaz et al., 2009). This type of ostracism is often practiced by a group belonging to the same nationality, region, or race. Physical ostracism is leaving the place on the arrival of an isolated employee or not allowing employees to enter a particular place. Social ostracism is the denial of entry of an employee into a group, alliance, or relationship. And organizational shunning is "the systematic exclusion of a person who was once an included member of the group" (Anderson, 2009, p. 36).

These different types of workplace ostracism have a damaging impact on both employees and organizations. Ostracism is reported to be linked with emotional exhaustion (Liu & Xia, 2016), mental health (Heaphy & Dutton, 2008), lower job satisfaction (O'Reilly et al., 2015), reduced workplace citizenship, poor performance (Wu et al., 2011), sleep quality (Chen et al., 2020), weakened social ties (Yang & Treadway, 2018), psychological detachment (Chen et al., 2020), poor wellbeing (Sonnentag & Fritz, 2015), and unacceptable in-role task behaviour (D. L. Ferris et al., 2015). These negative consequences of incivility and ostracism at the workplace necessitates exploration of an effective remedy for the same. And, in this pursuit, the present study proposes that workplace spirituality could combat toxicity at the workplace.

Workplace spirituality and toxicity at the workplace

The proposed relationship between toxicity and workplace spirituality is theoretically premised in 'Conservation of Resources Theory (Hobfoll, 1989), 'Job Demands-Resources Theory' (Bakker & Demerouti, 2007), 'Broaden-and-Build Theory of Positive Emotion' (Fredrickson, 2004), and the 'Effort-Recovery Model' (Meijman & Mulder, 1998). These theories suggest that workplace spirituality negates one's intention to indulge in toxic behaviour, and it also enhances resources for victims to cope with workplace toxicity. Based on the broaden-and-build theory, we proposed that workplace spirituality-induced positive emotions help one in taming toxicity. Different dimensions of workplace spirituality like gratitude and a sense of community may develop cognitive, psychological, and physical resources for an employee. And these resources help in combating workplace incivility and ostracism. Workplace spirituality develops one's psychological capital in terms of hope, optimism, self-efficacy, and resilience (Sarkar & Garg, 2020). Gratitude facilitates hope, resilience, and optimism (Witvliet et al., 2018). A sense of community builds social resources like social support, empathy, mutual learning, and growth, and social approval and affiliation (Pandey et al., 2019; Saxena et al., 2020). A sense of community and transcendence encourages constructively deviant work behaviours like organizational citizenship behaviour, pro-social work behaviour, raising voice and dissents, etc. (Garg & Saxena, 2020). And these positively deviant work behaviours promote creativity, team spirit, and healthy competition among employees. It also encourages the employee to bring issues and challenges of workplace toxicity in the open, which leads to their resolution.

In addition, a high level of job demands and lack of acquaintance with coping strategies are two prominent reasons for workplace incivility and ostracism. According to the conservation of resources theory, toxicity leads to high environmental demands, which dissipate valuable resources of employees for coping (Sguera et al., 2016). These resources include physical stamina, mental energy, cognitive focus, emotional stability, peace of mind, and spiritual growth (Sguera et al., 2016). Again, several constructs of workplace spirituality like swadharma, transcendence, and karma capital may augment employees' contextual resources (social support, autonomy, feedback, and development opportunities) and personal resources (psychological, affective), physical resources (mental and physical health) and intellectual resources (skills, knowledge, experience; Garg, 2020; Pandey et al., 2009; Sarkar & Garg, 2020). Scholars have reported that autonomy, feedback, skill, and knowledge variety are an integral aspect of swadharma's dimension of a meaningful life (Pandey et al., 2016, p. 2019; Garg, 2017a). Banking on the Person-organization-fit approach, Saxena et al. (2020) argued that swadharma and transcendence lead to a better fit of employees in organizational settings. Also, several researchers, including Singh and Chopra (2018), advocated that the adherence to Karma philosophy helps in ensuring long-lasting peace, inner mind enrichment, and job satisfaction. And all these resources are poised to develop several coping strategies for employees.

And finally, the effort-recovery model explains how workplace spirituality could help a victim of workplace toxicity. The theory states that an individual needs time and effort to recover. Deci and Ryan's self-determination theory (2000) states that a human being has three innate psychological needs- relatedness, autonomy, and competence. And, institutionalization of workplace spirituality could help in the daily satisfaction of these psychological needs. A sense of community and transcendence leads to the development of a feeling of relatedness (Garg, 2017c; Pandey et al., 2019). Swadharma ensures autonomy. And karma capital and gratitude aspire one to develop skills and competencies to perform an assigned task with a hundred percentage work efficiency and also to revert to the benefactor. Based on these theoretical foundations, the following hypotheses and models (Figure 1) are proposed.

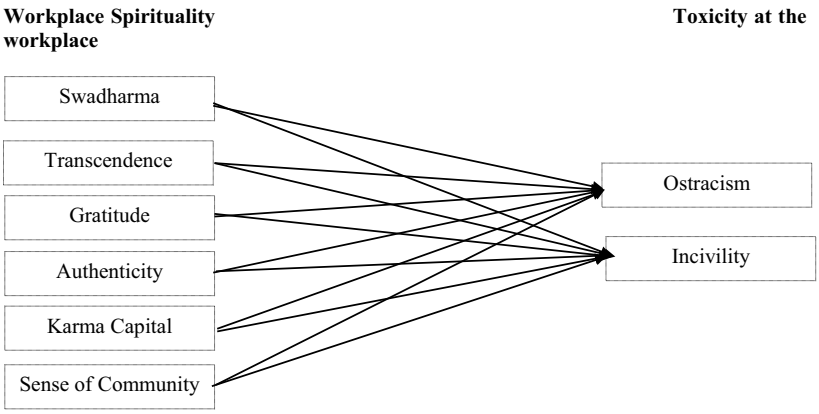


Figure 1. Proposed Model of Workplace Spirituality and Toxicity at the workplace. Workplace Spirituality Toxicity at the workplace

H1: Swadharma is negatively associated with workplace ostracism.

H2: Transcendence is negatively associated with workplace ostracism.

H3: Gratitude is negatively associated with workplace ostracism.

H4: Authenticity is negatively associated with workplace ostracism.

H5: Karma Capital is negatively associated with workplace ostracism.

H6: A sense of community is negatively associated with workplace ostracism.

H7: Swadharma is negatively associated with workplace incivility.

H8: Transcendence is negatively associated with workplace incivility.

H9: Gratitude is negatively associated with workplace incivility.

H10: Authenticity is negatively associated with workplace incivility.

H11: Karma Capital is negatively associated with workplace incivility.

H12: A sense of community is negatively associated with workplace incivility.

Research methodology

Objective and sample design

The primary objective of this study was to explore the effect of workplace spirituality on two dimensions of toxic work environments- incivility and ostracism. Six dimensions of workplace spirituality, namely Swadharma, Lokasangraha (Transcendence), Authenticity, Sense of community, Karma capital, and Gratitude, are explored here. The type and degree of workplace toxicity vary from one industry to another, so any investigation of incivility and ostracism has to be industry-specific. The current study selected the Indian oil and gas industry for examining the linkage between workplace spirituality and workplace toxicity. Besides, scholars like Saxena et al. (2020) highlighted different levels of workplace spirituality in public and private sector companies. Thus, the present study targeted only public sector oil and gas companies. The sample comprised 152 employees engaged in public sector oil and gas companies in India. The authors approached two hundred respondents, out of which 161 responded. Nine responses were discarded for being incomplete. Data was collected based on convenience sampling by using a structured questionnaire. Convenience sampling is often criticized for being biased. This bias has been reduced by integrating the non-probability sampling method with probability sampling. Thus, it was planned to collect responses from employees of different age groups, experiences, gender, and place of posting (onsite or off-site). It allowed the realization of the benefits of stratified sampling too. Respondents were approached either personally or through email. They were assured regarding the anonymity of responses and strictly academic usage of collected data. The sample characteristics are discussed in Table 1.

Measures

Workplace spirituality

It was quantified using a composite 23-items scale partially developed by Pandey et al. (2009) and Garg (2017b). The statements for the four "dimensions of Indian workplace spirituality (swadharma, authenticity, lokasangraha (Transcendence), sense of community) were taken from the scale developed by "Pandey et al. (2009). The other two dimensions (karma capital and gratitude) were taken from Garg (2017b). These items were rated on a five-point rating scale ranging from 1 (strongly disagree) to 5 (strongly agree). It included items like "People here try to avoid wastage of any kind," "People here are concerned about the natural environment while working here"), "when stuck with a problem, people here feel free to ask for (advice from colleagues/ advice from a superior/help from their colleagues." This consolidated scale has been used previously by Saxena et al. (2020), Garg (2020), and Garg et al. (2019).

Ostracism

It was measured using a 10-items scale developed by "D. L. Ferris et al. (2008). These items were rated on a seven-point Likert rating scale ranging from never (1) to always (7). The scale assesses the "actual experiences of specific behavior rather than general perspectives or imaginary reactions to the hypothetical scenarios. In included items like "Others ignored you at work," "Others left the area when you entered," and "Your greetings have gone unanswered at work."

Table 1. Description of Sample.

Variable	Item	Frequency	Percent (%)
Gender	Female	26	17.1
	Male	126	82.9
Age	Below 25 years	6	3.9
	26–35 years	52	34.2
	36–35 years	53	34.9
	46 years & above	41	27
Marital Status	Single	11	7.2
	Married	140	92.1
	Divorced	1	7
Education	Intermediate	3	2
	Under-Graduate	74	48.7
	Post-Graduate	75	49.3
Designation	Manager	82	53.9
	Executive	10	6.6
	Engineer	56	36.8
	Other Staff	4	2.6
Work Experience	Less than 1 year	4	2.6
	1–5 years	4	2.6
	5–10 years	34	22.4
	10–15 years	50	32.9
	More than 15 years	60	39.5
Posting	On-site	64	42.1
	Off-site	88	57.9

Source: Primary Data

Incivility

Workplace Incivility Scale (WIS; Cortina et al., 2001) was used for measuring the frequency of experiences of uncivil activities by respondents. Participants responded on a five-point Likert scale, starting from never (1), once in a while (2), sometimes (3), often (4), and many times (5), describing how frequently they had encountered each of this behavior from a supervisor or coworker during the past year at work. The scale tried to assess the actual experiences of specific behavior rather than general perspectives or imaginary reactions to the hypothetical scenarios. It included items like "Put you down or was condescending to you in some way," "Paid little attention to a statement you made or showed little interest in your opinion."

Reliability, validity, multicollinearity, and common method bias

Internal consistency reliability is calculated with the help of Cronbach's alpha. Researchers agree that the value of Cronbach's alpha measures should be greater than 0.70 to confirm reliability (George & Mallery, 2003). Table 2 states that the values of Cronbach's alpha for all six dimensions of workplace spirituality, workplace ostracism, and workplace incivility are greater than 0.7, which establishes good reliability. Also, convergent validity is accessed using Average Variance Extracted (AVE), which must greater than 0.50 to substantiate validity (Fornell & Larcker, 1981). Again, Table 2 reveals that AVE values for all variables of the study are greater than the desired value of 0.50, which confirms the validity of the data. Since workplace spirituality is being measured with the help of six dimensions, there are chances of emergence of the issue of multicollinearity. It has been investigated with the help of the Variance Inflation Factor (VIF values. According to Kline (2005), VIF is to be less than 5 to conclude the absence of multicollinearity. The values of VIF for swadharma, sense of community, authenticity, gratitude, karma capital, and transcendence are less than 5, which confirms the absence of multicollinearity, i.e., the independent variables are not related to each other.

Podsakoff et al. (2012) suggested that common method variance may arise when data is collected at the same time from the same sources. It could be minimized using both procedural and statistical remedies. Firstly, pre-validated measurement scales were used to measure workplace spirituality, ostracism, and incivility. Secondly, all the respondents were guaranteed confidentiality and also anonymity of the data. These small initiatives encourage accurate and positive responses. Apart from that, Harman's one factor was also used to check common method bias. When all the study variables were subjected to principal component analysis, no single factor was found to explain the majority of the variance. Hence, the probability of common method biases is neglected in the study.

Data analysis and results

Table 2 reports mean values, standard deviation, reliability, validity, and multicollinearity estimates of all the study variables. The two most experienced dimensions of workplace spirituality are gratitude (mean = 4.342) and karma capital (mean = 4.056). Also, the mean value for ostracism (mean = 1.686) and incivility (mean = 1.721). It is also observed that all values of Cronbach's alpha, Average Variance Explained, and Variance Inflation Index are well within prescribed

Table 2. Descriptive Statistics, Reliability, Validity and Multicollinearity Estimates.

Variables	Mean	S.D.	A	AVE	VIF
Swadharma	3.868	0.758	0.872	0.639	2.750
Authenticity	3.191	0.846	0.753	0.584	1.667
Transcendence	3.715	0.713	0.796	0.648	4.703
Karma Capital	4.056	0.642	0.793	0.527	3.074
Gratitude	4.342	0.584	0.703	0.550	2.158
Sense of Community	3.755	0.736	0.704	0.650	1.089
Workplace Spirituality	3.822	0.522	– – –	– –	– –
Workplace Ostracism	1.686	0.954	0.932	0.693	– – -
Workplace Incivility	1.721	0.736	0.940	0.534	– –

Source: Primary Data, SD- Standard Deviation, α- Cronbach's alpha, AVE- Average Variance Explained, VIF- Variance Inflation Index

ranges. And thus, they confirm the reliability and validity of scale and absence of multicollinearity. Table 3 discusses the correlation between different variables of the study. It depicts that correlation coefficients between six dimensions of workplace spirituality and ostracism and incivility are statistically significant with a negative sign. In a nutshell, it is concluded that workplace spirituality is negatively correlated with toxicity at the workplace.

Table 4 and Table 5 represent the results of multiple regression that considers six dimensions of workplace spirituality as independent variables and ostracism as a dependent variable. The findings report that all dimensions of workplace spirituality are negatively regressed with ostracism at the workplace (β(Swadharma) = −0.460, β(Transcendence) = −0.175, β(authenticity) = −0.108, β (Gratitude) = −0.032, β(karma capital) = −0.201, and β (sense of community) = −0.120). However, the regression coefficient is not statistically significant for a sense of community. It states that Hypotheses 1 to 5 are accepted. It also states that workplace spirituality could explain 59% variations in ostracism behaviour at the workplace. The following regression equation is derived from the findings of the study.

Ostracism = 1.412– 0.460 Swadharma – 0.108 Transcendence – 0.175 Authenticity – 0.032 Gratitude – 0.201 Karma Capital – 0.120 Sense of community

Table 6 and Table 7 show regression coefficients of the regression line that takessix dimensions of workplace spirituality as independent variables and incivility as a dependent variable. Here, it is being observed that all constructs of workplace spirituality are negatively regressed on incivility, and all regression coefficients are statistically significant at 0.05. Various regression coefficients are β (Swadharma) = −0.207, β(Transcendence) = −0.106, β (authenticity) = −0.175, β(Gratitude) = −0.056, β(karma capital) = −0.229, and β(sense of community) = −0.034. Also, these six dimensions of workplace spirituality could explain 70.9% variations in incivility at the workplace. The findings suggest the acceptance of hypotheses 7 to 12.

Discussion and conclusion

The undesirable consequences of workplace toxicity have attracted the attention of all stakeholders, including practitioners, academicians, psychologists, and policymakers. The present study appropriately tries to capture the imagination of researchers towards the role of workplace spirituality in combating toxicity at work. The study provides empirical evidence of a negativerelationship between six Indian dimensions of workplace spirituality and two manifestations of workplace toxicity, i.e., incivility and ostracism in the workplace. These results support the proposed hypotheses of the study. The findings of the study support the results of a few previous studies of the field. Lata and Chaudhary (2019) proposed a conceptual framework linking workplace spirituality with incivility. Sarkar and Garg (2020) reported a positive association between workplace spirituality and nonviolent work behaviour. DeWall et al. (2011) suggested that grateful moods are linked with less aggressive behaviour. Zhang (2018) concluded the moderating effect of workplace spirituality in the relationship between workplace ostracism and unethical prosocial behaviour.

The study contributes to the existing literature in three possible ways. First, the present study proposed an interesting and novel way to deal with toxicity in the workplace. This is one of the few studies that explored the role of workplace spirituality in taming incivility and ostracism. Moreover, the findings are based on the Indian conceptualization of spirituality at work. Swadharma, which signifies meaningful and meditative work assignments (Pandey et al., 2019), has already been practiced with the help of Hackman and Oldham's Job characteristics model. The model advocates that a job's motivation potential score depends on autonomy, feedback, task significance, and skill variety. And previous researchers have reported that these job characteristics restrict indulge in toxic work behaviour. A sense of community and transcendence make employees work in a close group, thereby breaking the shackles of the physical, verbal, or social form of workplace ostracism.

Table 3. Correlation Matrix".

Variables	S	A	T	KC	G	SC	WO	WI
S	1							
A	0.513**	1						
T	0.827**	0.593**	1					
KC	0.687**	0.482**	0.663**	1				
G	0.528**	0.288**	0.439**	0.586**	1			
SC	0.046**	0.139**	0.128*	0.188*	0.141*	1		
WO	−0.015**	−0.218**	−0.052**	−0.111**	−0.068**	0.089	1	
WI	−0.066**	−0.295**	−0.077**	−0.223**	−0.096**	−0.010**	0.555**	1

Source: Primary Data, ** Significant at 0.001, *Significant at 0.005, S- Swadharma, A- Authenticity,T- Transcedence KC- Karma Capital, G-Gratitude, SC- Sense of Community, WO- Workplace Ostracism,WI- Workplace Incivility

Table 4. Result of Multiple Regression (Dependent Variable- Workplace Ostracism).

	"Unstandardized Coefficients"		"Standardized Coefficients"		
Variables	B	Std. Error	Beta	t-value	Sig.
(Constant)	1.412	0.674	-	−1.872	0.032**
Swadharma	−0.281	0.227	−0.460	−1.238	0.011**
Authenticity	−0.197	0.117	−0.175	−1.690	0.043**
Transcendence	−0.144	0.233	−0.108	−0.621	0.036**
Karma Capital	−0.300	0.209	−0.201	−1.434	0.024**
Gratitude	−0.052	0.192	−0.032	−0.269	0.008**
Sense of Community	0.155	0.108	−0.120	−0.430	0.155

Source: Primary Data, "**Significant at 0.05 level (2-tailed)"

Table 5. Model Summary (Dependent Variable- Workplace Ostracism).

R	R^2	"Adjusted R^2"	"Std. Error"	F	Sig.
− 0.768	0.590	0.573	0.539	15.755	0.038**

Source: Primary Data, "**Significant at 0.05 level (2-tailed)"

Table 6. Result of Multiple Regression (Dependent Variable- Workplace Incivility).

	"Unstandardized Coefficients"		"Standardized Coefficients"		
Variables	B	Std. Error	Beta	t-value	Sig.
(Constant)	2.209	0.517	-	4.271	0.000**
Swadharma	−0.202	0.174	−0.207	−1.161	0.028**
Authenticity	−0.248	0.090	−0.285	−2.773	0.006**
Transcendence	−0.109	0.178	−0.106	−0.614	0.040**
Karma Capital	−0.262	0.160	−0.229	−1.636	0.014**
Gratitude	−0.071	0.147	−0.056	−0.480	0.032**
Sense of Community	−0.034	0.083	−0.034	−0.407	0.045**

Source: Primary Data, "**Significant at 0.05 level (2-tailed)"

Table 7. Model Summary (Dependent Variable- Workplace Incivility).

R	R^2	"Adjusted R^2"	"Std. Error"	F	Sig.
−0. 842	0.709	0.696	0.611	24.089	0.043**

Source: Primary Data, "**Significant at 0.05 level (2-tailed)"

Also, karma capital advocates complete work devotion without fear of uncertain results (Garg et al., 2019). And, lack of outcome-related anxiety takes away mistrust, suspicion, and work distress, and thereby, it counters toxic elements of work behaviour.

Second, this current study contributes an alternate perspective to workplace toxicity. It enriches the literature on workplace toxicity by suggesting a constructive and positive antidote to incivility and ostracism. Traditionally, undesirable work behaviours used to be quelled using restrictive practices like disciplinary proceedings, punishments, suspension, etc. These restrictive interventions were often criticized for being repressive and damaging. This study offers a more acceptable and beneficial way to deal with toxicity. Workplace spirituality is poised to provide multiple benefits to the modern organization. Third, the study tends to expand the horizons of workplace spirituality. The field of spirituality at work is often criticized for being inadequately researched (Saxena et al., 2020). Several researchers, including Pandey et al. (2019), Saxena et al. (2020), advocated extensive researches in the domain of workplace spirituality for the extension, exploration, and development of an adequate theoretical knowledge base and practical intervention for employee wellbeing. Spirituality's "inclusion in established research streams is long overdue" (Sheep, 2006, p. 371). In this background, the present study provides an appropriate description of the organizational values of spirituality. And thus, it contributes to the growing literature of workplace spirituality.

The rise in the instances of workplace toxicity has alarmed HR managers and researchers around the globe. Significant financial implications and human rights concerns have encouraged OD experts to a series of preventive measures. Several practitioners have advocated a stringent recruitment process for new employees to prevent potentially toxic people from entering the organization (see, Howard, 2002). Extensive use of psychological personality tests, text-analysis of resumes, reference checking, and verification of candidates' social media platforms could allow employers to look for dark personality traits among job candidates. Here, the study's findings suggest the inclusion of personality tests that explore candidates' spiritual intelligence and quotient. Specifically, the recruiters may examine gratitude traits, inclination towards meaningful work, or likeness to excel in a team assignment. These spiritual dimensions of personality will make employees capable of coping with workplace toxicity.

Few organizations adopt a 'zero-tolerance policy towards toxic work behaviour with strict adherence to organizational policies, and grievance, and disciplinary process (see, Flynn et al., 2008). Here, it is being suggested to institutionalize a comprehensive spiritual organizational culture. Such a spirituality-based organizational climate may be developed through internalizing spiritual values in the company's vision, mission, and policies (Saxena et al., 2020). Several renowned MNCs like Apple, Microsoft, Amazon, Google, etc., are already practicing specific spiritual interventions in their respective organizations. Similarly, based on the findings of the present study, few industry-relevant spiritual practices may be devised for the Indian oil and gas industry. It is suggested that both culprit and victim of workplace toxicity may be posted in Corporate Social Responsibility (CSR) department or in cause-related

marketing job profile to cherish and realize the tenants of transcendence. Several companies have started to send their employees on paid leaves to serve the oppressed strata of the society. Such practices have helped several organizations in the detoxification of their employees. Also, employees of oil and gas companies may be subjected to training in yoga, meditation, pranayama (breathing exercise), vipassana, self-introspection, mindfulness, etc. to help them understand and practice the essence of karma theory. All these interventions require the active support of top management and leaders. For inference, gratuitous behaviours are contagious, and that of managers has a deep and profound impact on employees (Fredrickson, 2004; Garg et al., 2020). So, gratuitous acts of leaders and managers will undoubtedly inspire other employees to follow suit.

Other organizations rely on the 'Employee assistance program' and 'Stress management plan' for detoxification of their workforce. It is being suggested that the effectiveness of any such program may be enhanced by adding certain spiritual dimensions to it. For inference, a stress management program may add some gratitude exercise. These gratitude interventions develop trait gratitude among practitioners. Some of the gratitude interventions are gratitude journals, gratitude letters and subsequent visits, meditation, and self-introspection, blessing-count exercise, mindfulness, opposition thinking, the act of kindness and compassion, and sharing positive and favourable events of one's life. Similarly, the importance of a sense of community could be realized with the help of group-training techniques like T-training, role-playing, group assignment, simulations.

Although the study provides some valuable insights into the association of workplace spirituality and toxicity at work, it is marred with certain limitations too. First, the findings of the study are based on a cross-sectional research design. Also, the entire data (independent and dependent variables) has been collected from the same source simultaneously. This may lead to an issue of common method variance. Also, both procedural and statistical remedies have been applied to minimize the common method bias. But, it cannot be overruled entirely. Future researchers are suggested to utilize repeated cross-sectional or longitudinal research design to deal with bias. Double source data could also be used for better results. Further, an experimental exploration in a laboratory or real-life setting could be another investigation tool for the future scholar. It is also observed that the sample comprises 83% of male respondents; thus, there might be some issue of the representative sample. However, McKinsey's 'Women in the Workplace report-2018′ highlighted that female employees comprise 15% of the total workforce of Indian oil and gas companies. Thus, the sample of the current study may be regarded as representative. Also, the present study was conducted in Indian settings. And therefore, empirical investigations in different national, social, and cultural backgrounds may yield different results. Future researchers are suggested to cross-verify the model of the present study in different cultural and national settings. Future research could also try to examine the potential mediating effect of psychological capital/resources, cognitive and social resources in the relationship of workplace spirituality and workplace toxicity.

In conclusion, the study proposed that workplace spirituality could help an organization in taming toxicity in the workplace. Using cross-sectional research design, the study establishes that Indian dimensions of workplace spirituality- swadharma, sense of community, authenticity, karma capital, gratitude, and transcendence- are negatively associated with workplace incivility and ostracism.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The author(s) reported there is no funding associated with the work featured in this article.

ORCID

Naval Garg http://orcid.org/0000-0001-6694-9663

References

Algoe, S. B. (2012). Find, remind, and bind: the functions of gratitude in everyday relationships. *Social and Personality Psychology Compass*, *6*(6), 455–469. https://doi.org/10.1111/j.1751-9004.2012.00439.x

Anderson, J. W. (2009). Organizational shunning: the disciplinary functions of "non-sense.". *Atlantic Journal of Communication*, *17*(1), 36–50. https://doi.org/10.1080/15456870802506140

Andersson, L. M., & Pearson, C. M. (1999). Tit for tat? the spiraling effect of incivility in the workplace. *Academy of Management Review*, *24*(3), 452–471. https://doi.org/10.2307/259136

Ashar, H., & Lane-Maher, M. (2004). Success and spirituality in the new business paradigm. *Journal of Management Inquiry*, *13*(3), 249–260. https://doi.org/10.1177/1056492604268218

Ashmos, D. P., & Duchon, D. (2000). Spirituality at work: A conceptualization and measure. *Journal of Management Inquiry*, *9*(2), 134–145. George, R. G., Sorenson, G. J., Bums, J. M. (2004). Encyclopedia of leadership (Vol. 4). London, UK: SAGE. https://doi.org/10.1177/105649260092008

Badrinarayanan, V., & Madhavaram, S. (2008). Workplace spirituality and the selling organization: A conceptual framework and research propositions. *Journal of Personal Selling & Sales Management*, *28*(4), 421–434. https://doi.org/10.2753/PSS0885-3134280406

Bakker, A. B., & Demerouti, E. (2007). The job demands-resources model: State of the art. *Journal of Managerial Psychology*, *22*(3), 309–328. https://doi.org/10.1108/02683940710733115

Bhawuk, D. P. S. (2011). *Spirituality and Indian psychology*. Springer-Verlag.

Chen, T., Dai, Z., Mo, P., Li, X., Ma, Z., Song, S., Chen, X., Luo, M., Liang, K., Gao, S., Zhang, Y., Deng, L., & Xiong, Y., et al. (2020). Clinical characteristics and outcomes of older patients with coronavirus disease 2019 (COVID-19) in Wuhan, China: A single-centered, retrospective study. *The Journals of Gerontology: Series A*, *75*(9), 1788–1795. https://doi.org/10.1093/gerona/glaa089

Cheung, T., & Yip, P. (2015). Depression, anxiety and symptoms of stress among Hong Kong nurses: A cross-sectional study. *International Journal of Environmental Research and Public Health*, *12*(9), 11072–11100. https://doi.org/10.3390/ijerph120911072

Cortina, L. M., Kabat-Farr, D., Leskinen, E. A., Huerta, M., & Magley, V. J., et al. (2013). Selective incivility as modern discrimination in organizations. *Journal of Management*, *39*(6), 1579–1605. https://doi.org/10.1177/0149206311418835

Cortina, L. M., Kabat-Farr, D., Magley, V. J., & Nelson, K. (2017). Researching rudeness: The past, present, and future of the science of incivility. *Journal of Occupational Health Psychology*, *22*(3), 299–313. https://doi.org/10.1037/ocp0000089

Cortina, L. M., Magley, V. J., Williams, J. H., & Langhout, R. D., et al. (2001). Incivility in the workplace: incidence and impact. *Journal of Occupational Health Psychology*, *6*(1), 64–80. https://doi.org/10.1037/1076-8998.6.1.64

Cortina, L. M., & Magley, V. J. (2009). Patterns and profiles of response to incivility in the workplace. *Journal of Occupational Health Psychology*, *14*(3), 272–288. https://doi.org/10.1037/a0014934

Daniels, K., Watson, D., & Gedikli, C., et al. (2017). Well-being and the social environment of work: A systematic review of intervention studies. *International Journal of Environmental Research and Public Health*, *14*(8), 918. https://doi.org/10.3390/ijerph14080918

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, *11*(4), 227–268. https://doi.org/10.1207/s15327965pli1104_01

DeWall, C. N., Lambert, N. M., Pond, R. S., Kashdan, T. B., & Fincham, F. D., et al. (2011). A grateful heart is a nonviolent heart. *Social Psychological and Personality Science*, *3*(2), 232–240. https://doi.org/10.1177/1948550611416675

Dotan-Eliaz, O., Sommer, K. L., & Rubin, Y. S., et al. (2009). Multilingual groups: effects of linguistic ostracism on felt rejection and anger, coworker attraction, perceived team potency, and creative performance. *Basic and Applied Social Psychology*, *31*(4), 363–375. https://doi.org/10.1080/01973530903317177

Estes, B., & Wang, J. (2008). Integrative literature review: Workplace incivility: Impacts on individual and organizational performance. *Human Resource Development Review*, *7*(2), 218–240. https://doi.org/10.1177/1534484308315565

Ferris, D. L., Brown, D. J., Berry, J. W., & Lian, H., et al. (2008). The development and validation of the workplace ostracism scale. *Journal of Applied Psychology*, *93*(6), 1348–1366. https://doi.org/10.1037/a0012743

Ferris, D. L., Lian, H., Brown, D. J., & Morrison, R., et al. (2015). Ostracism, self-esteem, and job performance: when do we self-verify and when do we self-enhance? *Academy of Management Journal*, *58*(1), 279–297. https://doi.org/10.5465/amj.2011.0347

Flynn, J., Valikoski, T. R., & Grau, J., et al. (2008). Listening in the business context: reviewing the state of research. *International Journal of Listening*, *22*(2), 141–151. https://doi.org/10.1080/10904010802174800

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*(1), 39–50. https://doi.org/10.1177/002224378101800104

Foster, S., & Foster, A. (2019). The impact of workplace spirituality on work-based learners. *Journal of Work-Applied Management*, *11*(1), 63–75. https://doi.org/10.1108/JWAM-06-2019-0015

Fredrickson, B. L. (2004). Gratitude, like other positive emotions, broadens & builds. In R. A. Emmons & M. E. McCullough (Eds.), *The Psychology of Gratitude* (pp. 145–166). Oxford University Press.

Garg, N., & Gera, S. (2019). Gratitude and leadership in higher education institutions. *Journal of Applied Research in Higher Education*, *12*(5), 915–926. https://doi.org/10.1108/jarhe-09-2019-0241

Garg, N., Jain, A., & Punia, B. (2020). Gratitude, social intelligence, and leadership among university teachers: Mediation and moderation analysis. *International Journal of Organizational Analysis*, *29*(2), 368–388. https://doi.org/10.1108/ijoa-01-2020-2016

Garg, N., Punia, B. K., & Jain, A., et al. (2019). Workplace spirituality and job satisfaction: exploring mediating effect of organization citizenship behaviour. *Vision: The Journal of Business Perspective*, *23*(3), 287–296. https://doi.org/10.1177/0972262919850928

Garg, N., & Saxena, A. (2020). Analyzing the inter-relation between workplace spirituality and constructive deviance. *Asian Journal of Business Ethics*, *9*(1), 121–141. https://doi.org/10.1007/s13520-020-00102-x

Garg, N. (2017a). Workplace spirituality and employee well-being: an empirical exploration. *Journal of Human Values*, *23*(2), 129–147. https://doi.org/10.1177/0971685816689741

Garg, N. (2017b). Workplace spirituality and organizational commitment: an empirical exploration. *Purushartha - A Journal of Management, Ethics and Spirituality*, *10*(2), 8–24. https://doi.org/10.21844/pajmes.v10i02.10566

Garg, N. (2017c). Workplace spirituality and organizational performance in indian context: mediating effect of organizational commitment, Work motivation and employee engagement. *South Asian Journal of Human Resources Management*, *4*(2), 191–211. https://doi.org/10.1177/2322093717736134

Garg, N. (2020). Exploring role of gratitude in developing teacher leadership in Indian universities. *International Journal of Educational Management*, *34*(5), 881–901. https://doi.org/10.1108/ijem-07-2019-0253

George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference. 11.0 update* (4th ed). Allyn & Bacon.

George, R. G., Sorenson, G. J., & Bums, J. M. (2004). *Encyclopedia of leadership* (Vol. 4). London, UK: SAGE.

Giacalone, R. A., & Jurkiewicz, C. L. (Eds.). (2003). *Handbook of workplace spirituality and organizational performance* (pp. 108–122). M.E. Sharpe.

Harrington, W. J., Preziosi, R. C., & Gooden, D. J., et al. (2001). Perceptions of workplace spirituality among professionals and executives. *Employee Responsibilities and Rights Journal*, *13*(3), 155–163. https://doi.org/10.1023/A:1014966901725

Heaphy, E. D., & Dutton, J. E. (2008). Positive social interactions and the human body at work: Linking organizations and physiology. *Academy of Management Review*, *33*(1), 137–162. https://doi.org/10.5465/amr.2008.27749365

Herr, R., Barrech, A., Riedel, N., Gündel, H., Angerer, P., & Li, J., et al. (2018). Long-term effectiveness of stress management at work: effects of the changes in perceived stress reactivity on mental health and sleep problems seven years later. *International Journal of Environmental Research and Public Health*, *15*(2), 255. https://doi.org/10.3390/ijerph15020255

Hitlan, R. T., Cliffton, R. J., & DeSoto, M. C., et al. (2006). Perceived exclusion in the workplace: the moderating effects of gender on work-related attitudes and psychological health. *North American Journal of Psychology*, *21*(8), 7–236.

Hobfoll, S. E. (1989). Conservation of resources: A new attempt at conceptualizing stress. *American Psychologist*, *44*(3), 513–524. https://doi.org/10.1037/0003-066x.44.3.513

Howard, S. (2002). A spiritual perspective on learning in the workplace. *Journal of Managerial Psychology*, *17*(3), 230–242. https://doi.org/10.1108/02683940210423132

Ito, M., & Kodama, M. (2005). The relationships of sense of authenticity with stress-responses and stress-copings. *The Japanese Journal of Health Psychology*, *18*(1), 24–34. https://doi.org/10.11560/jahp.18.1_24

Jalan, S., & Garg, N. (2021). Exploring linkages between workplace spirituality and occupational stress among indian police personnel. *South Asian Journal of Human Resources Management*, 232209372110596. Advance online publication. https://doi.org/10.1177/23220937211059626

Jawahar, I. M., & Schreurs, B. (2018). Supervisor incivility and how it affects subordinates' performance: A matter of trust. *Personnel Review*, *47*(3), 709–726. https://doi.org/10.1108/PR-01-2017-0022

Jurkiewicz, C. L., & Giacalone, R. A. (2004). A values framework for measuring the impact of workplace spirituality on organizational performance. *Journal of Business Ethics*, *49*(2), 129–142. https://doi.org/10.1023/B:BUSI.0000015843.22195.b9

Kern, J. H., & Grandey, A. A. (2009). Customer incivility as a social stressor: the role of race and racial identity for service employees. *Journal of Occupational Health Psychology*, *14*(1), 46–57. https://doi.org/10.1037/a0012684

Kline, R. B. (2005). *Principles and practice o f structural equation modeling* (2nd ed). Guilford Press.

Landry, L., & Neubauer, D. (2016). The role of the government in providing access to higher education: the case of government-sponsored financial aid in the US. *Journal of Education and Work*, *29*(1), 64–76. https://doi.org/10.1080/13639080.2015.1049027

Lata, M., & Chaudhary, R. (2019). Workplace spirituality and incivility at work: A conceptual framework. *Academy of Management Proceedings*, *2019*(1), 18394. https://doi.org/10.5465/ambpp.2019.65

Lim, S., & Cortina, M. (2005). Interpersonal mistreatment in the workplace: The interface and impact of general incivility and sexual harassment. *Journal of Applied Psychology*, *90*(3), 483–496. https://doi.org/10.1037/0021-9010.90.3.483

Liu, H., & Xia, H. (2016). Workplace ostracism: A review and directions for future research. *Journal of Human Resource and Sustainability Studies*, *04*(3), 197–201. https://doi.org/10.4236/jhrss.2016.43022

Markow, F., & Klenke, K. (2005). THE EFFECTS OF PERSONAL MEANING AND CALLING ON ORGANIZATIONAL COMMITMENT: AN EMPIRICAL INVESTIGATION OF SPIRITUAL LEADERSHIP. *International Journal of Organizational Analysis*, *13*(1), 8–27. https://doi.org/10.1108/eb028995

McClure, K. R., & Titus, M. A. (2018). Spending up the ranks? the relationship between striving for prestige and administrative expenditures at U.S. Public Research Universities. *The Journal of Higher Education*, *89*(6), 961–987. https://doi.org/10.1080/00221546.2018.1449079

McCormick, D. W. (1994). Spirituality and management. *Journal of Managerial Psychology*, *9*(6), 5–8. https://doi.org/10.1108/02683949410070142

McMillan, D. W., & Chavis, D. M. (1986). Sense of community: A definition and theory. *Journal of Community Psychology*, *14*(1), 6–23. https://doi.org/10.1002/1520-6629(198601)14:1<6::AID-JCOP2290140103>3.0.CO;2-I

Meijman, T. F., & Mulder, G. (1998). Psychological aspects of workload. In P. J. D. Drenth, H. K. Thierry, & C. J. De Wolff (Eds.), *Handbook of work and organizational psychology* (2nd ed, pp. 5–33). Psychology Press/Erlbaum.

Naylor, T., Willimon, W., & Osterberg, R., et al. (1996). The search for community in the workplace. *Business and Society Review*, *97*(1), 42–47.

O'Reilly, J., Robinson, S. L., Berdahl, J. L., & Banki, S., et al. (2015). Is negative attention better than no attention? the comparative effects of ostracism and harassment at work. *Organization Science*, *26*(3), 774–793. https://doi.org/10.1287/orsc.2014.0900

Oxford Dictionary. (2015). London, UK: Oxford University Press.

Oyeleye, O., Hanson, P., O'Connor, N., & Dunn, D., et al. (2013). Relationship of workplace incivility, stress, and burnout on nurses' turnover intentions and psychological empowerment. *JONA: The Journal of Nursing Administration*, *43*(10), 536–542. https://doi.org/10.1097/NNA.0b013e3182a3e8c9

Pandey, A., Gupta, R. K., & Arora, A. P., et al. (2009). Spiritual climate of business organizations and its impact on customers' experience. *Journal of Business Ethics*, *88*(2), 313–332. https://doi.org/10.1007/s10551-008-9965-z

Pandey, A., Gupta, R. K., & Kumar, P., et al. (2016). Spiritual climate and its impact on learning in teams in business organizations. *Global Business Review*, *17*(3_suppl)), 159S–172S. https://doi.org/10.1177/0972150916631208

Pandey, A., Gupta, V., & Gupta, R. K., et al. (2019). Spirituality and innovative behaviour in teams: examining the mediating role of team learning. *IIMB Management Review*, *31*(2), 116–126. https://doi.org/10.1016/j.iimb.2019.03.013

Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P., et al. (2012). Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology*, *63*(1), 539–569. https://doi.org/10.1146/annurev-psych-120710-100452

Pouliakas, K., & Theodossiou, I. (2013). THE ECONOMICS OF HEALTH AND SAFETY AT WORK: AN INTERDICIPLINARY REVIEW OF THE THEORY AND POLICY. *Journal of Economic Surveys*, *27*(1), 167–208. https://doi.org/10.1111/j.1467-6419.2011.00699.x

Rastogi, A., & Pati, S. P. (2014). Towards a conceptualization of karma yoga. *SSRN Electronic Journal*, 21(1), 51–63. Published. https://doi.org/10.2139/ssrn.2522914

Rastogi, R., & Dave, V. (2004). Managerial effectiveness: A function of personality type and organizational components. *Singapore Management Review*, *26*(2), 79–87. https://doi.org/10.1177/0971685815579985

Robinson, S. L., O'Reilly, J., & Wang, W., et al. (2012). Invisible at work. *Journal of Management*, *39*(1), 203–231. https://doi.org/10.1177/0149206312466141

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*, *31*(5), 709–728. https://doi.org/10.1108/IJCMA-11-2019-0217

Saxena, A., Garg, N., Punia, B., & Prasad, A., et al. (2020). Exploring role of indian workplace spirituality in stress management: A study of oil and gas industry. *Journal of Organizational Change Management*, *33*(5), 779–803. https://doi.org/10.1108/JOCM-11-2019-0327

Schilpzand, P., de Pater, I. E., & Erez, A., et al. (2016). Workplace incivility: A review of the literature and agenda for future research. *Journal of Organizational Behavior*, *37*(4), S57–S88. https://doi.org/10.1002/job.1976

Sguera, F., Bagozzi, R. P., Huy, Q. N., Boss, R. W., & Boss, D. S., et al. (2016). Curtailing the harmful effects of workplace incivility: the role of structural demands and organization-provided resources. *Journal of Vocational Behavior*, *95–96*(3), 115–127. https://doi.org/10.1016/j.jvb.2016.08.004

Sharma, S. (1999). Corporate gita: lessons for management. administration and leadership. *Journal of Human Values*, *5*(2), 103–123. https://doi.org/10.1177/097168589900500203

Sheep, M. L. (2006). Nurturing the whole person: the ethics of workplace spirituality in a society of organizations. *Journal of Business Ethics*, *66*(4), 357–375. https://doi.org/10.1007/s10551-006-0014-5

Singh, J., & Chopra, V. G. (2018). Workplace spirituality, grit and work engagement. *Asia-Pacific Journal of Management Research and Innovation*, *14*(1–2), 50–59. https://doi.org/10.1177/2319510X18811776

Sliter, M., Sliter, K., & Jex, S., et al. (2012). The employee as a punching bag: the effect of multiple sources of incivility on employee withdrawal behavior and sales performance. *Journal of Organizational Behavior*, *33*(1), 121–139. https://doi.org/10.1037/a0029862

Sonnentag, S., & Fritz, C. (2015). Recovery from job stress: The stressor-detachment model as an integrative framework. *Journal of Organizational Behavior*, *36*(S1), S72–S103. https://doi.org/10.1002/job.1924

Tricahyadinata, I., Hendryadi, S., Zainurossalamia, Z. A., S., & Riadi, S. S., et al. (2020). Workplace incivility, work engagement, and turnover intentions: multi-group analysis. *Cogent Psychology*, *7*(1),1743627. https://doi.org/10.1080/23311908.2020.1743627

Van der Walt, F. (2018). Workplace spirituality, work engagement & thriving at work. *SA Journal of Industrial Psychology*, *44*(1), a1457. https://doi.org/10.4102/sajip.v44i0.1457

Watkins, P. C. (2014). *Gratitude and the good life: Toward a psychology of appreciation*. Springer Science & Business Media.

Witvliet, C. V., Richie, F. J., Root Luna, L. M., & van Tongeren, D. R., et al. (2018). Gratitude predicts hope and happiness: A two-study assessment of traits and states. *The Journal of Positive Psychology*, *14*(3), 271–282. https://doi.org/10.1080/17439760.2018.1424924

Wu, L., Wei, L., & Hui, C., et al. (2011). Dispositional antecedents and consequences of workplace ostracism: an empirical examination. *Frontiers of Business Research in China*, *5*(1), 23–44. http://dx.doi.org/10.1007/s11782-011-0119-2

Yang, J., & Treadway, D. C. (2018). A social influence interpretation of workplace ostracism and counterproductive work behavior. *Journal of Business Ethics*, *148*(4), 879–891. https://doi.org/10.1007/s10551-015-2912-x

Zhang, S. (2018). Workplace spirituality and unethical pro-organizational behavior: the mediating effect of job satisfaction. *Journal of Business Ethics*, *161*(3), 687–705. https://doi.org/10.1007/s10551-018-3966-3

Rendiconti Lincei. Scienze Fisiche e Naturali
https://doi.org/10.1007/s12210-021-01040-x

RESEARCH PAPER

Comparison of TiO_2 catalysis and Fenton's treatment for rapid degradation of Remazol Red Dye in textile industry effluent

Harsh Pipil[1] · Shivani Yadav[1] · Harshit Chawla[1] · Sonam Taneja[1] · Manisha Verma[1] · Nimisha Singla[1] · A. K. Haritash[1]

Received: 13 September 2021 / Accepted: 20 December 2021
© The Author(s), under exclusive licence to Accademia Nazionale dei Lincei 2021

Abstract
The contamination of water bodies by toxic industrial effluents is a serious threat to environment and the exposed organisms. The treatment of carcinogenic azo dyes in wastewater of grossly polluting textile industry is a major challenge considering the persistent nature of chemical dyes against biological treatment. The present study explores efficacy of advanced oxidation processes—photocatalysis and photo-Fenton, towards degradation of Remazol Red dye in the textile industry effluent. It was observed that both processes can completely remove the colour and approximately 85% mineralization of the dye within reaction time of 60 min and 8 min, respectively. The economic analysis placed photo-Fenton as a cost-effective method with treatment cost of approx. 0.0090 US $/litre of wastewater containing Remazol Red dye. Although, Photocatalysis was relatively slow, it is substantially effective in removal/degradation of colour from textile effluent against the biological treatment. The study concludes that photo-Fenton and Photocatalysis are cost-effective and substantial treatment options for removal of toxicity arising from coloured textile effluents.

Keywords Remazol red · Photocatalysis · Photo-Fenton · Decolorization · Textile effluent

1 Introduction

The textile industry has become a significant contributor to water pollution, utilising millions of gallons of freshwater globally, and resulting in the generation of huge volumes of wastewater. The wastewater generated in washing, scouring, mercerizing, and bleaching during textile processing varies significantly, thus causing complex nature of textile effluent. As per an estimate, about 2 million litres of water are consumed to fabricate 10 tonnes of textiles (Yaseen and Scholz 2019). The composition of textile industry effluent varies with the scale of industry, manufacturing processes, type of technology used, and degree/type of treatment to wastewater. So, the industry has ended up as one of the most notorious water-polluting sector. The untreated or poorly treated textile effluent, released either on the land surface or a water body, could relocate a huge cost to the environment, disturbing the ecology and triggering human health problems (Ho et al. 2017). The nature of pollutants released depends on the chemicals that are used in different processes, thus often differs in the values of BOD (biochemical oxygen demand), pH, acidity, turbidity, suspended solids, temperature, etc. The suspended solids in the effluent combine with oily scum and interfere with the oxygen transfer mechanism in the air–water interface. The textile effluents also contain heavy metals, such as Cu, As, Zn, and Cr, capable of harming the environment (Eswaramoorthi et al. 2008). The colorants in the effluent can also cause diseases, such as irritation of skin and dermatitis, haemorrhage, nausea, ulceration of skin (Jothi 2015) etc. The residual dyes in the effluent may increase the BOD of the receiving water body, reducing the DO and reoxygenation process, thus, obstructing the development of photoautotrophic organisms.

The textile industry effluent is characterized by high pH, salinity, intensive colour and persisting odour. The use of azo (N≡N) dyes as colouring pigments for fibres has been a practice for centuries. Although studies have confirmed the ecotoxicity of azo dyes to a number of aquatic species, dyeing process still depends on 60–70% stock of azo dyes against the share of other pigments (Rajkumar and Muthukumar 2012). Synthetic dyes used to colour the textiles end up in the effluent, thus making wastewater toxic and

✉ A. K. Haritash
akharitash@dce.ac.in

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

Published online: 06 January 2022

carcinogenic. The frequently used dye for cellulose textiles is reactive Remazol Red. Remazol dyes are comparatively less reactive than other dyes, thus making them more stable in water. Moreover, the subsequent aerobic degradation of these dyes does not occur, since the masking groups are provided to make them color-fast (Ghaly 2014).

Conventional treatment processes like flocculation, chemical coagulation, simple sedimentation, and biological processes such as aerobic–anaerobic treatment, aerated lagoons, aerobic activated sludge, etc. have been used in the textile industries; however, these do not achieve complete degradation of the dyes (Ho et al. 2017). The implementation and selection of treatment method depends upon the operational skills, technological aids, knowledge, maintenance, and environmental concern as well. Whereas the photothermal dissociation under ambient conditions results in formation of aromatic amines which enhance toxicity and carcinogenicity (Saggioro et al. 2011), the dyes are resistant to aerobic microbial breakdown. Although anaerobic process is used to treat textile effluents, slow rate and incomplete degradation of the dyes result in residual toxicity in treated wastewater. Moreover, phase transfer of toxic dyes takes place into the sludge produced during biological treatment (Maia et al. 2014). On the other hand, physical processes like adsorption, reverse osmosis, etc. are cost/energy-intensive and produce toxic sludge (Rajkumar and Muthukumar 2012). Since there are limitations associated with physical and biological processes, advanced oxidation processes (AOPs) like Photocatalysis and Fenton's treatment have come-up as potential treatment options for rapid chemical oxidation of persistent organic pollutants. AOPs aim at producing highly oxidising reactive hydroxyl radicals ($OH^{\bullet}$) which can oxidise organic compounds, particularly unsaturated molecules like azo dyes, non-selectively and rapidly (Saatci 2010). Therefore, Advanced Oxidation Process (AOP) has become a potential treatment method to decolourize and degrade dyes from textile industry.

The Photocatalytic process relies on excitation of semiconducting oxides (TiO_2, ZnO, Fe_2O_3, ZnS and CdS) thereby promoting the movement of an electron and hole towards the semiconductor surface, where they take part in electron transfer reactions with solvent and the adsorbed species (Balzani et al. 2019). The semiconductor oxide acts as catalyst to generate holes and free radicals ($OH^{\bullet}$) for degradation of complex organic contaminants (Verma and Haritash 2020). TiO_2 is frequently used as an efficient photocatalyst for oxidation of wide range of organic compounds. Studies have reported that ZnO and TiO_2 were the most effective photocatalysts for the decolourization of dyes with slightly higher efficiency of ZnO (Palak et al. 2019; Akyol et al. 2004). The use of commercial grade TiO_2 report complete degradation of Remazol dye under photo-irradiation in a period of approximately 2 h (Saatci 2010), while the use of nano-TiO_2 (P25) reports complete degradation with higher reaction rate. As compared to the traditional chemical oxidation like chlorination, photocatalysis process is more effective for treatment of wastewater (Balzani et al. 2019). Similarly, the Fenton's process has also gained more attention (Asghar 2015) towards degradation of organic compounds. A mixture of ferrous (II) ions and hydrogen peroxide (H_2O_2) produces hydroxyl radicals in the acidic medium, reducing a significant amount of COD and total organic carbon TOC content. Ferrous (II) and hydrogen peroxide can be used for wastewater treatment owing to easy availability and non-toxic nature of iron, and easy to handle nature of 30% hydrogen peroxide (Modirshahla et al. 2007), respectively. Photo-Fenton, on the other hand, combines the Fenton reaction with UV radiation causing photochemical decomposition of hydrogen peroxide. Photo-Fenton is the preferred technique to generate hydroxyl radicals, resulting in an increased rate of degradation (da Rocha Santana 2017). According to a study (Saatci 2010), the photo-Fenton's process has higher oxidation rate than the Fenton's process with reported efficiency of 90% degradation of COD in a reaction time of 20 min. Low cost, simple use and rapid degradation make this process suitable for the treatment of textile wastewater. The other advantages of AOPs over conventional treatment methods are—non-selective and destructive action, and complete mineralization of organic pollutants. AOPs generate low to zero waste and use short-lived chemical species with high oxidizing power to degrade complex molecules. Therefore, the current study aims to evaluate and compare the performance of Photocatalysis, Solar-Photocatalysis, and Photo-Fenton's process to find an appropriate treatment method to decolourise/degrade textile industry effluent containing Remazol Red dye.

2 Materials and methods

2.1 Collection and characterisation of textile effluent

The effluent samples (20 in number) were collected from different textile industries located in Ghaziabad and Gautam Buddha Nagar districts, Uttar Pradesh (UP), India. The samples were collected in pre-rinsed virgin polypropylene plastic bottles of 1.0 L capacity during September to October, 2020 immediately after the first phase of COVID-19 lockdown in India. The samples were characterised for pH, electrical conductivity (EC), Total dissolved solids (TDS), using Labman make (LMMP 30 model) bench-top multiparameter; Total suspended solids (TSS) were determined gravimetrically; analysis for nitrate (NO_3^{2-}), was carried out spectrophotometrically (Lab India make UV 3092 Model) at a wavelength of 220 nm; and Biological Oxygen Demand (BOD_5),

Chemical Oxygen Demand (COD), Total Sulphides, and Ammonia (NH_3) were determined using standard methods as prescribed by APHA (2012). All the samples were analysed in triplicates and intra-lab comparison of analysis was performed to ensure the accuracy of analysis (≥ 95%).

2.2 Chemicals/reagents and experimental setup for dye degradation

The reactive textile dye Remazol Red obtained from Gayatri Dye Chem Industries, Gujarat, India was used to prepare synthetic wastewater (100 mg/l) using ultrapure (Type-1) water. The absorption spectrum of Remazol red dye was plotted in the wavelength range from 190 to 800 nm over a double beam UV–Vis spectrophotometer. The wavelength of maximum absorption (λ_{max}) was obtained at 520 nm; and standard curve of Remazol red dye was plotted in concentration range of 10 to 100 mg/l. Analytical grade (AR) TiO_2 (P25) was obtained from Evonik Industries, Germany; and Ferrous sulphate ($FeSO_4.7H_2O$) (99.5%) and hydrogen peroxide (H_2O_2) (30% w/v) were obtained from Central Drug House (CDH), India. The degree of mineralisation was measured over the TOC Analyser (TOC-L Shimadzu Make, Japan).

The experiments for dye degradation were performed using 200 ml extract of synthetic stock solution (100 mg/l), in triplicates, at room temperature (25°C) with continuous stirring at 150 rpm and air-sparging. The experiments were conducted to study the effect of varying concentrations of TiO_2 (0.1 g/l to 20.0 g/l) and H_2O_2 at pH 7.0 for photocatalytic treatment. Photo-Fenton experiments were studied to examine the effect of varying Fe^{2+} concentration (0.33 mM to 1 mM); effect of H_2O_2 (1 mM to 10 mM); and effect of varying pH (1.0–4.0) and optimized values were recorded from the same. Since Fenton's process operates in acidic pH range only, the experiments were performed in pH range of 1.0–4.0. A sample volume of 5 ml was extracted for analysis at a regular time interval varying from 5 to 15 min depending upon the treatment pace until the process got stabilized. The Photocatalytic and Photo-Fenton experiments were carried out in a fabricated UV-chamber having 8 UV tubes (Philips 36 W each), whereas solar-photocatlytic treatment was conducted under sunlight. The aggregated source intensity of the UV-chamber was 672 W/m^2 as explained by Verma and Haritash (2019). The results of different processes were compared and their efficiency towards complete removal of Remazol was evaluated. To study the mineralization of dye, initial and final values of TOC were determined. The percent removal/ degradation of dye was calculated using the following equation:

$$\text{Efficiency } (\%) = \left[\left(C_i - C_t\right)/ C_i\right] \times 100 \quad (1)$$

where C_i is the initial dye TOC concentration; and C_t is the dye TOC concentration at time 't'.

3 Results and discussion

3.1 Characteristics of textile industry effluent

The physico-chemical characteristics of textile effluent are given in Table 1. The average pH of untreated influent from was 7.5 against pH of 7.2 for treated effluent. It is reported that acidic conditions (low pH) or alkaline conditions (high pH) can alter the enzymatic structure and development of microorganisms in the receiving water-body, while pH > 8.5 can affect fish production (Edmund 1998). The average value of EC in textile effluent before and after treatment was 5.2mS/cm and 4.3mS/cm, respectively. High EC of the sample indicates the presence of ions like sodium, potassium, iron, etc. which are present in surfactants, detergents, and other chemicals used during the scouring process to remove natural impurities (Holkar et al. 2016). The average TDS of the treated effluent was 2147 mg/l. The high TDS has the potential to deteriorate the quality of surface water

Table 1 Characteristics of industrial wastewater from medium scale textile industries

Parameter	Unit	Influent		Effluent	
		Range	Mean ± SD	Range	Mean ± SD
pH	–	3.5–11.3	7.5 ± 1.79	3.1–9.7	7.2 ± 1.31
EC	mS/cm	1–11	5.2 ± 2.9	1–13	4.3 ± 3.3
TDS	mg/l	640–5505	2721 ± 1500.8	464–6660	2147 ± 1701
TSS	mg/l	8–698	179.9 ± 169.2	2–176	31.2 ± 40.4
BOD_5	mg/l	40–240	141.0 ± 80.4	20–200	79 ± 54
COD	mg/l	240–2400	835.0 ± 741.1	80–640	320 ± 175
Nitrate	mg/l	2–62	28.0 ± 16.9	0–16	5.7 ± 3.8
Total Sulphides	mg/l	1.2–88	13.7 ± 18.7	0–56	6.8 ± 12.5

EC electrical conductivity, *TDS* total dissolved solids, *TSS* total suspended solids, *BOD_5* 5-day biochemical oxygen demand, *COD* chemical oxygen demand, *SD* standard deviation

as a consequence of which the aquatic life gets adversely affected. The average TSS of effluent was 180 mg/l, nitrates were 28 mg/l, total sulphides as 13.7 mg/l, average BOD was 141 mg/l and COD was 850 mg/l. Due to the high BOD, the dissolved oxygen in receiving water may reduce significantly causing septic conditions that leads to death of fishes and other aquatic species. Chemical oxygen demand (COD) indicate the occurrence of organic as well as inorganic matter in the textile dyeing industry effluents. It comprises both the biodegradable and non-biodegradable portions of live bacterial attack, but average BOD/COD ratio < 0.3 indicates that it can be oxidized by strong chemical oxidants (Abbasi 1998; Tan et al. 2000; Chiron et al. 2000) thus indicating that treatment by chemical oxidation (AOPs) is a preferred method.

3.2 Photocatalytic degradation of Remazol Red dye

The TiO_2-assisted photocatalytic degradation confirmed significant removal of Remazol Red dye from the textile effluent. The effect of regulating factors (TiO_2, H_2O_2, and light) on efficiency of the process is as discussed below.

3.2.1 Effect of catalyst size, type of light, and Irradiation time

To optimize the effect of catalyst on degradation of Remazol Red dye, commercially available normal TiO_2 and nano-TiO_2 (P25) were used in the present study. The comparison of normal TiO_2 against nano-TiO_2, revealed that nano-TiO_2 has got better potential towards treatment of textile dyes (Fig. 1a). Since nano-TiO_2 has more surface area, absorption of energy (UV-associated) is more and so is the production of $OH^{\bullet}$ radicals. The nano-TiO_2 resulted in almost complete

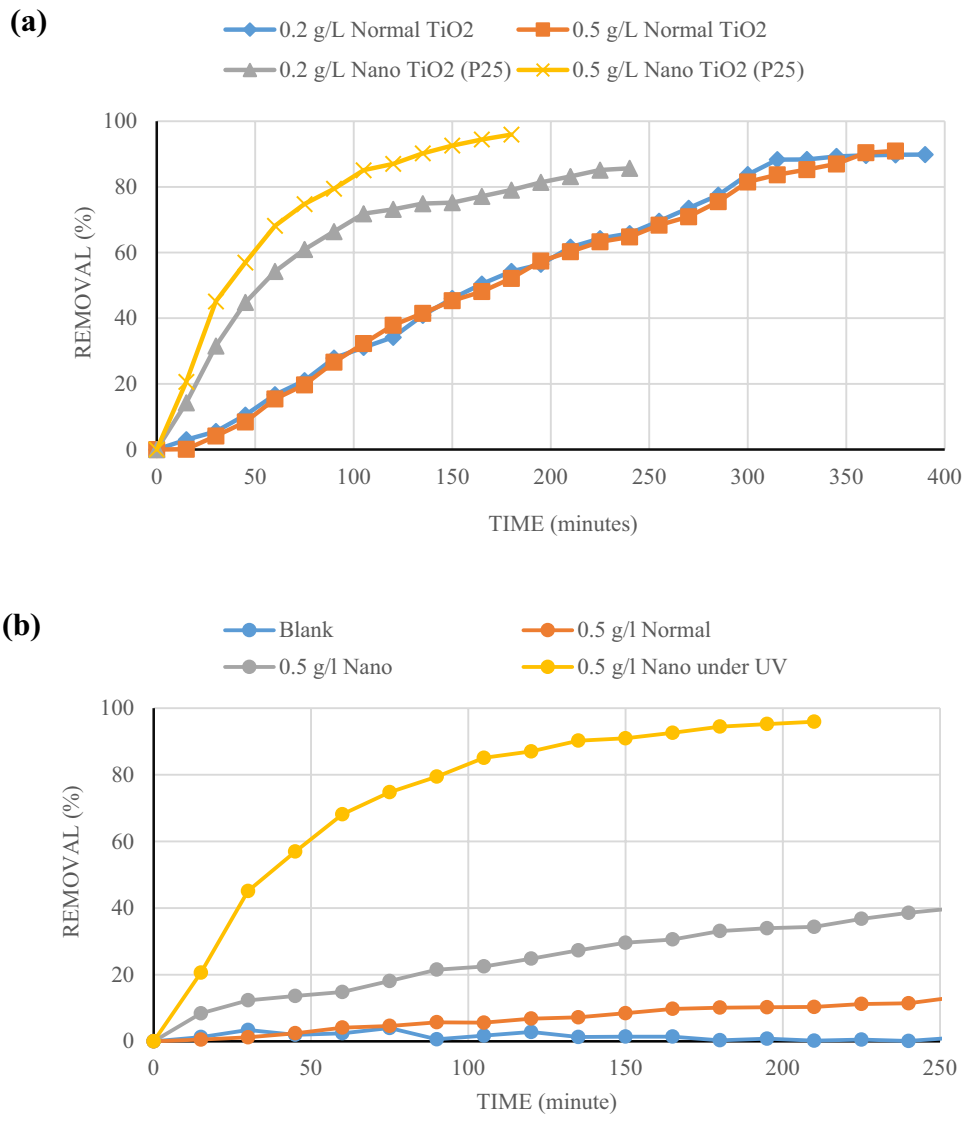


Fig. 1 **a** Photocatalytic treatment of Remazol dye against Normal TiO_2 and Nano TiO_2 (P25); **b** Photocatalytic treatment of Remazol Red dye under sunlight and UV source

removal of colour (at a dose 0.5 g/l) within 3.0 h, whereas normal TiO_2 could remove only 80% of the colour even after 5.0 h. Considering the edge of nano-TiO_2 over normal TiO_2, rest of the experiments were conducted using nano-TiO_2 for treatment of textile effluent.

Since, sunlight only consists of 3–5% UV light, degradation occurs slower (Davis and Huang 1990). The solar-photocatalytic experiment set of 0.5 g/l using normal TiO_2/H_2O_2/Sunlight and nanoTiO_2/H_2O_2/Sunlight showed 13% and 40% degradation efficiency in 255 min, respectively (Fig. 1b). Poor degradation rate could be attributed to the ambient conditions as the experiment was carried out in winters when intensity of sunlight was lower. Moreover, the lower fraction of UV (~5%) in total solar terrestrial radiation results in lower energy and weak activation of TiO_2. Since the energy associated with UV light is high, it can easily activate TiO_2 resulting in formation of $OH^{\bullet}$ radicals. It is observed that higher light intensity is directly proportional to generation of $OH^{\bullet}$ radicals thus facilitating the faster dye degradation.

Under the optimized concentration of TiO_2 (1.0 g/l), 98% of the Remazol dye was degraded in 60 min of the irradiation time. The photodegradation of the dye promoted by photocatalyst via the direct photolysis in 60 min of the irradiation time is responsible for degradation of only the molecules adsorbed on the surface of TiO_2. Increasing the catalyst concentration increases the turbidity of the solution hence strongly inhibiting the light penetration resulting in less degradation.

3.2.2 Effect of H_2O_2

To inhibit the problem of electron–hole recombination is to add another electron acceptor, i.e., H_2O_2, to the reaction. It enhances the degradation of compounds due to a more efficient generation of hydroxyl radical and inhibition of electron–hole recombination (Mahmoodi et al. 2006). The colour degradation was faster in the presence of H_2O_2 as it is a powerful oxidizing agent; and it increased the photodegradation efficiency. The experiment conducted with nano TiO_2 dose of 0.10 g/l, 0.20 g/l, and 0.50 g/l, and constant concentration of 5 mM H_2O_2 in the dye solution exhibited a faster rate than the experiment in the absence of H_2O_2. Dye degradation efficiency achieved in both the experiments were though similar (90% and 94%); however, it was acquired in 210 min in the presence of H_2O_2 compared to 315 min in its absence (Fig. 2). The addition of H_2O_2 with TiO_2/UV has a synergistic effect. The UV irradiation causes H_2O_2 to photolyze, generating two hydroxyl radicals (2OH•) (6) or hydrogen ions (7) for the production of hydroperoxyl radicals (5) (Konstantinou & Albanis 2004). Since, electron–hole recombination is hampered, the hydroxyl radicals are induced either directly (2) or indirectly (4) (Samsudin et al. 2015). Similar findings showing synergistic effect between photocatalytic activity from nano TiO_2 and H_2O_2 been reported by Karimi et al. (2014):

$$TiO_2(e^-_{CB}) + H_2O_2 \rightarrow TiO_2 + \cdot OH + OH^- \quad (2)$$

$$TiO_2(e^-_{CB}) + O2 \rightarrow TiO_2 + O^2 \quad (3)$$

$$O^{2-} + H_2O_2 \rightarrow OH^- + \cdot OH + O_2 \quad (4)$$

$$O^{2-} + H^+ \rightarrow HO_2\cdot \quad (5)$$

$$H_2O_2 \rightarrow 2 \cdot OH \quad (6)$$

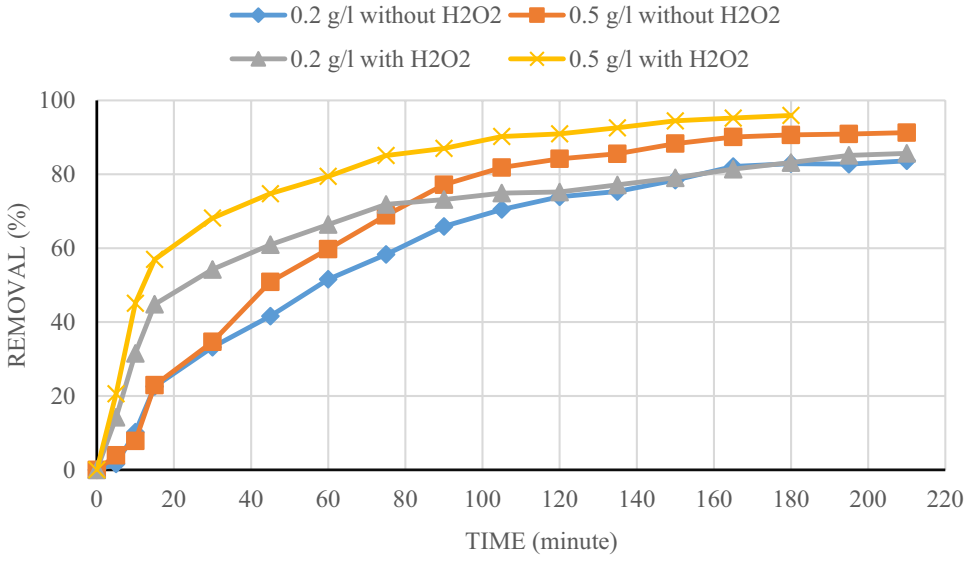


Fig. 2 Photocatalytic (TiO_2) treatment of dye in the presence and absence of H_2O_2

$$H_2O_2 \rightarrow O_2^{2-} + 2H^+ \quad (7)$$

where e_{CB}^- represents the electron in conduction band.

3.2.3 Effect of catalyst loading

It was observed that dye degradation rate was slower when the catalyst dose was less because of the lesser availability of active sites, thus, less surface area of the photocatalyst. In addition, the limited generation of hydroxyl radicals and superoxide ions demands a longer irradiation time to degrade the high-strength dye (100 mg/l). The degradation rate increased with an increase in the TiO_2 dose as more surface area was available for adsorption. Similar observations have been made in other studies too (Neppolian et al. 2002; Kaur and Singh 2007). It was noted that at 1.0 g/l TiO_2 almost complete degradation (~98%) was achieved in 60 min (Fig. 3). The results are in good agreement with the results of Maia et al. (2014), where 1.0 g /l of TiO_2 completely degraded Remazol Blue RGB and Remazol Carbon RGB and nearly 90% for the Remazol Yellow Gold RGB, in 120 min of photodegradation. Complete degradation of reactive dye, RB4 dye (anthraquinone dye) was achieved with TiO_2 dose of 1.0 g/l in 40 min of irradiation time as reported by Samsudin et al. (2015). At a higher dose, scattering of light caused by the opacity of the solution (cloudy suspension) affects the degradation efficiency (Carp et al. 2004). It, thus, reduces the number of photons absorbed resulting in fewer hydroxyl radicals generation from TiO_2 to oxidise the dye. The phenomena was observed while augmenting TiO_2 dose from 1.5 g/l to 20.0 g/l in the current study. While attaining the maximum degradation efficiency, the turbid solution formed decreased the efficiency with increased TiO_2 dose. This cloudy suspension was also reported in degradation of dyes like Reactive Red 180, Reactive Red 2 and Reactive Blue 19 (Hamlin et al. 1999; Lee and Pavlostathis 2004). Though the degradation efficiency was maximum with TiO_2 dose of 1.0 g/l in the current study, the rate kinetics was, however, higher with TiO_2 dose 10.0 g/l and 20.0 g/l also contributing to maximum degree of mineralisation of the order of 80% and 85%, respectively. The higher degradation efficiency while increasing TiO_2 dose may be because electron–hole pairs are generated when photons, the energy packets, are absorbed by the photocatalyst, exciting the electrons to jump from valence band to conduction band. These electron–hole pairs either react with the dye molecule directly or react with the surface-bound water to generate hydroxyl radicals. More the loading of TiO_2, more will be the availability of active sites for adsorption, hence, more electron–hole pair to bind with water resulting in more OH• radicals. The typical mechanism is represented by the following reaction equations:

$$TiO_2 + h\upsilon \rightarrow h^+ + e^- \quad (8)$$

$$O_2 + e^- \rightarrow O^{-2} \quad (9)$$

$$h^+ + H_2O \rightarrow OH^{\bullet}_{ad} + H^+aq \quad (10)$$

$$h^+ + \text{dye} \rightarrow \text{products} \quad (11)$$

$$OH^{\bullet}_{ad} + \text{dye} \rightarrow \text{products} \quad (12)$$

$$e^- + O_2 \rightarrow O_2(aq)^- \quad (13)$$

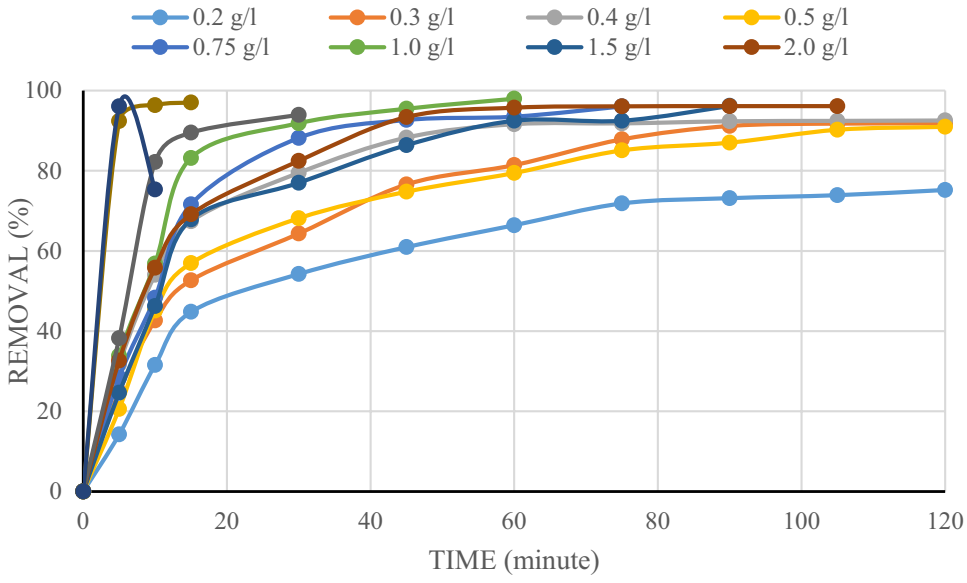


Fig. 3 Photocatalytic treatment of Remazol Red dye at varying doses (g/l) of Nano TiO_2

$$e^- + O_2 + H_2O \rightarrow HO_2(aq) + OH^+(aq). \quad (14)$$

3.3 Photo-Fenton experiment

The degradation of Remazol Red dye during Fenton's treatment was observed to be higher compared to photocatalysis. The regulating parameters of Fenton's treatment (Fe^{2+}, H_2O_2, pH etc.) were optimized to arrive at better efficiency towards dye degradation.

3.3.1 Influence of Fe^{2+} concentration

In Fenton's and Photo-Fenton's processes, ferrous ion (Fe^{2+}) is a significant factor determining the rate of the reaction. In the current study, the influence of variable concentrations of Fe^{2+} (0.33 mM; 0.5 mM; 1.0 mM) on decolorisation of Remazol Red dye was studied at pH 3.0. The concentration of H_2O_2 was kept constant at 5.0 mM, while the Fe^{2+} concentration was varied to maintain the ratio of 1:5, 1:10, and 1:15 for Fe^{2+}: H_2O_2. It was observed that Fe^{2+} at concentration of 0.5 mM (ratio 1:10 with H_2O_2), achieved complete decolourization in 8 min, while it was 87% and 96% for the Fe^{2+} dose at ratio of 1:5 and 1:15 which stabilized in the reaction time of 6 min and 11 min, respectively (Fig. 4). A similar study (Punzi et al. 2012) reported complete colour removal of Remazol Red RR at 0.25 mM iron concentration in less than 30 min. At lower iron concentration, the degradation is less because of limited availability of $OH^{\bullet}$ radicals to carry out oxidation. More iron concentration in the sample causes a slight reduction in the degradation rate of Remazol dye. $FeSO_4$ dissociates as Fe^{2+} which acts as the catalyst to dissociate H_2O_2 generating OH radicals but the high iron concentration has a scavenging effect on hydroxyl radicals (Verma and Haritash 2019).

3.3.2 Influence of H_2O_2 concentration

It was observed that degradation efficiency increased from 61 to 100% when H_2O_2 concentration was increased from 1.0 to 5.0 mM fixed iron concentration of 0.5 mM and pH 3.0. At maximal concentration of 10.0 mM, complete decolourisation was observed in about 10 min against the reaction time of 8.0 min at optimized H_2O_2 level of 5.0 mM (Fig. 5). Hence, the H_2O_2 concentration of 5.0 mM was considered as the optimized dose. Several studies on various textile dyes have also reported that decolourization after reaching its maximum efficiency, later stabilizes and has no effect even after increasing H_2O_2 concentration. This is due to the scavenging effect of hydroxyl radicals and generation of OH_2 (15) radicals at higher concentration of H_2O_2 (Chacó et al. 2006):

$$HO^{\bullet} + H_2O_2 \rightarrow H_2O + HO_2. \quad (15)$$

The TOC analysis also reported the highest mineralisation of about 85% in the experiment with Fe^{2+} concentration of 0.5 mM and H_2O_2 concentration of 5.0 mM as compared to others.

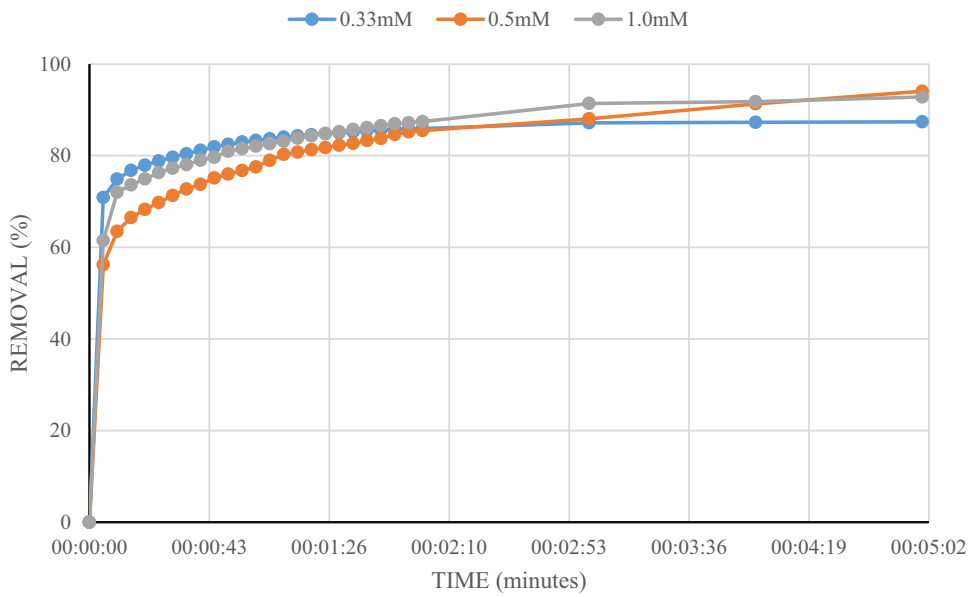


Fig. 4 Effect of Fe^{2+} concentration (mM) on degradation of Remazol Red dye during Photo-Fenton treatment

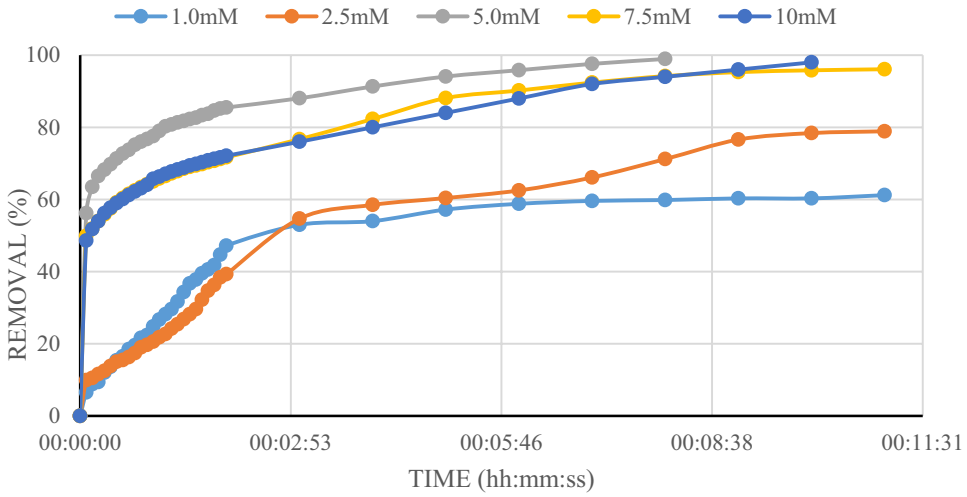


Fig. 5 Photo-Fenton treatment of dye at varying H_2O_2 concentration (mM) and pH 3.0

3.3.3 Influence of pH

The pH of the textile industry effluent governs the oxidation of organic impurities. Since the generation of •OH radicals is also pH dependent, under acidic conditions (2.0–4.0), maximum hydroxyl radicals are generated to carry out effective degradation. The current study considered the effect of pH values in the range of 1.0–5.0 with an optimized dose of Fe^{2+} and H_2O_2 ratio of 1:10. Quicker and complete decolorization (100%) of the dye was observed at pH 3.0 within 8 min of the reaction time. The complete degradation at pH 3.0 took almost half the time for maximum degradation than at pH 2.5 and 3.5 as indicated in Fig. 6. The results were in strong agreement with the optimum degradation of organic pollutants, azo dyes like AO7 (acid orange 7) (Sharma et al. 2016), Mordant red 73 dye, Amoxicillin (AMX) (Verma and Haritash 2019), and several other studies at pH 3.0. It is reported that the stability of H_2O_2 and Fe^{2+} is also more at pH 3.0 (Kwon et al. 1999). Moreover, the degree of mineralisation was also found maximum (~ 85%) at pH 3.0.

As the pH of the dye solution increased from 1.0 to 3.0, decolorization also increased; however, it decreased at pH 3.5 and 5.0. Degradation was minimum (93%) at pH 1.0 achieved in 40 min of reaction time. Complex iron

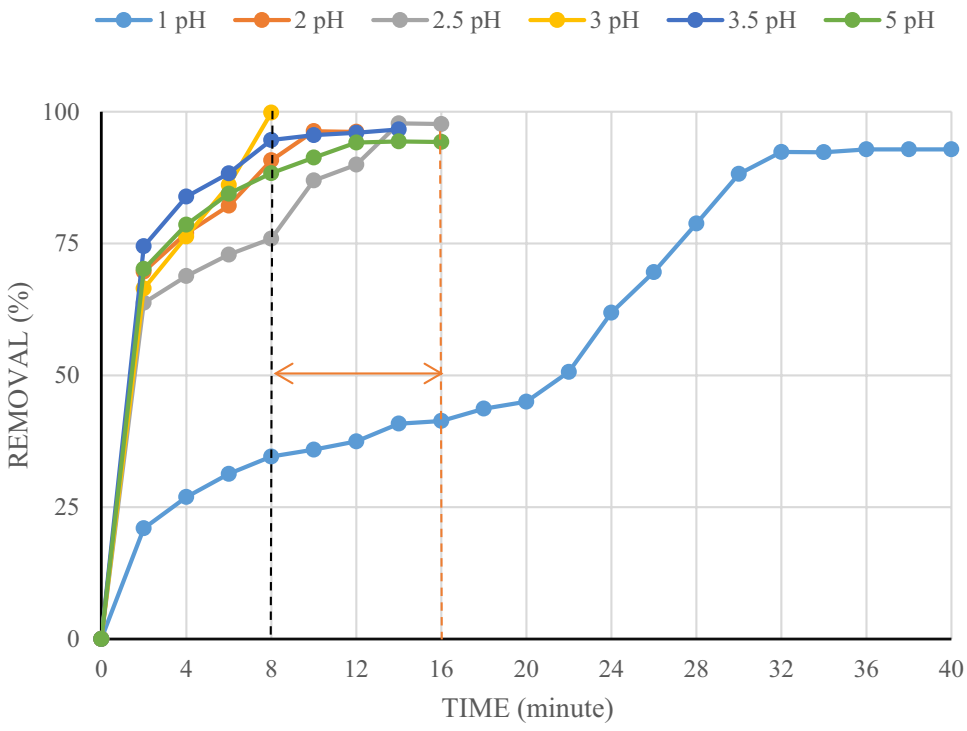


Fig. 6 Effect of pH on degradation of Remazol Red dye during Photo-Fenton treatment

Table 2 Economic analysis of the operating cost of experiments

Treatment method	Energy per hour (kWh)	Time for treatment (minute)	Rate of Energy (per kWh) (INR)	Cost of electricity (INR)	Chemical consumed	Cost of chemical per litre (INR)	Cost of treatment in per litre (UD$)
Photo-Fenton	0.3	8	8.50	0.34	$FeSO_4.7H_2O$ H_2O_2 (30% w/v)	0.07 0.25	0.0090
Photocatalysis	0.3	60	8.50	2.55	TiO_2 (P25) H_2O_2 (30% w/v)	18.14* 0.25	0.29

*Calculated considering 92% recovery of TiO_2

species and oxonion ($H_3O_2^+$) ions are formed and H^+ ions scavenge the hydroxyl radicals at lower pH values (Kwon et al. 1999). In addition, with $2.0 > pH > 4.0$, the oxidation potential of hydroxyl radicals decreases. On the contrary, the degradation efficiency at pH 5.0 was lesser (94%) than at pH 3.0, since both iron and H_2O_2 become unstable at alkaline conditions. Iron precipitates as hydroxide and does not react with H_2O_2 to form hydroxyl radicals, whereas hydrogen peroxide is destabilized to forms water and oxygen.

3.4 Recovery of TiO_2 and Economic Analysis

Since TiO_2 act as the catalyst in decolourization of dye, most of it can be recovered. However, some of the TiO_2 was lost during the experiments, while it goes through centrifuging before taking the observations over the spectrophotometer. A known amount of TiO_2 was added initially to the sample. After the sample got stabilised, it was oven dried at 105 ^{O}C. The recovery of TiO_2 was calculates as

$$TiO_2 \text{ recovered} = \frac{(w_1 - w_2)}{w_1} \times 100, \tag{16}$$

where, w_1 - initial weight of the TiO_2 added to sample, w_2 - final weight of the TiO_2 recovered post experiment.

Up to 70–97% of the TiO_2 added was recovered during the experiment under optimized conditions of photocatalysis. Therefore, optimization of TiO_2 contributes towards reducing the overall cost of treatment of the synthetic wastewater and making the process green and sustainable. Hence, recycling and recovery of TiO_2 could be foreseen as a sustainable approach for treating the wastewater.

The economic analysis of the operating cost of experiments for degradation of Remazol dye by photocatalysis and Photo-Fenton treatment was conducted. Since, Photo-Fenton and Photocatalytic degradation exhibited the maximum degradation efficiency; economic comparison of the respective methods was conducted with respect to reagents usage and energy consumption. The experiment were conducted under UV light (8 UV tubes) each having power rating of 36 W, magnetic stirrer of power rating 8.5 W; and air sparger of power rating 3.5 W, thus, making a total power consumption of 300 W. The price of commercial electricity in Uttar Pradesh Province is INR 8.5 for 1KWh. In case of Photocatalytic treatment, the maximum colour removal efficiency (98%) was obtained in 60 min for 1.0 g/l TiO_2 dosage. The nano-sized TiO_2 used in the experiment costs INR 22,680 per 100 g; however, only 8% of the amount added was consumed (92% of it was recovered) during the process. The 8% of the TiO_2 consumed costs INR 3.63 per 200 ml (INR 18.14 per litre of the solution). Each litre of wastewater sample of strength 100 ppm with respect to Remazol Red consumed 0.3 KWh of power in 1 h while achieving the maximum efficiency making the energy cost approximately INR 2.55/-. Hence, the total cost is 0.29 US $ per litre for 100 ppm synthetic dye solution of Remazol Red. In case of Photo-Fenton experiment, complete degradation (100%) was achieved in 8 min (~0.13 h). Hence, the total power of 0.04 KWh was consumed in 8 min, making the energy cost of 0.0090 US $ per litre, which is the total cost for Photo-Fenton experiment (Table 2). Hence, it is observed that Photo-Fenton is a far cheaper and efficient method for colour removal from textile industry effluent.

4 Conclusion

Based on the results obtained in the present study, it is concluded that TiO_2-photocatalysis and photo-Fenton process are efficient and sustainable methods for removal of colour (dye) from the textile industry effluent. Since production of $OH^{\bullet}$ radicals and their non-selective oxidative action is associated in both the processes, degradation of other organic impurities from industrial effluent is also expected. The optimization of regulating factors in both the processes can significantly enhance the efficiency towards degradation and reduce the treatment cost at the same time. The economic comparison based on the input of energy and materials revealed that photo-Fenton process is not only a rapid method for colour removal, but is also cost-effective, thus, making it a sustainable option for degradation of dyes in textile effluent.

Acknowledgements The authors acknowledge the help of textile industries in the collection of wastewater samples.

References

Abbasi SA (1998) Water quality sampling and analysis. Discovery Publishing House, Daryaganj

Akyol A, Yatmaz HC, Bayramoglu M (2004) Photocatalytic decolorization of Remazol Red RR in aqueous ZnO suspensions. Appl Catal B 54(1):19–24. https://doi.org/10.1016/j.apcatb.2004.05.021

Asghar AR (2015) Advanced oxidation processes for in-situ production of hydrogen peroxide/hydroxyl radical for textile wastewater treatment: a review. J Clean Prod 87:826–838

Balzani V, Pacchioni G, Prato M, Zecchina A (2019) Solar driven chemistry: towards new catalytic solutions for a sustainable world. Rend Fis Acc Lincei 30(3):443–452. https://doi.org/10.1007/s12210-019-00836-2

Carp O, Huisman CL, Reller A (2004) Photoinduced reactivity of titanium dioxide. Prog Solid State Ch 32(1):33–177. https://doi.org/10.1016/j.progsolidstchem.2004.08.001

Chacó JM, Leal MT, Sánchez M, Bandala ER (2006) Solar photocatalytic degradation of azo-dyes by photo-Fenton process. Dyes Pigm 69(3):144–150. https://doi.org/10.1016/j.dyepig.2005.01.020

Chiron S, Fernandez-Alba A, Rodriguez A, Garcia-Calvo E (2000) Pesticides chemical oxidation: state of the art. Water Res 34:366–377

da Rocha Santana RM (2017) Degradation and kinetic study of Reactive blue BF-5G and Remazol red RB 133% dyes using Fenton and photo-Fenton process. Rev Eletrônica Gest Educ Tecnol Ambient 21(2):104–118

Davis AP, Huang CP (1990) The removal of substituted phenols by a photocatalytic oxidation process with cadmium sulfide. Water Res 24(5):543–550. https://doi.org/10.1016/0043-1354(90)90185-9

Edmund MJ (1998) Understanding factors that affect pH and guide to alkalinity and pH control. Sea scope. Aquar Syst 5:35

Eswaramoorthi S, Dhanapal K, Chauhan D (2008) Advanced in textile waste water treatment: The case for UV-ozonation and membrane bioreactor for common effluent treatment plants in Tirupur, Tamil Nadu, India. Environment with People's Involvement & Co-ordination in India. Coimbatore, India

Ghaly AE (2014) Production, characterization and treatment of textile effluents: a critical review. J Chem Eng Process Technol 5(1):1–18

Hamlin JD, Phillips DAS, Whiting A (1999) UV/Visible spectroscopic studies of the effects of common salt and urea upon reactive dye solutions. Dyes Pigm 41(1):137–142. https://doi.org/10.1016/S0143-7208(98)00076-X

Ho SOL, Khairil YW, Mohammad A (2017) Comparative study on the biodegradation of mixed remazol dyes wastewater between integrated anaerobic / aerobic and aerobic sequencing batch reactors. Rend Fis Acc Lincei 28(3):497–501. https://doi.org/10.1007/s12210-017-0622-2

Holkar CR, Jadhav AJ, Pinjari DV, Mahamuni NM, Pandit AB (2016) A critical review on textile wastewater treatments: Possible approaches. J Environ Manage 182:351–366. https://doi.org/10.1016/j.jenvman.2016.07.090

Jothi D (2015) Application of enzyme extracted from aloe vera plant in chemical pretreatment of cotton knitted textile to reduce pollution load. World J Eng 03(03):37–44. https://doi.org/10.4236/wjet.2015.33b007

Karimi L, Zohoori S, Yazdanshenas ME (2014) Photocatalytic degradation of azo dyes in aqueous solutions under UV irradiation using nano-strontium titanate as the nanophotocatalyst. J Saudi Chem Soc 18(5):581–588. https://doi.org/10.1016/j.jscs.2011.11.010

Kaur S, Singh V (2007) TiO_2 mediated photocatalytic degradation studies of Reactive Red 198 by UV irradiation. J Hazard Mater 141(1):230–236. https://doi.org/10.1016/j.jhazmat.2006.06.123

Konstantinou IK, Albanis TA (2004) TiO2-assisted photocatalytic degradation of azo dyes in aqueous solution: kinetic and mechanistic investigations: A review. Appl Catal B 49(1):1–14. https://doi.org/10.1016/j.apcatb.2003.11.010

Kwon BG, Lee DS, Kang N, Yoon J (1999) Characteristics of p-chlorophenol oxidation by Fenton's reagent. Water Res 33(9):2110–2118. https://doi.org/10.1016/S0043-1354(98)00428-X

Lee YH, Pavlostathis SG (2004) Decolorization and toxicity of reactive anthraquinone textile dyes under methanogenic conditions. Water Res 38(7):1838–1852. https://doi.org/10.1016/j.watres.2003.12.028

Mahmoodi NM, Arami M, Limaee NY, Gharanjig K, Ardejani FD (2006) Decolorization and mineralization of textile dyes at solution bulk by heterogeneous nanophotocatalysis using immobilized nanoparticles of titanium dioxide. Colloids Surfaces A Physicochem Eng Aspects 290(1):125–131. https://doi.org/10.1016/j.colsurfa.2006.05.012

Maia CG, Oliveira AS, Saggioro EM, Moreira JC (2014) Optimization of the photocatalytic degradation of commercial azo dyes in aqueous TiO_2 suspensions. React Kinet Mech Catal 113(1):305–320. https://doi.org/10.1007/s11144-014-0737-8

Modirshahla N, Behnajady MA, Ghanbary F (2007) Decolorization and mineralization of C.I. Acid Yellow 23 by Fenton and photo-Fenton processes. Dyes Pigm 73(3):305–310. https://doi.org/10.1016/j.dyepig.2006.01.002

Neppolian B, Choi HC, Sakthivel S, Arabindoo B, Murugesan V (2002) Solar light induced and TiO2 assisted degradation of textile dye reactive blue 4. Chemosphere 46(8):1173–1181. https://doi.org/10.1016/S0045-6535(01)00284-3

Palak P, Chauhan S, Kumar K, Haritash AK (2019) Degradation of AY36 using TiO_2-UV photocatalytic system. Appl Chem Eng 2(1):404

Punzi M, Mattiasson B, Jonstrup M (2012) Treatment of synthetic textile wastewater by homogeneous and heterogeneous photo-Fenton oxidation. J Photochem Photobiol A 248:30–35. https://doi.org/10.1016/j.jphotochem.2012.07.017

Rajkumar K, Muthukumar M (2012) Optimization of electro-oxidation process for the treatment of Reactive Orange 107 using response surface methodology. Environ Sci Pollut Res 19(1):148–160. https://doi.org/10.1007/s11356-011-0532-2

Saatci Y (2010) Decolorization and mineralization of Remazol red F3B by Fenton and photo-fenton processes. J Environ Eng 136(9):1000–1005. https://doi.org/10.1061/(asce)ee.1943-7870.0000226

Saggioro EM, Oliveira AS, Pavesi T, Maia CG, Ferreira LFV, Moreira JC (2011) Use of titanium dioxide photocatalysis on the remediation of model textile wastewaters containing azo dyes. Molecules 16(12):10370–10386. https://doi.org/10.3390/molecules161210370

Samsudin EM, Goh SN, Wu TY, Ling TT, Hamid SBA, Juan JC (2015) Evaluation on the photocatalytic degradation activity of reactive blue 4 using pure anatase nano-TiO_2. Sains Malays 44(7):1011–1019. https://doi.org/10.17576/jsm-2015-4407-13

Sharma A, Verma M, Haritash AK (2016) Degradation of toxic azo dye (AO7) using Fenton's process. Adv Environ Res 5(3):189–200. https://doi.org/10.12989/aer.2016.5.3.189

Tan BH, Teng TT, Omar AKM (2000) Removal of dyes and industrial wastes by Magnesium Chloride. Water Res 34:597–601

Verma M, Haritash AK (2019) Degradation of amoxicillin by Fenton and Fenton-integrated hybrid oxidation processes. J Environ Chem Eng 7(1):102886. https://doi.org/10.1016/j.jece.2019.102886

Verma M, Haritash AK (2020) Photocatalytic degradation of Amoxicillin in pharmaceutical wastewater: A potential tool to manage residual antibiotics. Environ Technol Innov 20:101072. https://doi.org/10.1016/j.eti.2020.101072

Yaseen DA, Scholz M (2019) Textile dye wastewater characteristics and constituents of synthetic effluents: a critical review. Int J Environ Sci Technol 16(2):1193–1226. https://doi.org/10.1007/s13762-018-2130-z

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:03/Issue:11/November-2021 Impact Factor- 6.752 www.irjmets.com

CONCENTRATED SOLAR POWER (CSP) HYBRIDISATION: TRANSFORMING THE ENERGY SECTOR OF INDIA

Sourabh Yadav [*1], Shubham Sharma[*2], Shubham Maurya[*3], Sandeep Yadav[*4], Dr. J.P Kesari[*5]

[*1,2,3,4] Student, Department of Mechanical Engineering, Delhi Technological University Bawana Road, Delhi -110042, India.

[*5] Professor, Department of Mechanical Engineering, Delhi Technological University Bawana Road, Delhi -110042, India.

ABSTRACT

This paper is aimed to study and realize the enormous potential of Solar Concentrated Power to fill in the deficit of the energy scenario in and around the World with India at focus. We started with the literature review, which includes the advancement in technologies from 2010-21; this is followed by an essential discussion on the various types of solar thermal techniques such as Parabolic-trough system, Solar power tower, Parabolic dish system, and Fresnel reflectors. These technologies were studied along with their working, applications, history, and comparisons of one technology with others. Taking inspiration from different case studies around the world, we have also tried to study the current CSP projects of India and the potential in the form of future scope of CSP that India could achieve from its present scenario. With this paper, we have tried to stress that India has ample access to solar irradiance. India must utilize this privilege to the maximum extent, increasing the dependence on solar-based technologies for electricity generation during sunshine hours.

Keywords: Concentrated Solar Power; Solar power tower; Parabolic-trough system; Parabolic-dish system; Fresnel reflectors; Non-Conventional Energy Sources; CSP.

I. INTRODUCTION

Concentrated Solar Power generates electricity using the Sun's heat which is a natural energy source. CSP technology can generate thermal as well as electrical energy demands. CSP technology produces heat between 400°C to 1,000°C by focusing on the Sun's Direct Normal Irradiation (DNI) **[16].** There are four types of solar concentrator technologies: power tower systems, trough systems, linear fresnel systems, and parabolic dish systems. These technologies can be combined with fossil fuels and solar energy due to their thermal nature. A hybrid of these can significantly improve the value of CSP technology. It can do it by increasing the availability and the capability to dispatch and lowering the cost by improving the use of power block equipment and reducing any technology-related risk by pushing conventional fuel use, like if the collector needs repair **[17].** The Concentrated Solar Power Technology produces energy without any emissions and causes any pollution; its waste is materials that are not toxic, flammable, or otherwise dangerous: For the nearby residents, the whole system is not a threat or a source of noise **[18].**

The following text gives a brief description and significant developments of the four CSP systems, parabolic-trough, parabolic-dish, solar-power tower, and Fresnel reflector systems. The report includes the basics of different processes involved, brief history, comparisons, and examples that would help us understand the present CSP/CST plants worldwide.

2.1 Parabolic-Trough system:

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:03/Issue:11/November-2021 Impact Factor- 6.752 www.irjmets.com

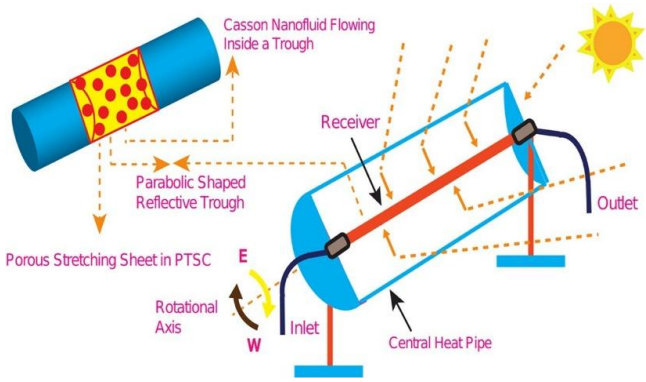


Figure 1: Schematic view of the Parabolic Trough system **[25]**

Initially, this system was used in southern California in 1984 by Luz International Co. as a 14MW parabolic-trough power plant. The system was combined with gas to make it run continuously, even during repairs **[26]**. These systems consist of various linearly interconnected parabolic troughs, steam turbines, super-heaters, re-heaters, pre-heaters, condensers, and steam-generator for electricity generation, as shown in figure 1.

The absorption tube provides a large absorption-coefficient of the irradiation from the Sun. This tube is coupled at the focus of the trough and maximizes the absorption of the irradiation, and reduces the energy loss from the absorption tube. This decrease in heat losses through parabolic trough are generally due to the low emission from the absorption tube. With a rotational carrier set upright, these follow the Sun east-west (EW) position; however, once this is oriented in the horizontal position, the tilt is changed to trace the variations in the Sun's height. Through this, mineral oils could be heated to a temperature of 400°C. Subsequently, the heated working fluid is transferred using pumps from the tube to the heat exchanger. Generally, these parabolic-trough systems are considered as good, and this is because these provide the best land-use factor of any solar-driven technology **[27].**

According to **[40],** at the world level, there are presently twenty active trough power plants worldwide: eleven in Spain, 5 in the USA, 2 in Iran, 1 in Italy, and 1 in Morocco. According to the same website data, twenty-seven trough power plants are getting constructed worldwide: twenty-two plants in Spain, 2 in the United States, 1 in India, 1 in Egypt, and 1 in Algeria. In the United States, the following parabolic power plants are operational: Solar Energy Generating Systems-354 MW (Mega Watts capacity), Martin Next Generation Solar Energy Center-75 MW, Keahole Solar Power-2 MW, Nevada Solar One-64 MW, and Saguaro Solar Power Station-1 MW **[40].** Currently, two parabolic trough power plants are operational in Iran: Yazd integrated solar combined cycle power station-17 MW and Shiraz solar power plant-0.25 MW. A 20MW plant is also functional in Morocco, and a 5-MW plant is operational in Italy **[40].**

2.2 Solar Power Tower:

At first, these power towers used water-working fluid, but today, solar power towers in the USA use molten nitrate salt instead of water. The reason for the same is the desirable characteristics like its non-toxic, non-flammable nature, and it's better to heat storage properties than water. Solar towers are very cost-effective when they get utilized for 50 MW -100 MW power. However, these systems require the maximum area per energy generated unit and more water requirement than other CSP/CST technologies. Optical characteristics of the heliostat, mirror cleanliness, precision of the tracking system are some of the parameters that affect the efficiency of a solar power tower **[28].**

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:03/Issue:11/November-2021 Impact Factor- 6.752 www.irjmets.com

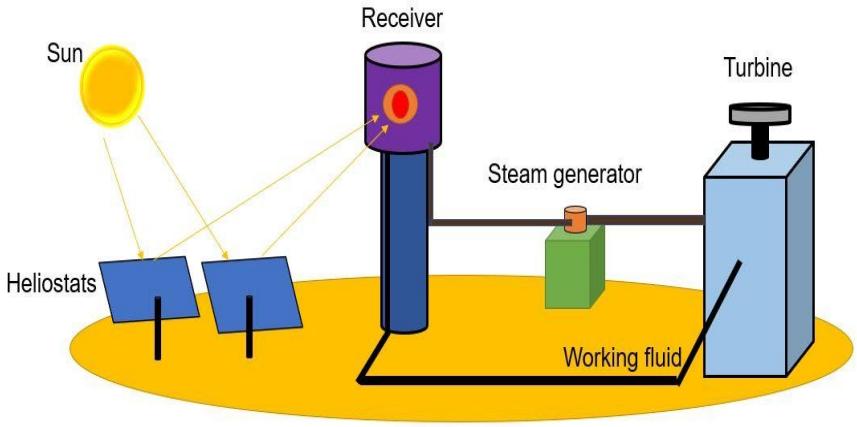


Figure 2: Schematic view of the Solar Tower System **[30]**

In a Solar-tower power plant, there are several heliostats (flat mirrors which concentrate solar irradiation on the receiver), a tower that houses a receiver machine with the working fluid, a generator that generates electrical energy. A receiver comprises the following: -ceramics and stabilized metals at high temperatures, -working fluid which could be water, -liquid sodium, -molten salt, -air. A steam generator converts mechanical energy into electrical energy (An electro-distribution grid then distributes the electricity generated). The exhaust steam from the turbine is then condensed into the condenser. This condensate is then pumped into the boiler machine, where it is again heated through the solar-receiver device **[29].**

There are currently five solar towers in operation in the world, according to the data found in **[40].** These include three tower-plant of Spain, 1 of the USA, and 1 of Germany. On the same website, there is a claim that 2-solar tower power plants are getting built worldwide: 1 in the USA and 1 in France **[40].** There are three CSP/CST plants in Spain: Planta Solar-20 and Gema-solar **[40].** There are also two solar tower power plants in the United States: Sierra Sun-Tower and Jülich Solar-Tower **[40].**

2.3 Parabolic dish system:

This dish system consists of parabola-shaped-reflectors, which look like a dish; a Stirling engine is also placed in the center and a generator machine for the electrical energy generation. The Stirling engine receives irradiation from the Sun during the daytime through the solar dishes which concentrate these irradiations. In a parabolic dish with (the concentration ratio of 2000), the focus reaches a temperature of 700◦C with working-fluid pressure of 200 bar **[32].** CSP/CST Technologies, specifically parabolic dish technology, is still in its early stages of development [33]. Currently, Parabolic-trough systems are the most dominant and mature technology for CSP/CST. Power-Towers follow these. However, Linear-Fresnel and Parabolic-dishes are in their early stages of development. Solar power towers worldwide have a capacity of 70-MW, while linear Fresnel systems have 31-MW in Spain and 4-MW in Australia [34]. Even though Parabolic-troughs or Tower-power plants have high efficiency, but the costs of electricity generation through Parabolic-dish are pretty high **[34].** A global parabolic dish with a capacity of 1.5-MW was installed in Arizona in 2010. Parabolic-Dish installed capacity was increased to 3-MW in 2013, with additional plants in Utah, and prototype dish-engines are currently operating in Arizona, Nevada, Colorado, and Spain **[34,35].**

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:03/Issue:11/November-2021 Impact Factor- 6.752 www.irjmets.com

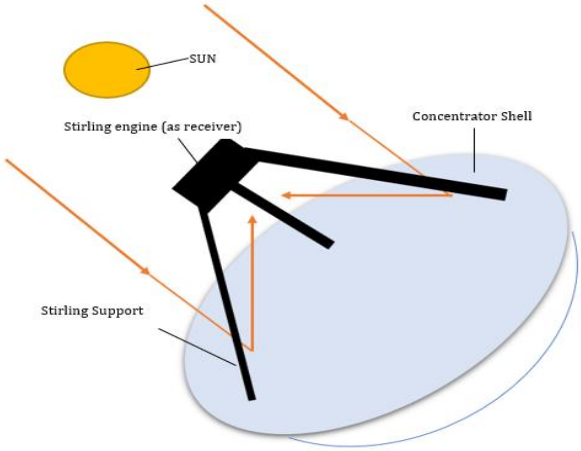


Figure 3: Schematic view of the Parabolic Dish System **[31]**

2.4 Fresnel reflectors:

With Fresnel-reflectors, solar-thermal power plants generate electricity by concentrating solar irradiation over cylindrical-parabolic reflectors, steam turbines, and generators **[37].** In the daytime, Fresnel-reflectors are directed towards the Sun automatically. This reflects the sunlight to the cylindrical-parabolic reflector, in the focus of which running-water long tubes are acting as the receiver system **[36].** Upon the reflection of the sunlight, water in these tubes evaporates and enters the steam turbine under pressure, which initiates the electricity generation in the electrical generator machine. The linear Fresnel reflector is a solar CST/CSP with a concentration ratio between 10-50, composed of segmented primary mirrors, placed close to the ground **[39].**

Currently, there are three active power plants with Fresnel-reflectors in the United States, Spain, and Australia based on the data found in **[40].** Active CSP plants are Kimberlina-5 MW plant in the United States, Puerto Errado 1 -1.4 MW plant in Spain and Liddell Power Station-2 MW plant in Australia. According to the same data, Spain is building the first power plant with Fresnel reflectors worldwide.

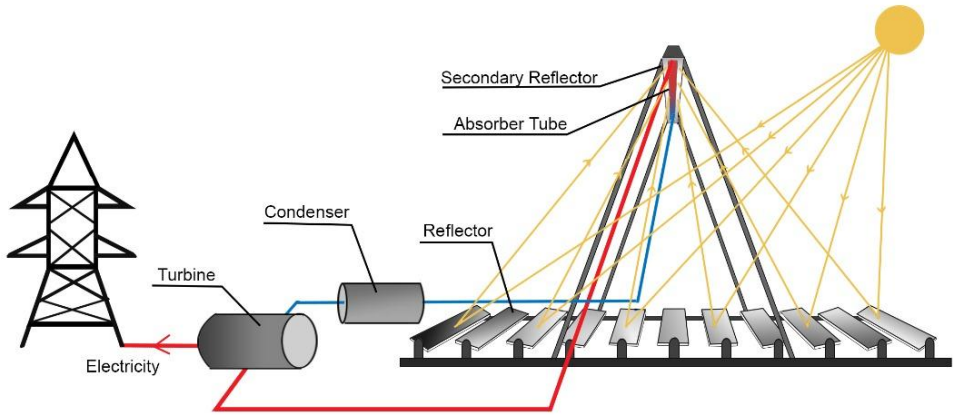


Figure 4: Schematic view of the CST/CSP plant with Fresnel reflectors **[38]**

II. LITERATURE REVIEW

Purohit et al. (2010) **[1]** evaluated solar concentrating power generation in India. For analyzing the techno-economic feasibility of concentrating-solar power technologies, they used two projects named PS-10 and ANDASOL-1 as a reference for this evaluation. They simulated these two systems in different locations. Results

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:03/Issue:11/November-2021 Impact Factor- 6.752 www.irjmets.com

showed that the site with an annual direct solar radiation of more than 1800 kWh per square meter is more suited for CSP projects installations.

Zhang et al. (2013) **[2]** reviewed Concentrated solar power plants in the year 2013. This paper briefly examines concentrated solar power technologies and the advantages of Solar tower collectors(STC). The results show the true potential of CSP in general and also define the design background of solar power collectors plants.

Soni et al. (2014) **[3]** carried out a techno-economic parametric assessment of solar power in India. This paper mainly focused on the required parameters for commissioning solar power projects in the Indian subcontinent using Concentrating Solar Power (CSP) and solar photovoltaic (PV). They classified the parameters obtained under both technologies and, at last, categorized them into technical and economic parameters. They also surveyed to understand these perceptions in the solar energy field and rank these 41 parameters. At last, the results indicated that investment cost and site location are the primary economic and technical factors in the area of solar energy.

Dunham et al. (2014) **[4]** provided a review on High-efficiency thermodynamic power cycles with their application to concentrated solar power systems, which mainly focused on high-efficiency single and combined processes. In this paper, they compared an estimate of an integrated receiver and power cycle operating Temperature for different cycles mentioned with the traditional approach of optimization based on Carnot efficiency. At last, results indicated that the conventional method is generally not efficient, which is used for optimizing the inlet temperatures and leads to a 1000C temperature shift from the Carnot cycle under any circumstances.

Sharma C. et al. (2015) **[5]** published a paper on India's solar thermal power generation assessment. Evaluating solar power for producing thermal energy and the selection of land for this work is crucial for implementing and designing appropriate measures. This paper presented a detailed analysis for calculating the potential of solar thermal power generation in India. The results conclude that for all the wastelands with 4m/s or more wind speeds for wind power, the estimated energy generation potential for a threshold DNI value of 1799 kWh/m2 is around 749 GW and 228 GW for a threshold DNI value of 2000 kWh/m2. At last, results obtained could be used for site selection for CSP plants.

Baharoon et al. (2015) **[6]** researched the historical development of concentrating solar power technologies to generate clean electricity efficiently. The conventional ways for generating electricity face two significant problems, Global warming, and depletion of fossil fuel reserves, which makes renewable technologies a better option to solve not present but future electricity problems also. This paper reviewed the historical evolution of CSP technologies to generate electricity, and all the current worldwide projects of CSP technologies were presented. It was found that the use of PTC technology in CSP was 95.7% and decreased to 73.4%. In contrast, the utilization of Tower power, Fresnel collector, and Stirling dish technologies in the operational projects were 2.24%, 2.07%, and 0%, respectively, and have increased to 20.82%,5.74%, and 0.052% for the projects which are under construction.

Otieno et al. (2016) **[7]** analyzed vital social and environmental risks in the development of concentrated solar power projects. The purpose of this research is to set out an investigation on the social and ecological risks in the development of CSP technologies and rank these risks according to their severity. And it was found that the risk of hampering local water resources was the maximum danger formerly and later deadening with a notch of moderate-high and reasonable, respectively.

Lorenzin et al. (2016**) [8]** provided a review on applying liquid metal as heat transfer fluid (HTF) in CSP/CST technologies. Among the next generation of CSP technologies, Solar Central Receiver Systems generally operates at about 1000 degrees Celsius to increase efficiency and reduce costs. So, to achieve these high temperatures, liquid metals can be used as HTF instead of conventional fluids like water, molten salts. In this paper, five liquid metals have been selected, and their compatibility with structural materials and integration with CFD software has been analyzed and discussed.

Naik et al. (2017) **[9]** researched the medium-temperature application of concentrated solar thermal technology. The authors illustrated the development and design of CSP/CST technology and its development in India. This paper also gives information about the need for these technologies, market potential, and barriers of

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:03/Issue:11/November-2021 Impact Factor- 6.752 www.irjmets.com

current CSP/CST technical installations, which are used for the generation of moderate temperatures about 80-249 °C.

Pelay et al. (2017) **[10]** reviewed energy storage systems for CSP/CST plants. This paper provided a comprehensive study of CSP/CST plants, their thermal energy storage systems(TES), and methods to improve heat and mass transfers. At last, this paper also proposed several new concepts for integrating TES systems and discussed various aspects of CSP technology with development around the world.

Dowling et al. (2017) **[11]** published a paper on the economic assessment of concentrated solar power technologies. According to this paper, the economic evaluations of CSP technologies have two dominant assessment methods. Many studies are grounded on the Levelized Costs Of Electricity (LCOE), but only a few consider the electricity market and time-varying meteorological conditions. At last, this research was towards the increasing reliability of economic assessments and to attenuate discrepancies.

Simsek et al. (2018) **[12]** researched Sustainability evaluation of Concentrated Solar Power (CSP) projects under the Clean Development Mechanism (CDM) by using Multi-Criteria Decision Method (MCDM). The key objective of this paper was to suggest a framework and help the decision-makers evaluate the sustainability of CDM technologies. In addition to that, this paper also shows the review of multi-criteria decision methods and sustainability indicators under the five factors: social, environmental, risk, technical, and economic factors, which is done using the MAUT method.

Mohammadia et al. (2019) **[13]** published a review on Hybrid concentrated solar power (CSP) desalination systems. This article provides an overview of desalination systems, studies on hybrid projects based on freshwater generation, and cogeneration of power and freshwater. It also proves that there are many ways to hybridize CSPs with desalination systems. Apart from that, they discussed potential ways to accelerate the commercialization of plants by cost minimization, maximization of thermal efficiency, and developing novel hybrid schemes.

Kumar et al. (2020) **[14]** researched the financial viability assessment of concentrated solar power technologies in India. For this, they studied the economic viability of three different CSP technologies, i.e., LFR, SPT, and PTC for 50MW for at least ten different locations in other regions with different climatic conditions. And results concluded that Jodhpur is one of the best-suited sites for the installation of the CSP/CST project, and A 50 MW Solar Tower CSP Plant has an LCOE of around 9.8 ₹/kWh, PBP-2.20 years, and NPV of ₹2626.68 million **[14],** which was found to best out of all selected locations.

Elbeh et al. (2021) **[15]** researched analyzing and optimizing concentrated solar power plants for application in desert regions. This paper reviews the CST technology for designing and optimizing CSP tower plants in areas with an arid climate. For this, two software, SolarPILOT and System Advisor Model, are used to analyze and optimize the whole CST project and its revenue. The outcomes confirmed that the feasibility of a CSP plant on 0.44 km2 of a solar field with 2736 heliostats produces at least 8 MWe power with 10 hours of thermal storage and hybrid steam condensing system.

III. CSP IN WORLD

It is evident from time and history that electricity production from fossil fuel burning (including coal, petroleum, natural gas.) creates many issues such as the sustainability or the future availability of these fuels. Besides this, burning fossils cause environmental problems like increasing the concentration of harmful gases produced from the burning, disrupting the environment's natural equilibrium. Therefore, to reduce the dependency on this disruption generating fossil fuels, the world is now shifting towards the use of renewable technologies, which uses natural fuels, such as solar energy, for electricity production, as negligence of these issues could lead to the occurrence of any global catastrophe **[19].** The term, Concentrated Solar Technology (CST) refers to the technology which involves the conversion of solar energy to electrical energy(electricity) by concentrating the sun irradiation, and CST or CSP plants (Concentrated Solar Power-Plant) uses this technology. Solar concentrators, steam turbines, and electricity generators make up CSP plants. In addition to parabolic dishes, heliostats, and troughs, Fresnel reflectors can also be used as solar concentrators **[20].** Concentrating solar power technology or CST focuses on the radiations from the Sun. Hence, it can drive the traditional gas and steam power-plant cycles or, in simple words, can solve mentioned problems. This concentrating solar power technology has been known for more than a hundred years.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:03/Issue:11/November-2021 Impact Factor- 6.752 www.irjmets.com

Moreover, we have advancements in technology with some major market-introduction programs triggered worldwide activities **[21].** Water, air, and solar energy conversion are used in low-temperature flat collectors (where Temperature <100°C). In moderate temperatures (100°C < Temperature < 400°C), collectors with concentrators and vacuum collectors, and in high temperatures (400°C < Temperature <4000°C), the solar-furnaces in CSP plants are used **[22,23].**

Potential (at world-level) of electricity generation through CST power-plants: NASA's DNI dataset (Direct Normal Irradiation) is used worldwide to understand the potential of electricity generation from CST **[24].** This dataset is based on twenty-two years of data collection; its characteristics include spatial-resolution of around 99 km (which is considered enough for understanding the potential of CSP plants on the world level), land cover like the non-permanent and permanent forest, waters, swamp, areas for agriculture, shifting sand-dune, which includes a sideline of 10 km; other than this it also contains the ice-glaciers, salt-pans, settlement, airports, gas & oilfields, protected zones/areas, restricted areas and, slope greater than 2 percent which is used as a site-exclusion criterion, these criteria are applied to identify the areas which are not suitable for construction of CST power plants. When DNI datasets and site exclusion maps are put together, one can obtain the global map of annual direct normal irradiance for potential CSP sites given. The world potential of CSP amounts to almost 3000000 TWhyr−1, which is about a hundred and fifty times the present world electricity consumption of 21000 TWhyr−1 **[24].**

IV. CSP IN INDIA

There are more than 1.3 billion people in India, and the country has a massive energy demand each day; to meet this need, it depends heavily on coal and oil-based thermal power plants. Because of the exhaustive nature of these fuels, India is investing more in solar energy to fulfill its energy requirement. According to a report by the Ministry of Statistics and Program Implementation India (MOSPI), renewable energy power generation has reached 78316.44 MW as of 2019, out of which Solar energy constitutes about 36% of it, which accounts for 28.18 GW. The Indian government in January 2010 inaugurated the Jawahar Lal Nehru National Solar Mission with a target of setting up 100 GW solar PV by the year 2022. The program was to structure India's growth as a global solar leader by deploying various policies and setting up resources to achieve the target.

The mission was to be completed in three phases; 1st Phase from (2010-2013) targeted to develop a total grid connection of 1000 MW covering rooftop systems and in a total of 200 MW off-grid solar applications PV. 2nd Phase (2014-2017) targeted a cumulative 4000-10000 MW of grid-connected solar PV, including rooftop and 1000 MW of off-grid solar applications PV. Other solar infrastructure development included 40 GW solar PV construction through solar parks and Ultra Mega solar power projects. 3rd Phase (2017-2022) targeted cumulative 100000MW on grid-connected solar PV and 2000 MW off-grid connected solar applications PV **[43].** India has a vast potential for solar energy because of its geography. Currently, India's primary source of energy generation is the conventional power plants that operate on fossil fuels. According to a report in 2019, coal contributed 75% of the total energy supply in India, and it also accounts for 50% of the entire installed capacity. Natural gas and oil are next to coal in terms of energy generation in India; in February 2021 total installed natural gas capacity was around 24 thousand Megawatt **[46].** The current scenario of climate change and lack of availability of fossil fuels in the future has necessitated India to look for alternatives to renewable sources of energy.

S.NO	Project Name	Location	Area(acres)	Technology	Capacity (MW)
1	Dhursar Project	Jaisalmer, Rajasthan	840	Linear Fresnel	125
2	Godawari Project	Jaisalmer, Rajasthan	370	Parabolic Trough	50
3	Megha Solar plant	Anantpur, Andhra Pradesh	600	Parabolic Trough	50

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:03/Issue:11/November-2021 Impact Factor- 6.752 www.irjmets.com

4	ACME solar tower	Bikaner, Rajasthan	12	Solar Tower	2.5
5	National Solar Thermal Power Facility	Gurugram, Haryana	NA	Parabolic Trough	1
6	Diwakar Project	Jaisalmer, Rajasthan	NA	Parabolic Trough	100
7	KVK Energy Solar Project	Jaisalmer, Rajasthan	NA	Parabolic Trough	100
8	Abhijeet Solar Project	Jaisalmer, Rajasthan	388	Parabolic Trough	50
9	Gujarat Solar One	Kutch, Gujarat	NA	Parabolic Trough	25
10	Dadri ISCC Plant	Dadri, Uttar Pradesh	NA	Linear Fresnel	14

Table 1: Details of CSP projects in India **[42]**

Under JNNSM, it was targeted to have 470 MW of CSP projects in India by the end of Phase – I, but by 2018 only 228.5 MW were operational. CSP projects have underperformed in Phase – I and Phase – II **[42].** Several reasons can be pointed out that led to the failure of the project. The first concern was the lack of data on DNI in certain parts of India. Accurate DNI data is critical for project planning. In the absence of accurate DNI data, the project planner relied on NASA's solar radiation data, 10-15% higher than actual DNI data. The Indian investors were also unfamiliar with CSP due to the lack of information on the feasibility of CSP in India during Phase I of JNNSM. The above issues led to investors' distrust in the project, which adversely affected the bidding process in Phase - I. Another barrier was the non-availability of any policy regarding the CSP. As a result of the government not setting specific tariffs for the project, the project's overall cost decreased without considering the upcoming uncertainties and challenges with CSP. This shows the inexperience of the developers in the project. CSP faced tough competition from SPV, which captured the attention of the Indian government, was heavily promoted, and massive investments were made in it. SPV is cheap, its operational and maintenance cost is meager. Also, investors were familiar with SPV's performance in the market. All of this gave SPV an edge over the newly introduced technology CSP. CSP being an emerging technology, Indian developers did not have any prior knowledge of CSP. Thus, it was required to import the technology from outside. Continuously increasing custom duties for equipment, raw materials, and processed materials posed a barrier to the completion of the project **[47].**

V. FUTURE SCOPE IN INDIA

Many studies and developments have been made in CSP to increase its efficiency, decrease its overall cost, and make it more feasible for different countries in the world. Hybridization, also known as cogeneration, is a recent development in the field of CSP that is in use by many countries around the world. It is possible by combining CSP with other fuel-based power plants. CSP is highly acknowledged globally because of its lower cost, higher efficiency, and more easy energy storage technology than SPV. CSPs are integrated with conventional power plants to increase their efficiency and lower the overall cost of the power plant. The arrangement reduces the emission from the plant and benefits the efficiency of the plant. The integration is bidirectional means we can hybridize CSP in a conventional power plant where most of the electricity generated is through the conventional one and CSP is used to superheat steam, or we can have CSP as the primary energy generation unit. The Indian energy sector is dominated by traditional plants, which means hybridization is crucial for reducing the country's dependency on conventional power plants. CSP can be combined with a wide range of energy sources, including thermal, biofuels, and geothermal. There are two methods of hybridizing: one by thermal integration where water is superheated by CSP and energy is generated using conventional, second one by thermos chemical integration **[44].**

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:03/Issue:11/November-2021 Impact Factor- 6.752 www.irjmets.com

Many CSP experts believe that hybridization can answer the problem of direct normal irradiance (DNI) and variable Indian climate. The average DNI in India is between 1750-1850 Kwh/m^2, which is very low compared to other countries with installed CSP technology. Spain (2100-2200), the US (2400-2700), South Africa (2500-2800), and Chile (2800-3400) are the countries with greater DNI than India. Due to the lower DNI, India requires larger solar field installations to generate the same energy as countries with higher DNI. During monsoon, days are generally rainy and cloudy, accounting for a reduced DNI of 4 Kwh/m^2, which certainly is not enough to generate sufficient steam that can run the turbine and generate electricity. Hybridization will help to provide support to these CSP installations to overcome these irregularities. Not only improved DNI, but some experts believe that with the given DNI, CSP technology in India by hybridization with conventional power plants can achieve a better output power quality. So, a fossil fuel-based power plant with some part of CSP installed is the new mantra for the revival and development of CSP in India.

Another benefit of installing hybrid CSP is the increased efficiency of the overall plant. Higher operating temperatures above 470 C were previously believed to be required to achieve higher efficiency from CSP. However, because of the low DNI and lack of fossil fuel backup, better efficiency for CSP in Phase - I of the JLNSS would not have been possible. Some experts recommend that a hybrid project as the second SECI technology demonstration would be more appropriate than a tower with high temperatures **[45].** For a country like India, which is hugely dependent on fossil fuel-based energy plants to fulfill their energy requirements, it is necessary to identify cost-effective renewable energy technologies for large-scale power generation. Hybridization can offer India a potential solution to its intermittency issues and peak load shortages, and significant potential in the future. Hybridization of PV and CSP is another critical advancement to keep an eye out for. It could be a possible solution for the development of CSP in India. The hybrid system can provide stable thermal energy storage and provide dispatchable power to off-grid. The hybrid technology has increased the overall generating efficiency, better performances than their parent technologies, and lower cost. Recent developments in hybrid technology have provided methods like the PV-topping method and spectral beam splitting method (SBS) that aims at fully utilizing solar energy **[48].**

VI. CONCLUSION

The paper gives a detailed overview of CSP technology, its growing demand around the world, and the current scenario in India. CSP can pose a potential solution for India's cleaner and renewable energy crisis. Here are some main highlights of the above work.

- Despite the effort of the Indian government, CSP constitutes only a tiny percentage of energy generation in India that accounts for 228.5 MW as in 2018.
- Unavailability of DNI data, lack of background research of the technology, and less interest of investors in the technology are the primary reasons for India's restricted growth. On top of these issues, it faces tough competition from SPV in the market, which is currently a significant area of interest of the Indian government.
- Hybridization with existing coal plants, low-cost finance, e-loans, and reliable DNI data, if provided, can ensure successful penetration of CSP through the country.
- In the end, it can be concluded that despite several barriers, CSP has a bright future in India. The total installed capacity of CSP is forecasted to reach 1.3 GW by 2024.

VII. REFERENCES

[1] Purohit, I., & Purohit, P. (2010). Techno-economic evaluation of concentrating solar power generation in India. Energy Policy, 38(6), 3015–3029.

[2] Zhang, H. L., Baeyens, J., Degrève, J., & Cacères, G. (2013). Concentrated solar power plants: Review and design methodology. Renewable and Sustainable Energy Reviews, 22, 466–481.

[3] Soni, M. S., & Gakkhar, N. (2014). Techno-economic parametric assessment of solar power in India: A survey. Renewable and Sustainable Energy Reviews, 40, 326–334.

[4] Dunham, M. T., & Iverson, B. D. (2014). High-efficiency thermodynamic power cycles for concentrated solar power systems. Renewable and Sustainable Energy Reviews, 30, 758–770.

[5] Sharma, C., Sharma, A. K., Mullick, S. C., & Kandpal, T. C. (2015). Assessment of solar thermal power

generation potential in India. Renewable and Sustainable Energy Reviews, 42, 902–912.

[6] Baharoon, D. A., Rahman, H. A., Omar, W. Z. W., & Fadhl, S. O. (2015). Historical development of concentrating solar power technologies to generate clean electricity efficiently – A review. Renewable and Sustainable Energy Reviews, 41, 996–1027.

[7] Otieno, G. A., & Loosen, A. E. (2016). An analysis of key environmental and social risks in the development of concentrated solar power projects.

[8] Lorenzin, N., & Abánades, A. (2016). A review on the application of liquid metals as heat transfer fluid in Concentrated Solar Power technologies. International Journal of Hydrogen Energy, 41(17), 6990–6995

[9] Naik, H., Baredar, P., & Kumar, A. (2017). Medium temperature application of concentrated solar thermal technology: Indian perspective. Renewable and Sustainable Energy Reviews, 76, 369–378.

[10] Pelay, U., Luo, L., Fan, Y., Stitou, D., & Rood, M. (2017). Thermal energy storage systems for concentrated solar power plants. Renewable and Sustainable Energy Reviews, 79, 82–100.

[11] Dowling, A. W., Zheng, T., & Zavala, V. M. (2017). Economic assessment of concentrated solar power technologies: A review. Renewable and Sustainable Energy Reviews, 72, 1019–1032.

[12] Simsek, Y., Watts, D., & Escobar, R. (2018). Sustainability evaluation of Concentrated Solar Power (CSP) projects under Clean Development Mechanism (CDM) by using Multi Criteria Decision Method (MCDM). Renewable and Sustainable Energy Reviews, 93, 421–438.

[13] Mohammadi, K., Saghafifar, M., Ellingwood, K., & Powell, K. (2019). Hybrid concentrated solar power (CSP)-desalination systems: A review. Desalination, 468, 114083.

[14] Kumar, S., Agarwal, A., & Kumar, A. (2020). Financial viability assessment of concentrated solar power technologies under Indian climatic conditions. Sustainable Energy Technologies and Assessments, 100928.

[15] Elbeh, M. B., & Sleiti, A. K. (2021). Analysis and optimization of concentrated solar power plants for application in arid climate. Energy Science & Engineering, 9(6), 784–797.

[16] "Concentrating Solar Power, Global Outlook 09, Why Renewable Energy is Hot", published by Greenpeace International, SolarPACES and ESTELA, 2009. Available at: http://www.greenpeace.org/raw/content/international/press/re ports/concentrating-solar-power-2009.pdf

[17] Müller-Steinhagen, H. & Trieb, Franz. (2004). Concentrating solar power, - A review of the technology. Ingenia. 18. 43-50.

[18] M. Brenna, F. Foiadelli, M. Roscia and D. Zaninelli, "Evaluation of solar collector plant to contribute climate change mitigation," 2008 IEEE International Conference on Sustainable Energy Technologies, 2008, pp. 198-202, doi: 10.1109/ICSET.2008.4747001. ICSET 2008. 198 - 202. 10.1109/ICSET.2008.4747001.

[19] Ecotricity. The end of fossil fuel. Available from: http://www.ecotricity.co.uk/%20our-green-energy/energy-independence/the-end-of-fossil-fuels

[20] Pitz-Paal R. High-temperature solar concentrators. In: Blanco Galvez J, Malato Rodriguez S, editors. Solar energy conversion and photo energy systems. Encyclopedia of life support systems (EOLSS). Oxford, UK: Developed under the Auspices of the UNESCO, Eolss Publishers; 2007.

[21] Trieb F, Schillings C, O'Sullivan M, Pregger T, Hoyer-Klick C. "Global potential of concentrating solar power". In: SolarPaces conference. Berlin; 2009. Google Scholar

[22] Al-Kayiem, H., & Mohammad, S. (2019). Potential of Renewable Energy Resources with an Emphasis on Solar Power in Iraq: An Outlook. Resources, 8(1), 42. https://doi.org/10.3390/resources8010042

[23] Fernández-García, A., Zarza, E., Valenzuela, L., & Pérez, M. (2010). Parabolic-trough solar collectors and their applications. Renewable and Sustainable Energy Reviews, 14(7), 1695–1721. https://doi.org/10.1016/j.rser.2010.03.012.

[24] Müller-Steinhagen, H. (2013). Concentrating solar thermal power. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 371(1996), 20110433. https://doi.org/10.1098/rsta.2011.0433

[25] Wasim Jamshed, Suriya Uma Devi S, Rabia Safdar, Fares Redouane, Kottakkaran Sooppy Nisar &

Mohamed R. Eid (2021) Comprehensive analysis on copper-iron (II, III)/oxide-engine oil Casson nanofluid flowing and thermal features in parabolic trough solar collector, Journal of Taibah University for Science, 15:1, 619-636, DOI: 10.1080/16583655.2021.1996114

[26] Trieb F, Schillings C, O'Sullivan M, Pregger T, Hoyer-Klick C. "Global potential of concentrating solar power". In: SolarPaces conference. Berlin; 2009. Google Scholar

[27] Yılmaz, İ. H., Mwesigye, A., & Göksu, T. T. (2020). Enhancing the overall thermal performance of a large aperture parabolic trough solar collector using wire coil inserts. Sustainable Energy Technologies and Assessments, 39, [100696]. https://doi.org/10.1016/j.seta.2020.100696

[28] Pavlović, T. M., Radonjić, I. S., Milosavljević, D. D., & Pantić, L. S. (2012). A review of concentrating solar power plants in the world and their potential use in Serbia. Renewable and Sustainable Energy Reviews, 16(6), 3891–3902. https://doi.org/10.1016/j.rser.2012.03.042

[29] ZAVOICO, ALEXIS B. (2001). Solar Power Tower design basis document, revision 0. https://doi.org/10.2172/786629

[30] Cruz, N. C., Salhi, S., Redondo, J. L., ÁLvarez, J. D., Berenguel, M., & Ortigosa, P. M. (2018). Design of a parallel genetic algorithm for continuous and pattern-free heliostat field optimization. The Journal of Supercomputing, 75(3), 1268–1283. https://doi.org/10.1007/s11227-018-2404-8

[31] Hafez, A., Soliman, A., El-Metwally, K., & Ismail, I. (2016). Solar parabolic dish Stirling engine system design, simulation, and thermal analysis. Energy Conversion and Management, 126, 60–75. https://doi.org/10.1016/j.enconman.2016.07.067

[32] Pheng, L. G., Affandi, R., Ab Ghani, M. R., Gan, C. K., Jano, Z., & Sutikno, T. (2014). A Review of Parabolic Dish-Stirling Engine System Based on Concentrating Solar Power. TELKOMNIKA (Telecommunication Computing Electronics and Control), 12(4), 1142. https://doi.org/10.12928/telkomnika.v12i4.1132

[33] Buck, R, Heller, P, & Koch, H. Receiver development for a Dish-Brayton system. United States. (1996).

[34] Irena. Renewable Energy Technologies: Cost Analysis Series. (2012). https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2013/Renewable_Power_Generation_Costs_in_2012_summary.pdf?la=en&hash=548B1D4A7BEAF616A19B26D8DF07011A8B8F49E7

[35] Fernández-García, A. & Zarza, E. & Valenzuela, L. & Pérez, M., 2010. "Parabolic-trough solar collectors and their applications," Renewable and Sustainable Energy Reviews, Elsevier, vol. 14(7), pages 1695-1721, September.

[36] Bellos, E. (2019). Progress in the design and the applications of linear Fresnel reflectors – A critical review. Thermal Science and Engineering Progress, 10, 112–137. https://doi.org/10.1016/j.tsep.2019.01.014

[37] Abbas, R., Muñoz-Antón, J., Valdés, M., & Martínez-Val, J. (2013). High concentration linear Fresnel reflectors. Energy Conversion and Management, 72, 60–68. https://doi.org/10.1016/j.enconman.2013.01.039

[38] Soomro, Mengal, Memon, Khan, Shafiq, & Mirjat. (2019). Performance and Economic Analysis of Concentrated Solar Power Generation for Pakistan. Processes, 7(9), 575. https://doi.org/10.3390/pr7090575

[39] Lovegrove, K., & Stein, W. (2012). Concentrating Solar Power Technology: Principles, Developments and Applications (Woodhead Publishing Series in Energy) (1st ed.). Woodhead Publishing.

[40] List of solar thermal power stations. (2021, November 23). Retrieved November 23, 2021, from https://en.wikipedia.org/wiki/List_of_solar_thermal_power_stations

[41] National Statistical Office Ministry of Statistics and Programme Implementation Government of India. (2020b). Energy Statistics 2020 by Ministry of Statistics and Program Implementation India (MOSPI).

[42] Santamarta, J. (2019). Concentrated solar power had a global total installed capacity of 6,451 MW in 2019. HELIOSCSP. https://helioscsp.com/concentrated-solar-power-had-a-global-total-installed-capacity-of-6451-mw-in-2019/

[43] Jawaharlal Nehru National Solar Mission (Phase I, II and III) – Policies. (2021). IEA. https://www.iea.org/policies/4916-jawaharlal-nehru-national-solar-mission-phase-i-ii-and-iii

[44] Concentrated Solar Power (CSP): A Potential Solution for India | CAG. (2020). https://www.cag.org.in/newsletters/public-newsense/concentrated-solar-power-csp-potential-

solution-india

[45] Llamas, D. (2016). Concentrated solar power (CSP) in India: An outlook to 2024. HELIOSCSP. https://helioscsp.com/concentrated-solar-power-csp-in-india-an-outlook-to-2024/

[46] S Deshmukh, S., Mishra, O., Agrawal, V., & Anwar, K. (2019). Hybrid-CSP in India: Technological and Economic Aspects. International Journal of Renewable Energy Research. Published. https://www.ijrer.org/ijrer/index.php/ijrer/article/download/8561/pdf

[47] Hybrid: The new mantra for CSP revival in India | Reuters Events | Renewables. (2015). https://www.reutersevents.com/renewables/csp-today/technology/hybrid-new-mantra-csp-revival-india

[48] Jaganmohan, M. (2021, June 14). Fossil fuel industry in India - statistics & facts. Statista. https://www.statista.com/topics/5747/fossil-fuel-industry-in-india/#dossierKeyfigures

[49] Shirole, A., Wagh, M., & Kulkarni, V. (2021). Thermal Performance Comparison of Parabolic Trough Collector (PTC) Using Various Nanofluids. International Journal of Renewable Energy Development, 10(4), 875–889. https://doi.org/10.14710/ijred.2021.33801

[50] Ju, X., Xu, C., Han, X., Zhang, H., Wei, G., & Chen, L. (2017). Recent advances in the PV-CSP hybrid solar power technology. SOLARPACES 2016: International Conference on Concentrating Solar Power and Chemical Energy Systems. Published. https://doi.org/10.1063/1.4984480

2021 IEEE International Conference on Engineering, Technology & Education (TALE) | 978-1-6654-3687-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/TALE52509.2021.9678531

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 31,2022 at 08:51:17 UTC from IEEE Xplore. Restrictions apply.

immersive, an effective SGs are required. This in turn, maintain a correct balance between learning and gaming mechanics. At this point, the *Learning Mechanics-Game Mechanics* (LM-GM) framework [11], [12] can be used as a possible solution. Use of LM-GM framework helps us balance the learning and gaming objectives.We can easily analyse the linkages amongst the selected elements from this framework. The LM-GM tool can also facilitate a critical review of already designed games and provide clues for achieving a better balance between pedagogical objectives with playfulness of the game.

While the dynamics of game should ensure a balance between Learning mechanics and Game mechanics and thus ensure achievement of pedagogical goal of learning while still providing a fun and flow to the player; the litmus test of any serious game after its design is assessment. The assessment has various components like user experience, engagement of the player by way of immersion in game play and its utility in terms of teaching the core concepts. Learning outcome (LO) is a difficult construct but very crucial to rate a serious game. It helps us to comment on utility of the game in achieving our stated objectives [13]. There is lack of methods to assess the effectiveness of SG [14] in terms of learning outcome. This is primarily because of lack of concrete methodologies available for such intangible measurement. Questionnaires have been used extensively to seek suggestions from players to improve SGs [15]. A potential source for assessing the learning outcome can be the captured player data as the game progresses [16]. Many authors have reported use of this approach in assessing LO [17]. We attempt to use the approaches suggested by [18] to achieve an objective assessment of LO.

The major contributions of our paper are:

1) We attempt to design our games with help from LM-GM framework.
2) We evaluate the effectiveness of the four games using both objective measure; by capturing the player data using Unity analytics and subjective measure;by embedding a questionnaire to elicit user experience (UX).

The organization of our paper is described below. Section II discusses the Open Web Application Security Project (OWASP)-based concepts and brief description of our four games. Section III presents the design considerations of the four OWASP-based games using LM-GM framework. Section IV discusses the way to assess the Learning Outcome (LO) of the four games. Section V addresses the limitations of our research. Finally, the concluding remarks and future directions are placed in Section VI.

II. Open Web Application Security Project (OWASP)

Secure web application development needs suitable guidelines to guide the process. Without topical and relevant guidelines,we may not be able to design useful systems [19]. Stuart *et. al.*, [20] have been addressing this issue in a very comprehensive manner to cover system, network, and software security. A systematic approach has been defined by them to guard against online cyber-attacks. Darothy [21] has highlighted the importance of well defined tools for the defence of online systems.

Thus we see that web security has been of concern for more than two decades and with increased uptake of online systems to support e-commerce, e-governance and e-learning the vulnerabilities towards cyber threats have increased significantly. To minimize the impact of cyber threats on the Web Application, the testing methodologies like Penetration testing [22], *etc.*, are used. The testing methodologies follow a standard approach that helps in increasing the security of the Web Application and with thus hardened Web Application, the vulnerabilities of a cyber threat reduce. In this paper, the OWASP-identified vulnerabilities for the Web Application are considered. OWASP [23] is a non-profit foundation that works on the basis of improving the security of software. Apart from improving the software security, OWASP has been listing Top 10 vulnerabilities of web applications since 2003. The professional approach adopted by OWASP makes its recommendations for countering the web application threats a very useful source of content while training the web application developers and this is the rationale for design of serious games by us based on the latest OWASP recommendations.

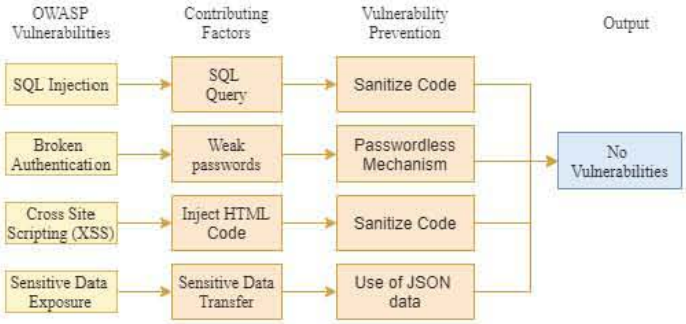


Fig. 1. OWASP Vulnerabilities [22]

The Top 10 vulnerabilities of OWASP are shown in Table I [24]. From the Top 10 vulnerabilities of the OWASP, the four vulnerabilities are illustrated in Fig. 1 [22].

In Fig 1, the 'OWASP Vulnerabilities' illustrate the four vulnerabilities that are considered in our paper. The 'Contributing Factors' in Fig. 1 provides a way to perform the cyber attack on the Web Application. The 'OWASP Prevention' provides the solution to minimize the impact of the cyber-attack on the Web Application. Finally, at the 'Output' in Fig. 1, the impact of the OWASP vulnerabilities is reduced on the Web Application.

The four vulnerabilities of OWASP are discussed below:

A. SQL Injection

In an SQL Injection attack [24], the attackers inject malicious SQL queries (or malicious programs) on the Web Application. After implementing the malicious SQL queries on the Web Application, the attacker accesses the user's confidential data in an unauthorized way. As a result of the

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 31,2022 at 08:51:17 UTC from IEEE Xplore. Restrictions apply.

TABLE I
OWASP TOP 10 VULNERABILITIES [24]

Vulnerabilities
SQL Injection.
Broken Authentication & Session Management.
Cross Site Scripting (XSS).
XML External Entities.
Broken Access Control.
Sensitive Data Exposure.
Security Misconfiguration.
Insecure Deserialization.
Using Components with known vulnerabilities.
Insufficient Logging & Monitoring.

SQL Injection, the data breach occurs. Hence, the solution is needed to reduce the impact of the SQL Injection attack on the Web Application.

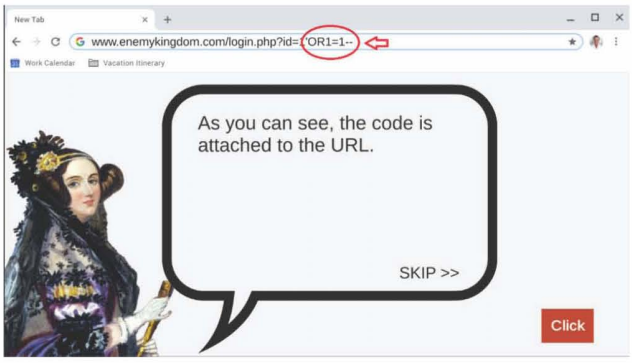


Fig. 2. SQL Injection Attack

As shown in the above Fig. 2, on the execution of the URL, the Web Application becomes vulnerable to the SQL Injection attack. Some of the real-world examples of the SQL Injection attacks are:

- In 2014[1], the Tesla website faced a SQL Injection attack. With the SQL Injection attack performed by the attacker, the attacker gains administrative privileges and steal user data from the vulnerable Web Application.
- In 2019[1], the Fortnite website also faced the attack of SQL Injection. With the SQL Injection attack performed by the attacker, the attacker gains access to the user accounts in an unauthorized way.

The solution is to strip special characters from the SQL queries or improve the validation and sanitization of the Web-based information. The various forms of SQL Injection attacks are Time Based SQL Injection, Error-based SQL Injection, *etc.* Each of the SQL Injection attacks models in the game is in the form of levels. The levels enable the players to understand SQL Injection attacks. With the proper understanding of the SQL Injection attacks, the vulnerability of the Web Applications reduces. With help from this game the player learns to implement/prevent the SQL Injection attacks.

B. Broken Authentication

A Broken Authentication attack [24] allows the attacker to access the user account, compromised passwords, *etc.* As the attacker accesses the user's data in an unauthorized way, the authenticity of the user's data is 'broken'. The attacker performs various attacks to break the authenticity of the user's account are password spraying attacks[2], brute force attacks, *etc.*

- In the brute force attack, the attacker uses the trial and error method to log in to the Web Application.
- The attacker performs the password spraying attack only when weak passwords like password123, query123! *etc.*, are used by the user.

Hence, there must be some solution to reduce the impact of the broken authentication attack on the Web Application. The real-world examples of the password spraying attack are:

- On 23rd September 2017[3], the Shopify website faced the problem of the password spraying attack. The vulnerability of the Shopify Website was exploited when the attacker logged in to the Shopify store in an unauthorized way. The impact of the password spraying attack on the Shopify Website was: (a) user was unable to log in to the user account. (b) user was unable to advertise the product from the Shopify website. In the end, the attacker received a $20,000 reward.

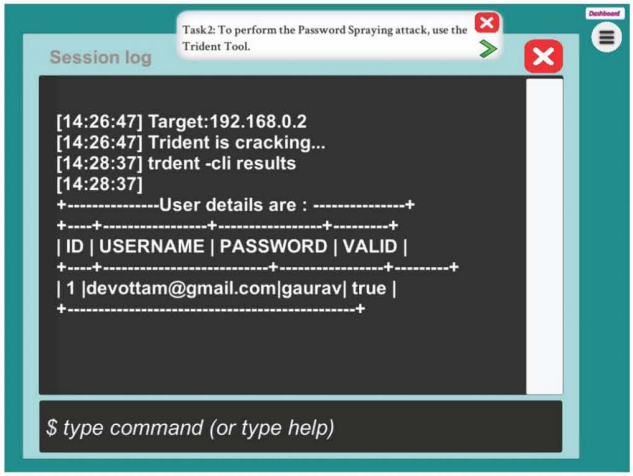


Fig. 3. Password Spraying Attack

To reduce the impact of the password spraying attack on the Web Application, solution like Double Octopus Passwordless

[1]https://www.neuralegion.com/blog/sql-injection-attack/.

[2]https://doubleoctopus.com/security-wiki/threats-and-tools/password-spraying/amp/.

[3]https://hackerone.com/reports/270981.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 31,2022 at 08:51:17 UTC from IEEE Xplore. Restrictions apply.

platform[2], *etc.*, is used. The Double Octopus Passwordless platform allows the user to log in to the website without a password. The password spraying attack is modeled in the game with the Trident Tool[4]. The command used with the Trident Tool is in Fig. 3. From Fig. 3, the player learns the way to perform the password spraying attack.

C. Cross Site Scripting (XSS)

Cross-site scripting (XSS) [24] is a type of injection security attack in which an attacker injects malicious script on a vulnerable website. Cross-site scripting attacks happen when an untrusted source injects the malicious code into the Web Application. The real-world examples of the XSS attacks were:

- In January 2019, the attacker performed the XSS attack on the Fortnite website. The purpose of the XSS attack was to expose the data of the 200 million players through a breach point[5]. Hackers exploited a webpage vulnerability to employ an XSS attack when victims clicked on a link. XSS attacks inject malicious code into vulnerable web applications.

The prevention methods like input sanitization and encoding prevent the XSS attack from being performed on the Web Application:

- Input sanitization[6] describes cleansing and scrubbing user input to prevent it from jumping the fence and exploiting security holes.
- Encoding[7] is transforming data from one format into another format.

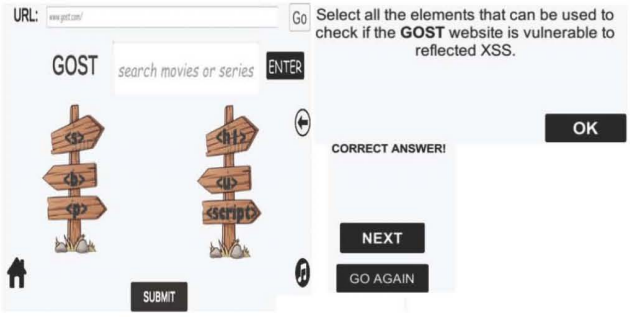


Fig. 4. Reflected XSS Attack

The three types of XSS attacks are: reflected[8], stored[8], and DOM XSS[9]. Reflected XSS attacks, also known as non-persistent attacks, occur when a malicious script is reflected off a web application to the victim's browser. Stored XSS, also known as persistent XSS, occurs when the user-supplied input is stored and then rendered within a web page. DOM XSS stands for Document Object Model-based Cross-site Scripting. A DOM-based XSS attack is possible if the web application writes data to the Document Object Model without proper sanitization.

Each of the three XSS attacks models in the game is in the form of levels. The levels enable the players to understand XSS attacks. With the proper understanding of the XSS attacks, the vulnerability of the Web Applications reduces. With help from this game the player learns to perform and prevent the XSS attack as shown in Fig. 4.

D. Sensitive Data Exposure

The Sensitive Data Exposure attack [24] allows the attacker to access the user's sensitive information from the Web Application in an unauthorized way. The Sensitive Data Exposure attack may occur due to weak encryption techniques, flaws in the software, or the user's information are uploaded to the erroneous database. As a result of the Sensitive Data Exposure attack on the Web Application, the data breach occurs. Hence, the solution is needed to prevent the Sensitive Data Exposure attack. The real-world examples of the Sensitive Data Exposure attack were:

- Google's Project Zero found an issue in Cloudflare's[10] edge servers that made it possible to dump memory potentially containing sensitive data, some of which were cached by search engines. The security bug was named Cloudbleed. Cloudflare had acknowledged the memory leak was started on 22^{nd} September 2016, and a private key between Cloudflare machines was leaked.

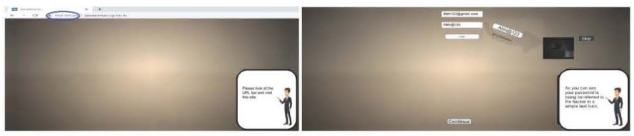


Fig. 5. Sensitive Data Exposure Attack

Hence, the Sensitive Data Exposure attack is prevented as: the first step is to stop the user from storing the sensitive information on the web. The sensitive information that is transmitted online must be processed, classified, or stored in any proper location. To understand the Sensitive Data Exposure, the gamification approach is followed with different levels. In the end of the game, the player learns to perform the Sensitive Data Exposure Attack as shown in Fig. 5.

III. Design Considerations

The four web-based games, *i.e.*, SQL Injection, Broken Authentication, XSS, and Sensitive Data Exposure are developed using Unity [25]. Unity is acting as our development platform and also a test bed to elicit objective player data using its 'Unity analytics' feature. Unity can be deployed on various current operating systems.

The purpose of the four web-based games is to teach the players about Web Application Security. The target audience

[4]https://hakin9.org/trident-automated-password-spraying-tool/.

[5]https://securityboulevard.com/2020/01/hacks-and-breaches-of-2019-a-year-in-review/.

[6]https://www.cloudways.com/blog/prevent-laravel-xss-exploits/.

[7]https://www.indusface.com/blog/xss-examples-prevent/.

[8]https://owasp.org/www-community/attacks/xss/#stored-and-reflected-xss-attacks.

[9]https://owasp.org/www-community/attacks/DOM_Based_XSS.

[10]https://bugs.chromium.org/p/project-zero/issues/detail?id=1139.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 31,2022 at 08:51:17 UTC from IEEE Xplore. Restrictions apply.

for the four web-based games is system administrators/entry-level IT users. The four games are web-based games.The four playable games are accessible on IIT Delhi website[11].

Now, we discuss the User Interface (UI) and User Experience (UX) for the four web-based games and describe in detail the adaptation of LM-GM framework to facilitate design of our games.

The key challenge in design of serious games is in maintaining a balance between the learning and gaming part. The User Interface (UI) and User Experience (UX) combine to create Game-play Experience (GX).User interface (UI) is anything through which user can interact. This can include the display screen, keyboard, mouse, and desktop display. This is how users interact with the game.User experience (UX) is the art of a game design mechanics to ensure that player interaction with an evolved game will be as immersive as possible. Both UI and UX are very important for the game and work together. A game with good UX and UI, also creates a good GX. With a good GX, people will enjoy the game [26].

A successful user experience of serious games depends greatly on how designers deal with the lack of prior knowledge of users about gaming. Designers could facilitate users by incorporating practice levels and a run through of game mechanics and navigation involved. Moreover, by adding adequate and meaningful direct feedback along the way users will be able to learn from their mistakes and adapt to game play with relative ease[12]. The LM-GM framework helps us achieve the balance between learning and gaming objectives in a structured way as described below. For all the four games described in this paper we chose common elements from the pool of LM-GM framework. An attempt was made to optimise the UI and UX for these games by using iterative feedback from a student of Design discipline in our team.

Making use of LM-GM framework [11] we have designed our four games. The LM-GM framework [11] has large number of learning and gaming mechanics as shown in Fig. 6 [12].As game designers we have freedom to choose appropriate set of mechanics from this global set to achieve crucial balance between learning and gaming. One can use this framework to critically review the existing games and attempt improvements by way of better balance between learning and gaming activities. Designers can also describe the dynamics of chosen mechanics during actual game-play.

For application of LM-GM framework in our context we have to identify appropriate gaming and learning mechanics from Table II those suitable for our games. Starting with few gaming and learning mechanics as template, we progress through the game for suitable delivery of the content.

We adapt the steps described in [27] to design our games under LM-GM framework. The game context is set through an instructional backstory that defines the player objectives and rationale for subsequent actions. This is achieved using

TABLE II
LM-GM BASED ANALYSIS OF *OWASP-based* GAME [12], [27]

Game Mechanics	Learning Mechanics	Implementation	Usage
Cut scene/ Story	Instructional	Animation	Backstory telling the game objectives, mechanics & outcome.
Tutorial Cascading Info.	Guidance Tutorials	Levels	The Player is guided through informative graphics to complete tasks.
Simulate Response Experiment	Observation, Analyse, Experimentation Modelling	Player must enter correct commands to complete the required tasks.	Tasks like attack -based command provide a sense of achievement.
Movement Time pressure	Activity/ Tasks	Navigate player through levels, time constraints on levels	Provides a sense of progress, immersion, urgency and game mastery
Strategy Planning	Explore Modelling	The flexible design of the level and tasks to allow different strategies to emerge	A deeper understanding of OWASP-based game mechanics through modeling
Levels, Feedback, Assessment	Feedback, Motivation, Reflect, Assessment	Advance to the next level. The score shows completion time. Feedback prompts	Reinforces a sense of understanding & progress. Player can take actions.
Behavioral Momentum	Repetition, Reflection	Gameplay repeats through levels.	Reinforces behavior change for Web Application secured development.

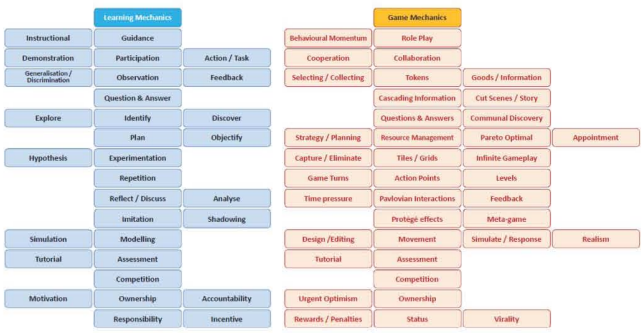


Fig. 6. Learning and game mechanics used as the basis to construct the LM-GM map for a game [12]

[11]http://gost.iitd.ac.in/serious_games/pages/ser.html.

[12]https://uxplanet.org/getting-serious-about-the-ux-of-serious-games-c12382e900e.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 31,2022 at 08:51:17 UTC from IEEE Xplore. Restrictions apply.

TABLE III
LM-GM FRAMEWORK [12], [27]

Learning Mechanics		Game Mechanics	
a) Instructional	h) Simulation	1) Cut scene/Story	9) Planning
b) Tutorial	i) Feedback	2) Tutorial	10) Levels
c) Observation	j) Motivation	3) Cascading info.	11) Feedback
d) Experiment	k) Assessment	4) Simulate/ Response	12) Assessment
e) Modelling	l) Reflection	5) Movement	13) Behavioral Momentum
f) Analyze	m) Repetition	6) Time pressure	
g) Action/task	n) Explore	7) Capture/elimination	
		8) Strategy	

cut scenes. Then the player is introduced to the core mechanics through a short tutorial becoming familiar with the user interface and main controls (cascading information). The core loop of the OWASP based games are to make the Web Application secured in levels, where the player explores each game task (behavioral momentum), tries to understand its structure and how to efficiently secure the Web Application using a simulate/response approach to observe, experiment and analyze the different commands used to study the behavior of the OWASP-identified attack under time constraints. The end of each level provides feedback to the player on their progress (score achieved). This approach ensures that students have to learn and fully-understand to complete each stage.

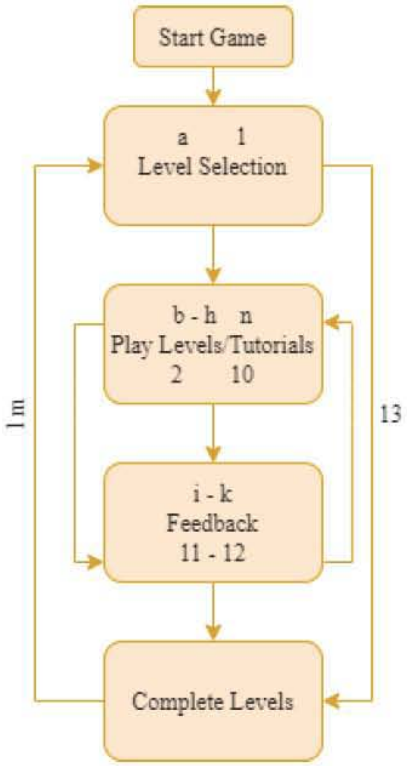


Fig. 7. Game Map of OWASP-based games using LM-GM

We have prepared a game map based on LM-GM framework as shown in Fig. 7. The map gives a dynamic view of the relationships between the LMs and GMs (defined in Table III) in the SGs flow of actions. It can be observed that the game iterates through these steps for each learning outcome for different levels. Table III list the mechanics used in our game from the exhaustive list of the LM-GM described in [11]. The LM-GM framework has thus helped us achieve the reinforcing nature of learning experienced by the player as he progresses through the game.

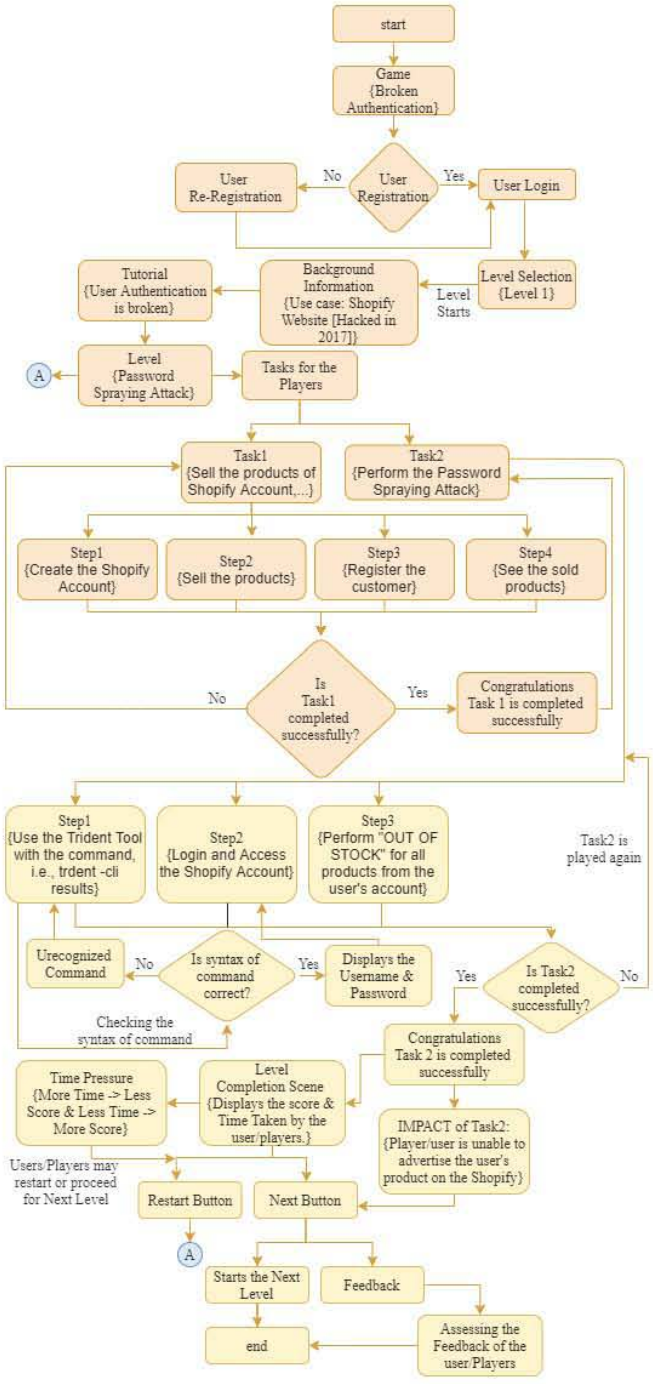


Fig. 8. Illustrative Flowchart

While designing these games a detailed flowchart for each

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 31,2022 at 08:51:17 UTC from IEEE Xplore. Restrictions apply.

stage of game play was prepared as a design document. This flow chart was peer reviewed for possible improvements and an illustrative flowchart is shown in Fig. 8. The steps of the Fig. 8 are considered while choosing the elements from the LM-GM pool. All four games have similar dynamics and the mapping of OWASP based games with LM-GM framework illustrates their iterative dynamics.

IV. Learning Outcome Assessment

LM-GM framework can not help us measure the LO; as highlighted in literature [28]. A limitation of LM-GM is that it does not expose the connection between concrete mechanics and the high-level educational objectives that the game is supposed to attain. In this research we attempt to measure it through the use of player data captured during game play. After evaluating the game mechanics through the lens of LM-GM framework in Section III above, we attempted measurement of the learning outcome of a player adapting the methodology suggested in [18]. Learning outcome is one of the key attributes of a serious game and unless we achieve it in clearly demonstrable manner, we may fall short of our main objective.

In these game, we have divided the games in two phases, namely; teaching phase and learning phase. Both phases have multiple learning goals. The Initial Assessment Score (IA) corresponds to the teaching phase and the Final Assessment Score (FA) corresponds to the testing phase. We have adapted the approach described in [18], [27] to compute these scores. We have been able to collect player data from only eight players. The objective assessment of the player has been computed by using time taken to complete a task and other attributes of the captured data have not been used in this study. The results of the objective scores for set of eight players consisting of 32 data points are clustered in four categories namely; Active Learners, Slow Learners, Masters and Outliers based on the change in values of IA with respect to Initial Threshold (IT) and FA with respect to Final Threshold (FT). This classification of players into four categories for four games are done using k-means clustering [29], is shown in Fig 9. The Active Learners are one who display good performance during FA as against a Slow Learners who display low performance in both IA and FA. Masters are characterised by higher score in IA as they seem to have background knowledge. Outliers are those players who show poor scores in FA even after getting good scores during IA phase. This clustering would be more meaningful when number of players is large, i.e., greater than 30 and therefore, the clustering shown in Fig 9 is only for illustration purpose.

We attempt to get inputs from player on user experience in all four games by using a built-in questionnaire. Player describes his experience on a five-point Likert scale using various attributes as suggested by a study in [30].Subjective feedback from players are compiled in Table IV.

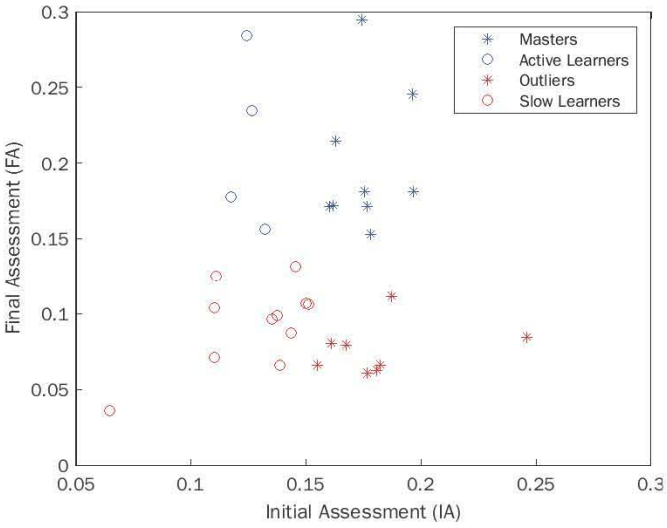


Fig. 9. Clustering of eight players for four games

V. Limitations

Our research paper has some limitations.Though our games are able to capture the objective player data using Unity analytics and subjective feedback through a questionnaire shown in Table IV, we have been able to measure the learning outcome of these games using these artefacts only for eight number of players because of non-availability of players due to on-going Covid pandemic. Although we capture player data on various attributes, in this research we have utilized only time taken by player to complete a level for the purpose of computing Learning Outcome. We plan to involve additional students to facilitate more reliable analysis of our game efficacy.

VI. Concluding Remarks and Future Directions

This paper has attempted to describe design of serious games to teach four of the Open Web Application Security Project (OWASP) top 10 vulnerabilities namely;SQL injection, Cross site scripting (XSS), broken authentication and sensitive data exposure.

The focus on UI and UX has been suitably emphasised and the game design process had regular interaction with a student expert in the design aspects. The user interface went through various iterations to ensure that basic tenets of design are manifested in these games. In this research we have attempted to combine the captured player data and the feedback by player using a Likert Scale on various attributes of the game. The paper has highlighted need of a balance between pedagogy and playfulness using LM-GM framework.We have mapped the learning mechanics elements of our games with the gaming mechanics elements using LM-GM framework and thus attempted to demonstrate the balance between the two.

Our games have provision to assess learning outcome of these games using both subjective feedback from player and objective data using Unity analytics. We plan to estimate learning outcome by involving adequate number of players, in

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 31,2022 at 08:51:17 UTC from IEEE Xplore. Restrictions apply.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 31,2022 at 08:51:17 UTC from IEEE Xplore. Restrictions apply.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 31,2022 at 08:51:17 UTC from IEEE Xplore. Restrictions apply.

Arabian Journal of Geosciences (2022) 15: 213
https://doi.org/10.1007/s12517-021-09410-3

ORIGINAL PAPER

Delineation of groundwater recharge potential zones for its sustainable development utilizing GIS approach in Kabul basin, Afghanistan

Santosh Kumar Singh[1] · **Ali Reza Noori**[1,2]

Received: 11 February 2021 / Accepted: 28 December 2021 / Published online: 20 January 2022
© Saudi Society for Geosciences 2022

Abstract
The Kabul basin is a plateau situated in the east-central part of Afghanistan, covering its capital. Groundwater has a vital role in drinking purposes and agricultural activities in the basin. Accordingly, to fulfill the drinking water and irrigation requirements and address sustainable water resources management for future development, discovering the groundwater recharge potential zones (GWRPZs) is essential. An attempt was undertaken to delineate the GWRPZs combined with the application of remote sensing (RS), geographic information system (GIS) integrated with analytical hierarchy process (AHP). Various influential variables, including geology, geomorphology, soil, lineament and drainage density, precipitation, land use and land cover, slope, elevation, and water depth, were utilized to produce different thematic maps. The raster formats of different thematic features were prepared in the ArcGIS environment. A statistical analysis was done based on the AHP methodology to determine the weights and ratings of various themes. Based on the results, eventually, five categories of the GWRPZs have been recognized. Very poor (8 km^2, less than 1%), poor (187 km^2, 7%), moderate (610 km^2, 23%), good (1084 km^2, 42%), and very good (715 km^2, 27%). The validation of outcomes was done with the utilization of water level records. The kappa coefficient (0.81), and overall accuracy (88.57%) indicate a moderate to good correlation between the result outcomes and the well's water level records. Results offer further insights. The GWRPZs map can be utilized by local authorities to develop sustainable water resources management and artificially recharge groundwater.

Keywords Groundwater · AHP · RS and GIS · Kabul basin · Water scarcity

Introduction

Water is named the vital substance. In most nations, many factors such as population growth, urbanization, water overuse, global warming, climate change, and miss management have significantly challenged this substance. If one looks at the past decades, its drastic decline is evidence of climate change impacts on water bodies. Population growth, urbanization, irrigation, and industrialization are influential parameters in increasing water demands. Rising demand brings under pressure all water resources. However, it will specifically and severely influence groundwater in the areas where groundwater is the sole source (Kundzewicz and Döll 2009; Kanakiya et al. 2014; Abijith et al. 2020). About 30% of overall freshwater is released from groundwater worldwide. Groundwater is the main source of water supply for more than 2 billion people in the globe (Haritash et al. 2018; Achu et al. 2020b). In most of the arid and semiarid regions, water demand widely depends on groundwater.

Climate change is the most pressing issue confronting the earth today. Climate change has a variety of consequences on many environmental and biological domains, and its effects on water resources and the hydrological cycle, in particular, may be highly complex and dynamic. Over the past century, global temperatures have risen by 1 to 2 °C

Responsible Editor: Broder J. Merkel

✉ Ali Reza Noori
alirezanoori_phd2k19@dtu.ac.in

Santosh Kumar Singh
sksinghdce@gmail.com

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

[2] Department of Water Supply and Environmental Engineering, Faculty of Water Resources and Environmental Engineering, Kabul Polytechnic University, Kabul, Afghanistan

(Ahsan et al. 2018), leading to global warming, melting the glaciers and rising global water levels, declining freshwater resources, and increasing numbers of severe weather events (e.g., floods, droughts, storm). Due to a combination of limited adaptation ability and extensive exposure to climate changes, Afghanistan is regularly rated amongst nations highly susceptible to climate change (Aich et al. 2017). During 1951 and 2010, Afghanistan had a warming trend of 1.8 °C, much greater than the global average (Aich et al. 2017). Afghanistan is projected to warm by an average of 2 °C between 2006 and 2050 and then by 2.7 to 6.4 °C by 2099. According to Aawar et al. (2019), in the Kabul basin, precipitation increased 4.88–30.42 mm/year between 2000 and 2018, and temperatures increased significantly at a different rate (− 0.02) to (0.71) °C/year between 2008 and 2018. Regardless of the direction of precipitation, rising temperatures are impacting ecosystems, livelihoods, and notably water resources. Since the changes in precipitation are small in the basin, this does not cause changes in the runoff or groundwater recharge. However, the amount of rainfall that causes surface runoff will cause groundwater recharge and is a potential source for artificial groundwater recharge through rainwater harvesting and other mechanisms. An increase in temperature will lead to negative impacts on water resources due to increase of evapotranspiration rates in the region.

Due to the rise in the human population, developing countries face intense pressure on safe drinking water (Singh and Deepika 2017). The Kabul city, a major part of the Kabul basin with a population of more than 4 million (CIA 2020; NSIA 2020), is one of those communities. The city served nearly all its water needs by extracting groundwater (Jawadi et al. 2020; Noori and Singh 2021). Therefore, currently, groundwater is the only attractive choice for supplying the drinking and hygienic requirements of society. Previous feasibility studies to expand the water supply infrastructure indicated a projected population of 4,089,000 in 2015, and their water demand is approximated about 123.4 Mm^3/year. However, in Kabul city, the groundwater projected supply rate is estimated at 44 Mm^3/year (Zaryab et al. 2017). At the Kabul basin, besides the drinking and residential use of groundwater, this resource also plays an essential role in agricultural and domestic use.

The decrease in groundwater levels in Kabul depends on different parameters. For instance, impacts of climate changes, overexploitation, increase in population, and reduction of recharge zones have been made as prominent key factors in numerous studies (Banks 2002; Qureshi 2002; Uhl and Tahiri 2003; Tünnermeier and Houben 2005; Saffi 2007; Houben et al. 2009; DACAAR 2011; Pell Frischmann 2012). According to Mack et al. (2013), the mean decline rate in groundwater levels in the city is measured at 1.5 m/year from 2008 to 2012 and 0.07 m/year from 2004 to 2008. Meanwhile, in most recent years, an excessive decrease in groundwater level has been reported. An evaluation from Zaryab et al. (2017) indicates that groundwater tables have decreased in the foothills by more than 10 m and in the city's central plains by 5 to 6 m. The water levels in the upper Kabul basin regions decreased by more than 15 m from 2003 to 2016. From 1982 to 2003, the groundwater level was reduced by a mean of 6 m, averaging 0.28 m per year. From 2003 to 2016, there has been a mean of 15 m of groundwater level reduction, which averages 1.15 m per year. The average annual decrease in groundwater levels between 2008 and 2016 is 1.7 m. Among different causes of decreasing groundwater, the population growth and overexploitation of groundwater can be highlighted as the leading ones, which have doubled since the 1990s.

After the new government came to power in the country (2001–2002), utmost Afghan refugees returned, and many settled in the capital. Also, most of the people moved from provinces to the capital for employment and education opportunities. The increasing population in the capital caused the city's unplanned development, increasing water demands, and pollution sources such as wastewater and solid wastes. These increasing factors challenged both the quality and quantity of groundwater. Additionally, a variety of other components are frequently associated with the Kabul water crisis. These components include the lack of government management systems, the overexploitation of water resources, the lack of proper policy and action plans for groundwater extraction and major extractors, and industrial and humanitarian interventions. Therefore, any study that supports the sustainable management of groundwater and water resources is essential. To identify the condition of subsurface water resources and support the artificial recharge of groundwater that was recently adopted from the Afghan government side (KMARP 2017), discovering potential recharge zones of groundwater is crucial.

Remote sensing (RS) and geographic information systems (GIS) hybrid applications offer a powerful method for interpreting possible groundwater zones adding that minimize time consumption and human resources. Utilizing a scope of compelling factors, like geology, the density of lineament and drainage, slope, geomorphology, precipitation, land use and land cover, soils, elevation, and water depth, gives precision and diminishes the odds of human mistakes. The combination of primary and secondary data sets is also possible in RS and GIS (Etikala et al. 2019).

Recently, numerous scholars had attempted RS and GIS to determine groundwater potential zones by applying different approaches, alongside analytical hierarchy process (AHP) (Muralitharan and Palanivel 2015; Abijith et al. 2020; Achu et al. 2020a; Luo et al. 2020), Influence factor (IF) technique (Selvam et al. 2016; Etikala et al. 2019; Siddi Raju et al. 2019), linear equation approach (Al-Abadi et al.

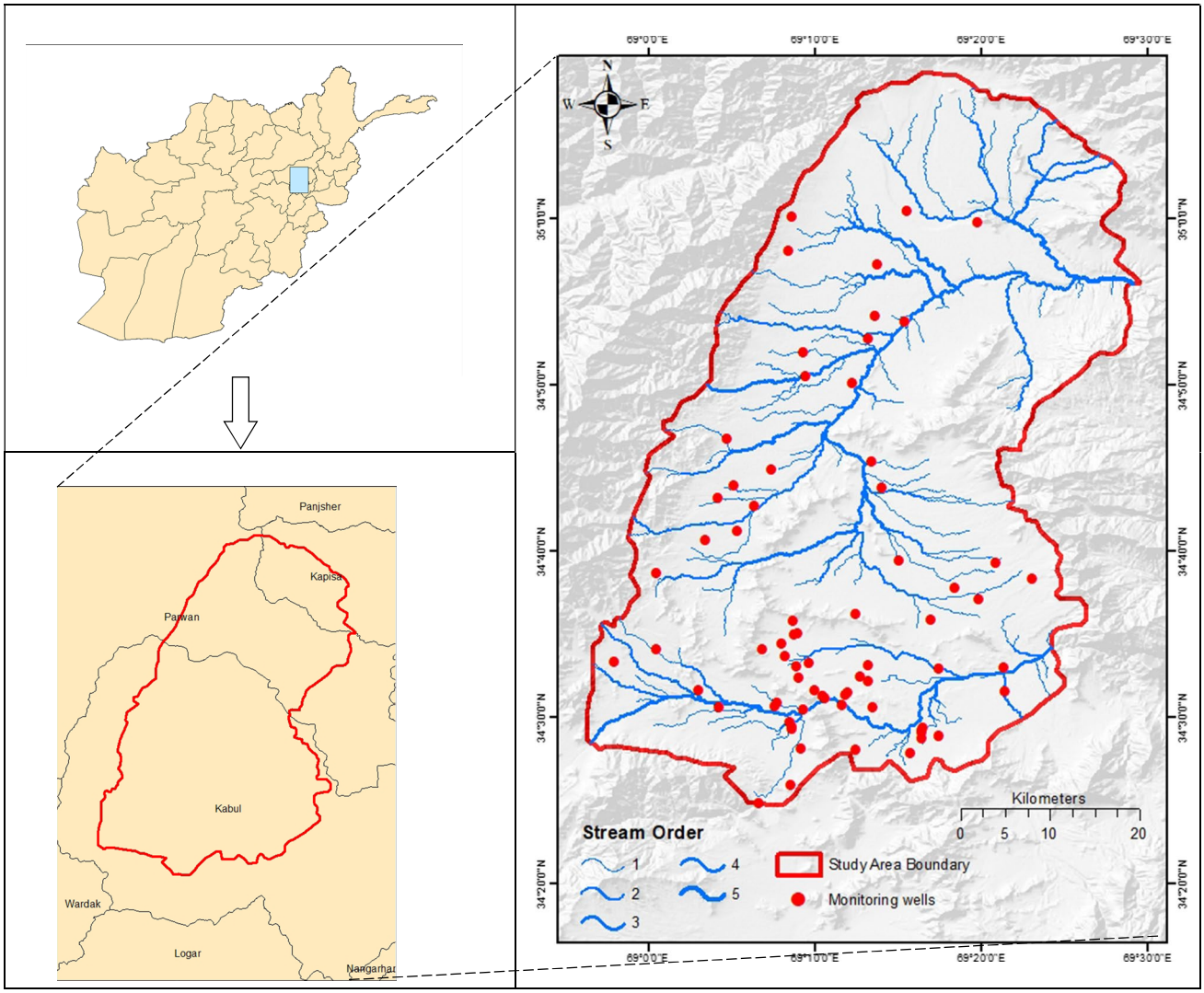


Fig. 1 Location map and watershed of the Kabul basin, Afghanistan

2016; Arnous 2016), probabilistic approach (Manap et al. 2014; Dadgar et al. 2017), geophysical technique (Abuzied and Alrefaee 2017; Kolandhavel and Ramamoorthy 2019; Mpofu et al. 2020), and multi-criteria decision analysis (MCDA) (Chowdhury et al. 2010; JHARIYA et al. 2016; Akinlalu et al. 2017). AHP approach is count as an accurate and cost-effective technique for decoding potential groundwater recharge areas. Therefore, this approach is considered for the current study.

This study aims to delineate groundwater recharge potential zones using the integration of RS, GIS, and AHP method in a semi-arid environment in Kabul basin, Afghanistan. The accuracy of the study was evaluated using overall accuracy and Kappa accuracy method. A 9 year (2004–2013) groundwater level data, recorded by the Afghanistan Geological Survey (AGS) with support of the United States Geological Survey (USGS), was used for accuracy assessment. The GWRPZ map facilitates groundwater resources management to develop sustainable water resources management and facilitate artificial recharge plans and mechanisms.

Study area

Kabul basin is a geologic plain including Afghanistan's capital city, located north–south, surrounded via Pghman mountains westerly and the Saffi mountains easterly. The total area of the study region is around 2600 km^2 and is situated geographically between 34° 25′ 00″ to 35° 10′ 00″ North latitudes and 68° 50′ 00″ to 69° 33′ 00″ East longitudes (Fig. 1). The basin's elevation varies between 1367 m in the northeast to 3054 m in the east and west above sea level. The annual average precipitation in the basin is measured at around 330 mm per year (source: meteorological data from

National Water Affairs Regulatory Authority (NWARA), data recorded 2008–2018). The temperature ranges with a monthly mean of a maximum of 32 °C in July and a minimum of − 7 °C in January (Zaryab et al. 2017). This area of study is a favorite destination and interest (Broshears et al. 2005; Mack et al. 2009, 2010, 2013, 2014; Taher et al. 2013). This study area is located among three neighboring provinces of Kabul, Parwan, and Kapisa.

The slope varies from 0 to 73° between east–west-north within the basin. Loess, conglomerate, and sandstone, followed by gneiss, fan alluvium and colluvium, gneiss and granite, gabbro, and monazite, basalt lava, limestone, and dolomites are considered to be the initial geology of the basin. Geomorphological study area mostly covered by Holocene River valley and intermountain basin, followed by high-mountain relief, mountain relief. Most types of soils were remarked into a well-drained and moderately drained class. With the rapidly growing urbanization in Kabul, the city is situated among the 100 largest cities globally (City Mayors 2020; Wikipedia contributors 2020a). Its land use and land cover (LULC) indicators were classified as barren rocks and barren land, agriculture and vegetation areas, settlements, some minor natural forests, water bodies, and marshland permanently inundated.

Methodology and data used

ArcGIS.10.6 was utilized for digitization and preprocessing of multiple thematic maps needed for this study. Various data sources were employed. The implemented approach is presented in (Fig. 2). For this study, ten different influential parameters were selected and produced in the form of raster data, including geology, drainage density, land use and land cover, slope, geomorphology, soil types, lineament density, rainfall, elevation, and water depth. All of these variables are relevant and influential components in evaluating the potential groundwater recharge zone. The geological map was digitized and georeferenced from the "geologic map of quadrangle 3468, Chak Wardak-Syahgerd (509), and Kabul (510) quadrangles, Afghanistan" (Bohannon and Turner 2005). The geomorphological thematic map was digitized

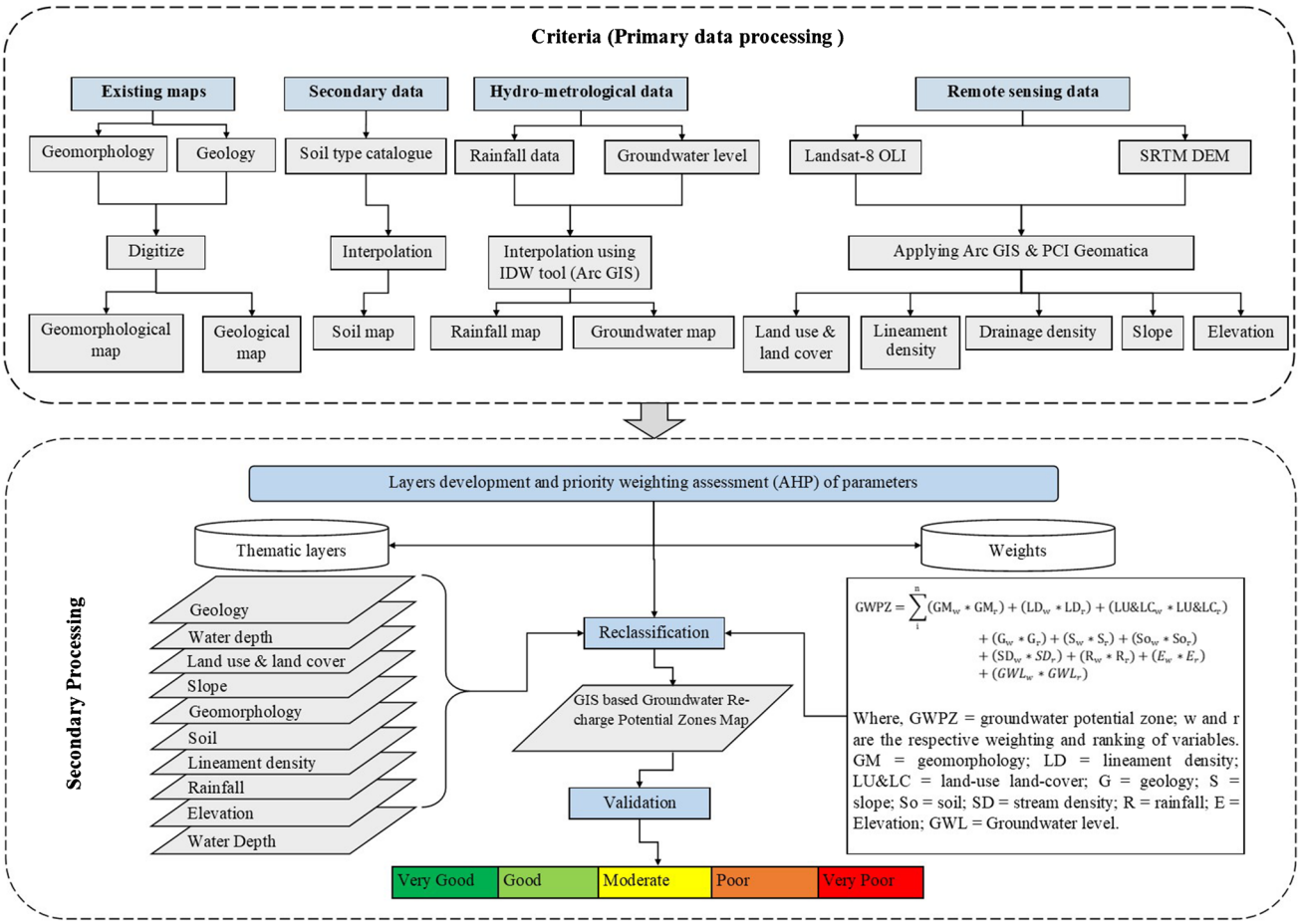


Fig. 2 Flowchart of delineating groundwater recharge potential zones

and extracted from the country geomorphological map prepared by the Afghanistan Information Management System (AIMS) at a scale of 1:6,000,000. The soil map was interpolated based on soil data sources from the Afghanistan soil catalog (Ahmadzai and Omuto 2019).

Landsat-8 OLI satellite data was utilized to configure the land use and land cover map of the study area. The required Landsat data was obtained from USGS Portal (https://earthexplorer.usgs.gov/) using the address. In the ArcGIS 10.6 environment, a "supervised classification strategy" was used by creating training samples for each LULC category. A post-classification smoothing step was conducted using the majority filter tool in the ArcGIS environment to reduce and merge the tiny pixel distribution of LULC classes in the research area.

Shuttle radar topography mission (SRTM) version, 3.0. Global 1 arcsec data, was downloaded and applied to create the slope thematic layer. Landsat-8 OLI data (Landsat scene identifier: LC08_L1TP_153036_20200704_20200708_01_T1; date of acquisition: 02–08-2020) was downloaded to provide the lineament thematic layer with the integration PCI Geomatica Banff and ArcGIS 10.6 software. The rainfall map was created based on meteorological data (2008–2018) from NWARA Afghanistan. An inverse distance weighting (IDW) interpolation technique has been employed to determine its spatial dissemination in the ArcGIS environment. As well as the elevation map was developed using SRTM-digital elevation model (DEM)-30 m data in an ArcGIS environment. A part of 9 years (2004–2013) of groundwater level data, recorded by AGS with support of the USGS, was used to create the groundwater level thematic layer. The study area's watershed map was generated in ArcGIS 10.6 environment employing DEM.

The weights for each factor were assigned based on the priority scale using AHP. Different influential presentive variables had been integrated into the ArcGIS environment to discover the GWRPZs. The consolidated method's findings were characterized in five categories: very good, good, moderate, poor, and very poor zones. The accuracy of the adopted methodology and its effectiveness for delineating groundwater recharge potential zone was obtained utilizing groundwater level data. The overall accuracy assessment and kappa accuracy assessment methods were utilized for the validation of the groundwater recharge potential zones map.

Analytical hierarchy process

To assess and interpret complicated and complex decisions, Thomas L. Saaty proposed a systematic multi-criteria approach called the AHP. (Wikipedia contributors 2020b). This systemic approach analyzes and addresses dynamic decision-making issues through analysis, comparative decisions, and priority synthesis (Achu et al. 2020b). Each variable's ranking is assigned in this method between one to five for equal and extreme importance. All of the attributes are evaluated based on the paired comparison.

The current investigation will figure out the relevance of the variables which affect groundwater potential. Weights obtained from this process are assigned to the prepared thematic maps in the ArcGIS environment using an overlay analysis approach. For the judgment of interaction across each thematic layer, a pairwise comparison matrix operates. The maximum eigenvalue is extracted and normalized from the pairwise matrix to evaluate each variable's relative significance. Accordingly, a weight comparison between the parameters in terms of their impact on groundwater occurrence was carried out based on professional judgment and analysis of the relevant publications. Eventually, the following equations will provide the consistency ratio and index of consistency accordingly (Abijith et al. 2020; Muniraj et al. 2020).

$$CR = \frac{CI}{RI} \tag{1}$$

CI—display for consistency index, CR—indicate consistency ratio, and RI—represent the random index,

$$CI = \frac{\lambda \max - n}{n - 1} \tag{2}$$

where λ_{max} illustrates the matrix's maximum eigenvalue while n is the square matrix order, the RI depends on the number of components. It should be noted that the weights provided through this matrix are only reliable when CR is less than 10%. Table 1 represents the matrix of pairwise comparisons.

Results and discussion

The rankings of each parameter and weight in AHP are illustrated in Table 2. Every sub-feature within a parameter is multiplied by its weights. For each parameter, the AHP weights are given per subclass influences applied to the ranks, and the ranks are numbered from 1 to 9. The weights were measured in such a manner that each parameter's effect was significant.

Geology

The porosity and penetrance of the rocks are clear indicators of groundwater occurrence and recharge in a given area. Whereas geological configurations have a detrimental effect on groundwater occurrence, the geological map is significant (Maheswaran et al. 2016). The thematic geological map

Table 1 Pairwise comparison matrix and normalized AHP wights of thematic layers

Themes	Geology	Drainage density	Land cover	Slope	Geomorphology	Soil	Lineament	Rain	Elevation	Water depth	Normalized AHP weights
Geology	1.00	1.00	2.00	3.00	3.00	2.00	2.00	2.00	1.00	1.00	0.146
Drainage Density	1.00	1.00	1.00	2.00	1.00	3.00	1.00	2.00	2.00	1.00	0.132
land cover	0.50	1.00	1.00	1.00	2.00	2.00	2.00	1.00	1.00	2.00	0.114
Slope	0.50	0.50	1.00	1.00	1.00	2.00	2.00	2.00	1.00	1.00	0.096
Geomorphology	0.33	1.00	0.50	1.00	1.00	2.00	1.00	1.00	1.00	1.00	0.084
Soil	0.50	0.33	0.50	0.50	0.50	1.00	3.00	2.00	0.50	0.50	0.068
Lineament	0.50	1.00	0.50	0.50	1.00	0.33	1.00	0.50	0.33	0.50	0.057
Rain	0.50	0.50	1.00	0.50	1.00	0.50	2.00	1.00	0.25	0.25	0.061
Elevation	1.00	0.50	1.00	1.00	1.00	2.00	3.00	4.00	1.00	3.00	0.136
Water depth	1.00	1.00	0.50	1.00	1.00	2.00	2.00	4.00	0.33	1.00	0.106

was digitized and georeferenced from the "geologic map of quadrangle 3468, Chak Wardak-Syahgerd (509) and Kabul (510) quadrangles, Afghanistan" (Bohannon and Turner 2005). The common geological forms found in the study area are conglomerate and sandstone, loess, limestone, fan alluvium and colluvium, gabbro and monazite, basalt lava, gneiss and granite, sandstone and siltstone, and ultramafic intrusions (Fig. 3a). The thematic map is ranked based on the regulation of geological formations' permeability and groundwater effects. Loess, conglomerate, and sandstones are good aquifers materials due to their high weathering and fractures. Maximum values were classified for loess and conglomerate, and minimum ratings were allocated to gneiss, granite, and basalt lava. The geological characteristic is assigned to a corresponding AHP weight of 0.14.

Drainage density

Precipitated water, glacial, or snowmelt is carried to the plains through drainage. It has an essential role in a given area as the source of groundwater recharge. Hydrogeological faults, fractures, and joints are essential factors for the recharge of groundwater, and they are perceived as guiding criteria that can impact a basin's drainage systems. Therefore, drainage operates as a transmission tool for the passage of snowmelt and glacial melting. Drainage density is defined as the ratio of total stream lengths and the area in a watershed. The small runoff will happen within the neighborhoods with high density due to their higher conductivity (Pinto et al. 2017; Abijith et al. 2020). The extraction of drainage density has been done in the ArcGIS environment by processing SRTM-DEM data (Fig. 3). The importance of drainage density and the occurrence of groundwater are closely associated. The thematic map of drainage density developed into five subclasses. The extracted interval values of 2.0–3.2 km/km^2 were highlighted with high ranking and 0.0–0.3 km/km^2 with a minor value. The drainage density influential layer is assigned to a corresponding weight of 0.12.

Land use and land cover

Land use and land cover are crucial aspects that significantly affect the recovery and occurrence of groundwater (Machiwal et al. 2011; Martin et al. 2017). The development of cities and the impermeability of the land surface will reduce groundwater recharge. Remotely sensed Landsat-8 satellite data was interpreted to generate a LULC thematic map (Fig. 4). A supervised classification technique was adopted. The study area has a mountainous and plain bare land and bare rock area, agricultural and vegetation area, settlements, small natural forest, water bodies, and marshland permanently inundated. Settlements diminish the effect on groundwater recharge, while vegetation-covered areas have more excellent groundwater recharge opportunities. Therefore, bare land and settlement areas are weak in terms of groundwater availability, but agricultural areas and vegetation are very appropriate. For assigning weights and ranks, the highest ranks are appointed on agriculture, followed by marshland, natural bar lands, and the slightest rank is given to settlements. The resultant AHP value of 0.11 is assigned to the LULC feature.

Slope

The slope has a substantial impact on runoff and represents the elevation difference in a given area. Meanwhile, it has a graceful role in the occurrence and recharge of groundwater. Areas with steeper slopes are not considered valuable substrates for groundwater recharge, as steeper slopes cause rapid runoff, are associated with erosion, and reduce water infiltration into the soil. Nevertheless, on the contrary, lands with lower slopes and flat areas are a good bed for

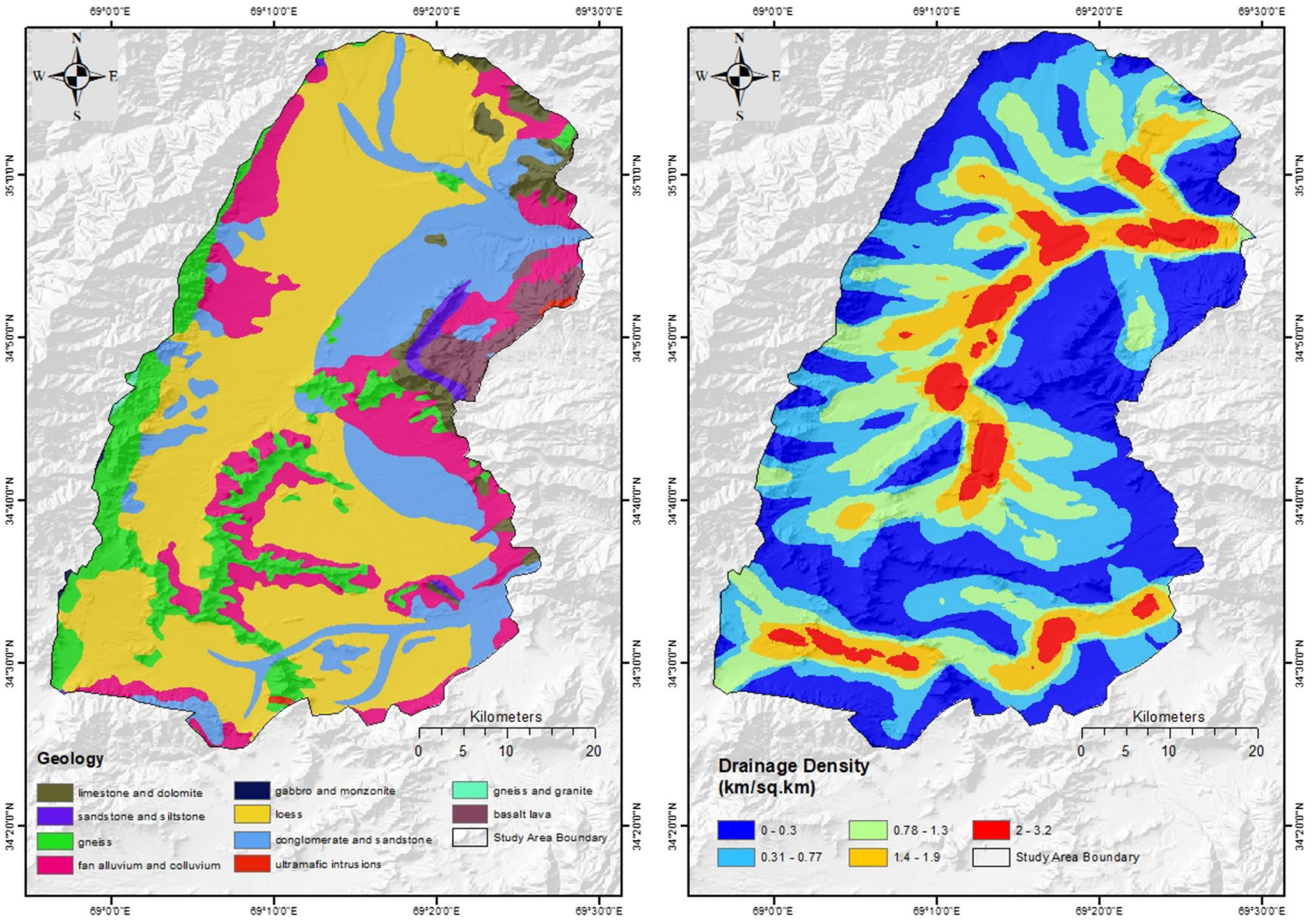
Table 2 AHP ranking and weighting with classes of parameters

Parameter and AHP weights	Rank	Parameter and AHP weights	Rank
Geology (14)		Geomorphology (8)	
Limestone and dolomite	2	High mountain relief	1
Sandstone and siltstone	4	Mountain relief	7
Gneiss	5	Intermountain basin	9
Fan alluvium and colluvium	5	Holocene river valleys	8
Conglomerate and sandstone	9	Soil (7)	
Ultramafic intrusion	4	None	2
Gabbro and monzonite	3	Moderately well-drained	5
Loess	9	Well-drained	8
Gneiss and granite	2	Lineament density (9)	
Basalt lava	1	0.00 -0.079	1
Drainage density (12)		0.08–0.21	3
0.00–0.30	1	0.22–0.36	4
0.31–0.77	3	0.37–0.58	7
0.78–1.30	4	0.59–1.10	8
1.40–1.90	4	Rainfall (8)	
2.00–3.20	5	298–320.1	6
Land use and land cover (11)		320.2–328.3	8
Water bodies	4	328.4–337.7	9
Agriculture area	9	337.8–350.2	7
Marshland	8	350.3–365.5	5
Barren land	6	Elevation (12)	
Barren rock	4	1370–1650	9
Natural forest	3	1660–1850	8
Settlement	4	1860–2030	7
Slope (9)		2040–2270	6
0.00–1.44	5	2280–3050	3
1.45–2.59	4	Water depth (10)	
2.6–4.31	3	2.6–7.7	9
4.32–10.60	1	7.8–13	7
10.70–73.30	1	14–21	6
		22–32	2
		33–53	1

groundwater recharge because, in these areas, runoff moves at a slower speed, and there is enough time for the infiltration process and provides the situation for groundwater recharge (Martin et al. 2017; Pinto et al. 2017). SRTM data was interpreted in the ArcGIS environment to delineate the slope map. The study area is surrounded by mountains whose slope decreases towards the center of the basin. The lateral parts and some central portions of the basin have a high slope, and the majority of the regions have a gentle slope with outstanding qualifications for groundwater. Five different slope categories are presented in the results of the classification of the slope map in the basin (Fig. 4). The lower slope from 0 to 1.44° has been given a high rank, and the lowest rank is allocated to 10.7–73.3°. The derived weight of 0.09 is assigned to the slope feature.

Geomorphology

A geomorphological map displays significant geomorphic units, landforms, and underlying geology. Morphology and its features exhibit a substantial impact on development, filtration, and water recharging into the earth's subsurface (Kumar and Pandey 2016; Etikala et al. 2019). A thematic geomorphological layer has been digitized and



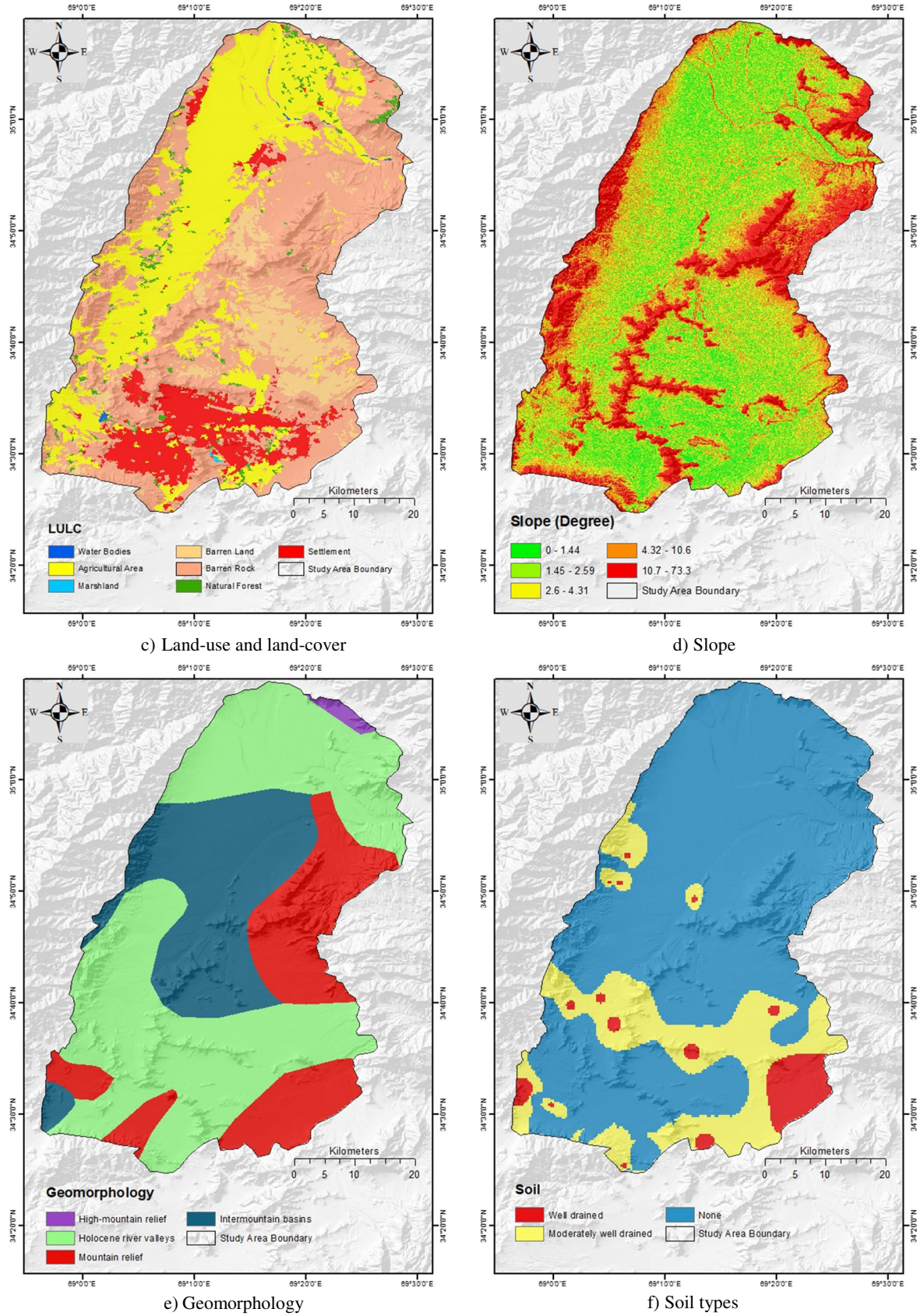
a) Geology b) Drainage density

Fig. 3 Geology and drainage density maps of the basin. **a** Geology. **b** Drainage density. **c** Land use and land cover. **d** Slope. **e** Geomorphology. **f** Soil types. **g** Lineament density. **h** Rainfall. **i** Elevation. **j** Groundwater depth

extracted in ArcGIS 10.6 from the country geomorphological map prepared by AIMS (Fig. 4). The basin has high-mountain relief, Holocene river valley, intermountain basin, and mountain relief. High-mountain relief in the northern part can have more runoff and less groundwater potential. Holocene river valley and intermountain basin having a gentle slope with medium vegetation, widely cover most parts of the basin. These geomorphologic points can be the appropriate infiltration beds for groundwater recharge potential. Mountain relief covers some parts of the area in the east and south. Holocene river valley and intermountain basin assigned high rank followed. The lowest grade was given to high mountain relief. The derived weight of 0.08 is assigned to the geomorphological feature.

Soil

Soil forms as a consequence of the erosion process from fragments of rocks. It is the only medium in which groundwater is recharged, so its attributes are essential for discovering groundwater potential. Soils cause groundwater recharge and storage in terms of their storage capacity and absorption rates, depending on the soil type (Abijith et al. 2020). The required data were extracted from Afghanistan Soil Catalogue (Ahmadzai and Omuto 2019) to delineate the soil map. Soil is labeled as per the drainage ranges, showing wetness period and well-drained, moderately well-drained, and none is observed in the basin. Most of the study area is covered by a well-drained soil class, mostly farmland (Fig. 4). High and least ranks were



c) Land-use and land-cover

d) Slope

e) Geomorphology

f) Soil types

Fig. 4 LULC, slope, geomorphology, and soil maps of the study area

assigned to well-drained and class of none accordingly. The derived weight of 0.07 is given to the soil feature.

Lineament density

In the crystalline terrain, curved or linear geological formations are called lineaments and are massively crucial in groundwater potential recharge and flow. Primary lineaments and their convergence regions will serve as potential high-yield groundwater areas. Its role in the recharging of groundwater is also undeniable. (Thapa et al. 2017; Siddi Raju et al. 2019). Lineament density map (Fig. 5) is prepared from Landsat-8 OLI data (Landsat Scene Identifier: LC08_L1TP_153036_20200704_20200708_01_T1; Date of acquisition: 02–08-2020) using PCI Geomatica Banff and ArcGIS 10.6. The high to low-density scale is usually associated with the groundwater capacity. The alignments of lineament in the current study are N-S, NW–SE, NE-SW, and E-W. The intersection areas of lineaments have a high groundwater potential. The north, north-central, west, and southwestern parts of the study area represent wider concentrations of lineament. The areas possessing larger lineament density (0.59–1.1 km.km^{-2}) were assigned as high rank, and accordingly, areas with minor values of lineament density (0–0.079 km.km^{-2}) were assigned low scale values. The derived weight of 0.09 is given to the lineament feature.

Rainfall

Rainfall seems to be the only main groundwater recharge source and the water cycle's primary water source. It counts as the critical explanation for recharging groundwater where water from cracks and soil penetrates the earth's subsurface. Other attributes of precipitation, such as duration and frequency, also affect runoff and recharge (Abijith et al. 2020). Ten years daily (2008–2018) precipitation data were obtained from NWARA Afghanistan, and its spatial distribution was interpolated in the ArcGIS environment using the IDW technique (Fig. 5). The mean annual precipitation is obtained at 330 mm in the Kabul basin. According to average annual precipitation, the research area is subdivided into five distinct regions. As the whole basin is in a moderate state, the records show somewhat higher precipitation in the basin's southwestern parts. High rainfall area is ranked with a higher value of 350–365 mm, but the lower rank is assigned to 298–320 mm. The resulting derived wight of 0.08 is allocated to the precipitation feature.

Elevation

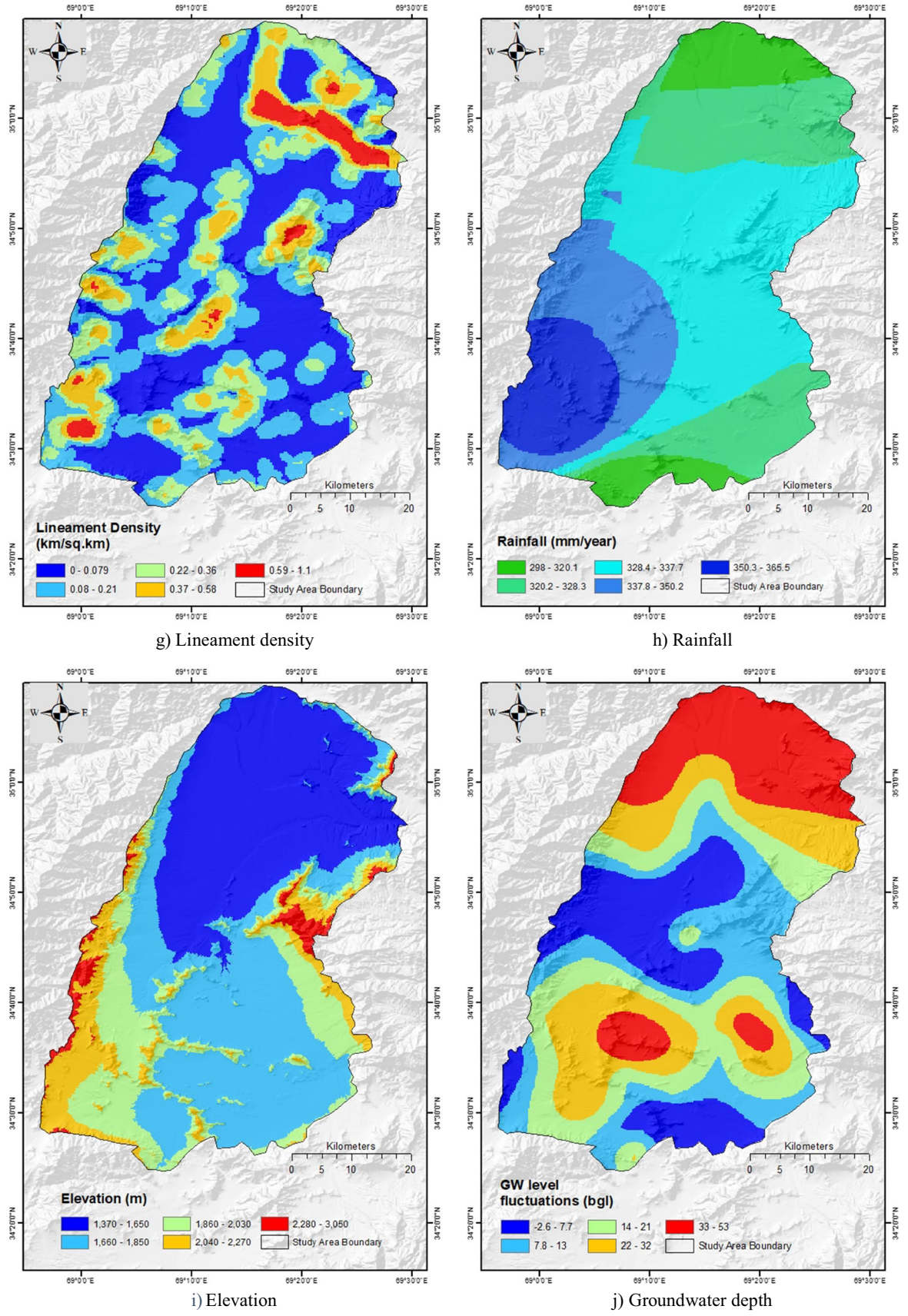
The altitude variable is one of the primary factors in the differentiation of GWRPZ. There is an indirect correlation between elevation and infiltration rates (Thapa et al. 2017). The 30-m resolution SRTM-DEM data has been utilized to delineate the elevational map of the research area. The highest elevation is differentiated mostly in the western and southwestern sections of the study region. The eastern and northeastern parts also have high altitudes. Regions with moderate and low elevations generally cover the Kabul basin's central parts (Fig. 5). According to elevation from mean sea level, the basin is grouped into five subclasses, including 1370–1650, 1660–1850, 1860–2030, 2040–2270, and 2280–3050 in the arcGIS environment. Strong preference is assigned to lowland, followed by moderate and high elevation areas. The resulting weight of 0.12 is allocated to the elevation feature.

Groundwater levels

Long-term statistics on the variance of groundwater tables give insights into the prospects for groundwater. It is also regarded as one of the crucial variables for the discovery of possible groundwater recharge areas. The AGS recorded nine-years monthly monitoring (2004–2013) of groundwater levels in the Kabul basin supported by USGS. Out of the 71 observation wells, only one of them is found beyond this analysis's study boundaries, which was excluded from the calculations. Recorded data from forty observational wells out of seventy were randomly selected to create the groundwater level thematic layer. The water level data of all 70 observations well stations were utilized to validate the final produced GWRPZ map. The data's interpretation revealed that, during observation, the minimum and maximum water depth ranged from – 2.6 to 7.7 m and 33 to 53 m below ground level, respectively (Fig. 5). The areas with shallow groundwater levels ranked higher weights, and, conversely, less weight was given to the more in-depth groundwater classes. The resulting weight of 0.10 is allocated to the groundwater level feature.

Groundwater potential zone mapping

Finally, weights were allocated to the input variables utilized the AHP method for GWRPZ delineation. All influential layers have been converted to a raster format and multiplied by using ArcGIS 10.6 with derived weights and ranks. Ultimately, the final output of processed thematic layers was obtained in five different categories. Finally, the findings were summarized and identified as very poor, poor, moderate, good, and very good (Fig. 6). The following equation was considered for the development of the ultimate GWRPZ map:



g) Lineament density h) Rainfall

i) Elevation j) Groundwater depth

Fig. 5 Lineament density, rainfall, elevation, and groundwater depth maps of the study area

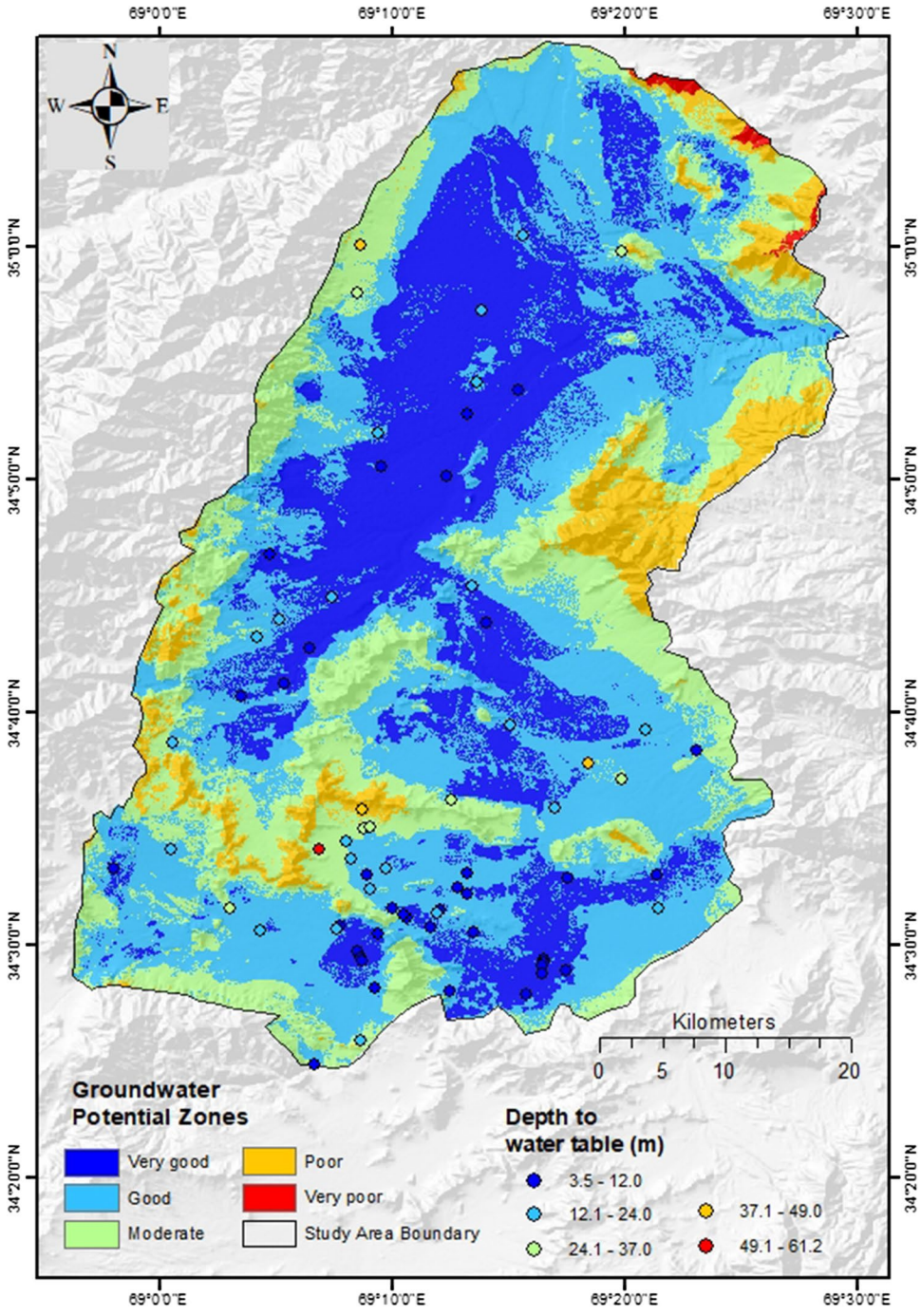


Fig. 6 Groundwater recharge potential zones map using AHP

$$\begin{aligned}\mathrm{GWRPZ} = \sum_{i}^{n} &\left(GM_w \times GM_r\right) + \left(LD_w \times LD_r\right) \\ &+ \left(LU\text{and } LC_w \times LU\text{and } LC_r\right) + \left(G_w \times G_r\right) + \left(S_w \times S_r\right) \\ &+ \left(So_w \times So_r\right) + \left(SD_w \times SD_r\right) + \left(R_w \times R_r\right) \\ &+ \left(E_w \times E_r\right) + \left(GWL_w \times GWL_r\right)\end{aligned} \tag{3}$$

where GWRPZ = groundwater recharge potential zone; w = weighting, and r = ranking of the variables. GM = Geomorphology, LD = Lineament Density, $LU\&LC$ = Land-use and Land-cover, G = Geology, S = Slope, So = Soil, SD = Stream Density, R = Rainfall, E = Elevation, GWL = Groundwater level.

A region with a weak and fragile potential is located in the basin's periphery, dominated by hills, slopes, and mountainous areas. The center portion of the basin has a moderate to high groundwater recharge potential. Geology has received a high priority and lower priorities assigned for soil and geomorphology. The most significant groundwater potential zones identified are intermountain basin

Table 3 Uncertainty matrix of groundwater recharge potential zones

Numbers	GWRPZ	Very good	Good	Moderate	Poor	Very poor	Total	Correct samples
1	Very good	34	4	0	0	0	38	34
2	Good	1	19	1	0	0	21	19
3	Moderate	0	0	7	0	0	7	7
4	Poor	0	0	1	2	0	3	2
5	Very poor	0	0	1	0	0	1	0
	Total	27	35	8	0	0	70	62

Overall accuracy = 62/70 = 88.57%

Kappa coefficient = 0.81

and Holocene River valley in geomorphology, loess, conglomerate, sandstone geologically, cropland in land use, and land cover well-drained soil in soil types throughout the basin. The center parts of the basin illustrate groundwater's intense potential, with a gentle slope and high drainage density. Rainfall has an essential function in the recharging of groundwater. High rainfall occurs in the southwest of the basin; thus, southwestern regions have adequate groundwater potential levels. The obtained outcomes from the AHP indicate that the percentage of the basin under very poor are less than one percent, poor (7%), moderate (23%), good (42%), and very good (27%). There has been discovered high and very high GWRPZs in the most parts of the basin, which are located along with the river courses.

Validation of results

The result has been validated using nine years (2004–2013) recorded data of the static water level for 71 monitoring wells through the AGS department supported with USGS. The monthly collected data has been classified with the minimum, average, and maximum values. Mean values of each monitoring well have been used to validate the results. Except for one monitoring well, all others were located within the research area's boundaries (Fig. 6). The average value of water depth in the observed monitoring wells ranged between 3.5 and 53 m. According to the observation of static water level in the well, the mentioned monitoring wells were classified into five classes viz. – 2.7–7.7 m, 7.8–13 m, 14–21 m, 22–32 m, and 33–53 m that was attributed to as very good, good, moderate, poor, and very poor (Fig. 6). Accuracy assessment is determining the relationship between the resulting groundwater potential zone map and field observations. An uncertainty matrix is used to determine reliability (Table 3). The following formula is applied to determine the overall accuracy (Siddi Raju et al. 2019).

$$\text{Overall accuracy} = \frac{\text{No. of correct MWL}}{\text{Total No. of MWL}} = \frac{62}{70} = 88.57\% \quad (4)$$

where MWL = monitoring well location.

The agreement's kappa coefficient is used since 1980 in remote sensing as an index to remark the accuracy of classified images for thematic layers. The Kappa coefficient of the agreement is calculated from the following formula (Foody 2020).

$$k = \frac{p_o - p_e}{1 - p_e} \quad (5)$$

where p_o is the percentage of the overall precision and p_e is the percentage of appropriate agreement with the values measured. The overall accuracy and kappa coefficient values are obtained at 88.57% and 0.81, respectively. The accuracy assessment in both overall and kappa assessment methods indicates a very good association among potential groundwater recharge zones and well's record values.

Conclusion

It is essential to study and evaluate groundwater's recharge potential areas and delineate it in places where water scarcity is challenging. This analysis examines the reconnaissance of groundwater recharge potential zonation in the Kabul basin, Afghanistan. This research strongly indicates that the hybrid application of both GIS and remotely sensed data, in addition to the analytical hierarchy process, provides a powerful tool to assess the zones of groundwater potential and recharge. Ten different thematic variables were analyzed in the ArcGIS environment with various numerical weightage values in the basin. These variables included geology, geomorphology, land use and land cover, lineament density, drainage density, soil, slope, rainfall, elevation, and water depth. Static groundwater level records have been utilized to acquire precision and reliability for discovering groundwater recharge potential zones. According to the final output

of the results, most parts of the study area are covered by a reasonable and very good capacity of groundwater recharge potential zones. Based on the results, eventually, five categories of the GWRPZs have been recognized. According to the statistics, the area is divided as a zone with a very poor recharge potential (8 km^2, less than 1%), poor (187 km^2, 7%), moderate (610 km^2, 23%), good (1084 km^2, 42%), and very good (715 km^2, 27%). The analysis and final developed map outputs would help manage groundwater in the Kabul basin to promote artificial recharge and sustainable development of water resources.

Availability of data and material The data that support the findings of this study are available from the corresponding author upon reasonable request.

Code availability Not applicable.

Declarations

Conflict of interest The authors declare that they have no competing interests.

References

Aawar T, Khare D, Singh L (2019) Identification of the trend in precipitation and temperature over the Kabul River sub-basin : a case study of Afghanistan. Model Earth Syst Environ 5:1377–1394. https://doi.org/10.1007/s40808-019-00597-9

Abijith D, Saravanan S, Singh L et al (2020) GIS-based multi-criteria analysis for identification of potential groundwater recharge zones - a case study from Ponnaniyaru watershed, Tamil Nadu, India. HydroResearch 3:1–14. https://doi.org/10.1016/j.hydres.2020.02.002

Abuzied SM, Alrefaee HA (2017) Mapping of groundwater prospective zones integrating remote sensing, geographic information systems and geophysical techniques in El-Qaà Plain area, Egypt. Hydrogeol J. https://doi.org/10.1007/s10040-017-1603-3

Achu AL, Reghunath R, Thomas J (2020a) Mapping of groundwater recharge potential zones and identification of suitable site-specific recharge mechanisms in a tropical river basin. Earth Syst Environ 4:131–145. https://doi.org/10.1007/s41748-019-00138-5

Achu AL, Thomas J, Reghunath R (2020b) Multi-criteria decision analysis for delineation of groundwater potential zones in a tropical river basin using remote sensing, GIS and analytical hierarchy process (AHP). Groundw Sustain Dev 10:100365. https://doi.org/10.1016/j.gsd.2020.100365

Ahmadzai H, Omuto C (2019) Afghanistan soil catalogue. FAO, Kabul

Ahsan S, Bokhari A, Ahmad B et al (2018) Future climate change projections of the Kabul River basin using a multi-model ensemble of high-resolution statistically downscaled data. Earth Syst Environ. https://doi.org/10.1007/s41748-018-0061-y

Aich V, Akhundzadah NA, Knuerr A et al (2017) Climate change in Afghanistan deduced from reanalysis and coordinated regional climate downscaling experiment (CORDEX) - South Asia simulations. Climate. https://doi.org/10.3390/cli5020038

Akinlalu AA, Adegbuyiro A, Adiat KAN et al (2017) Application of multi-criteria decision analysis in prediction of groundwater resources potential : a case of Oke-Ana, Ilesa Area Southwestern, Nigeria. NRIAG J Astron Geophys 6:184–200. https://doi.org/10.1016/j.nrjag.2017.03.001

Al-Abadi AM, Pourghasemi HR, Shahid S, Ghalib HB (2016) Spatial mapping of groundwater potential using entropy weighted linear aggregate novel approach and GIS. Arab J Sci Eng. https://doi.org/10.1007/s13369-016-2374-1

Arnous MO (2016) Groundwater potentiality mapping of hard-rock terrain in arid regions using geospatial modelling : example from Wadi Feiran. Hydrogeol J. https://doi.org/10.1007/s10040-016-1417-8

Banks D (2002) Guidelines for sustainable use of groundwater in Afghanistan

Bohannon RG, Turner KJ (2005) Department of the interior (U.S.). In: USGS. http://pubs.usgs.gov/of/2005

Broshears RE, Akbari MA, Chornack MP et al (2005) Inventory of ground-water resources in the Kabul Basin, Afghanistan. USGS

Chowdhury A, Jha MK, Chowdary VM (2010) Delineation of groundwater recharge zones and identification of artificial recharge sites in West Medinipur district, West Bengal, using RS, GIS and MCDM techniques. Environ Earth Sci 1209–1222. https://doi.org/10.1007/s12665-009-0110-9

CIA (2020) The World Fact Book. https://www.cia.gov/library/publications/the-world-factbook/geos/af.html. Accessed 30 May 2020

City Mayors (2020) Largest cities in the world and their mayors – 1 to 150. http://www.citymayors.com/statistics/largest-cities-mayors-1.html. Accessed 1 Sep 2020

DACAAR (2011) National groundwater monitoring wells network in Afghanistan. DACAAR, Kabul

Dadgar MA, Zeaieanfirouzabadi P, Dashti M, Porhemmat R (2017) Extracting of prospective groundwater potential zones using remote sensing data, GIS, and a probabilistic approach in Bojnourd basin, NE of Iran. Arab J Geosci. https://doi.org/10.1007/s12517-017-2910-7

Etikala B, Golla V, Li P, Renati S (2019) Deciphering groundwater potential zones using MIF technique and GIS: a study from Tirupati area, Chittoor District, Andhra Pradesh, India. HydroResearch 1:1–7. https://doi.org/10.1016/j.hydres.2019.04.001

Foody GM (2020) Explaining the unsuitability of the kappa coefficient in the assessment and comparison of the accuracy of thematic maps obtained by image classification. Remote Sens Environ 239:111630. https://doi.org/10.1016/j.rse.2019.111630

Haritash AK, Aggarwal A, Soni J et al (2018) Assessment of fluoride in groundwater and urine, and prevalence of fluorosis among school children in Haryana, India. Appl Water Sci 8:1–8. https://doi.org/10.1007/s13201-018-0691-0

Houben G, Niard N, Tünnermeier T, Himmelsbach T (2009) Hydrogeology of the Kabul Basin (Afghanistan), part I: aquifers and hydrology. Hydrogeol J 17:665–677. https://doi.org/10.1007/s10040-008-0377-z

Jawadi HA, Sagin J, Snow DD (2020) A detailed assessment of groundwater quality in future development. Water 1–19

Jhariya DC, Kumar T, Gobinath M et al (2016) Assessment of groundwater potential zone using remote sensing, GIS and multi criteria decision analysis techniques. J Geol Soc India 88:481–492

Kanakiya RS, Singh S, Sharma J (2014) Determining the water quality index of an urban water body Dal Lake, Kashmir, India. IOSR J Environ Sci Toxicol Food Technol 8:64–71. https://doi.org/10.9790/2402-081236471

KMARP (2017) Kabul managed aquifer recharge project preparation. In: KMARP. https://kmarp.com/home. Accessed 2 May 2020

Kolandhavel P, Ramamoorthy S (2019) Investigation of groundwater potential zones in NandiAru Sub Basin, Tamilnadu, India — an integrated geophysical and geoinformatics approach. Arab J Geosci

Kumar A, Pandey AC (2016) Geoinformatics based groundwater potential assessment in hard rock terrain of Ranchi urban environment, Jharkhand state (India) using MCDM-AHP techniques. Groundw Sustain Dev 2–3:27–41. https://doi.org/10.1016/j.gsd.2016.05.001

Kundzewicz ZW, Döll P (2009) Will groundwater ease freshwater stress under climate change ? Hydrol Sci J. https://doi.org/10.1623/hysj.54.4.665

Luo D, Wen X, Zhang H et al (2020) An improved FAHP based methodology for groundwater potential zones in Longchuan River basin, Yunnan Province, China. Earth Sci Informatics 13:847–857. https://doi.org/10.1007/s12145-020-00469-2

Machiwal D, Jha MK, Mal BC (2011) Assessment of groundwater potential in a semi-arid region of India using remote sensing, GIS and MCDM techniques. Water Resour Manag 25:1359–1386. https://doi.org/10.1007/s11269-010-9749-y

Mack TJ, Akbari MA, Ashoor MH et al (2009) Conceptual model of water resources in the Kabul Basin, Afghanistan. USGS

Mack TJ, Chornack MP, Coplen TB et al (2010) Availability of water in the Kabul Basin, Afghanistan. USGS

Mack TJ, Chornack MP, Taher MR (2013) Groundwater-level trends and implications for sustainable water use in the Kabul Basin, Afghanistan. Environ Syst Decis 33:457–467. https://doi.org/10.1007/s10669-013-9455-4

Mack TJ, Chornack MP, Verstraeten IM (2014) Chapter 11 sustainability of water supply at military installations, Kabul Basin, Afghanistan. pp 199–226

Maheswaran G, Geetha Selvarani A, Elangovan K (2016) Groundwater resource exploration in Salem district, Tamil Nadu using GIS and remote sensing. J Earth Syst Sci 125:311–328. https://doi.org/10.1007/s12040-016-0659-0

Manap MA, Nampak H, Pradhan B et al (2014) Application of probabilistic-based frequency ratio model in groundwater potential mapping using remote sensing data and GIS. Arab J Geosci 711–724. https://doi.org/10.1007/s12517-012-0795-z

Martin SL, Hayes DB, Kendall AD, Hyndman DW (2017) The land-use legacy effect: towards a mechanistic understanding of time-lagged water quality responses to land use/cover. Sci Total Environ 579:1794–1803. https://doi.org/10.1016/j.scitotenv.2016.11.158

Mpofu M, Madi K, Gwavava O (2020) Groundwater for sustainable development remote sensing, geological, and geophysical investigation in the area of Ndlambe municipality, Eastern Cape Province, South Africa : implications for groundwater potential. Groundw Sustain Dev 11:100431. https://doi.org/10.1016/j.gsd.2020.100431

Muniraj K, Jesudhas CJ, Chinnasamy A (2020) Delineating the groundwater potential zone in Tirunelveli Taluk, South Tamil Nadu, India, using remote sensing, geographical information system (GIS) and analytic hierarchy process (AHP) techniques. Proc Natl Acad Sci India Sect A Phys Sci 90:661–676. https://doi.org/10.1007/s40010-019-00608-5

Muralitharan J, Palanivel K (2015) Groundwater targeting using remote sensing, geographical information system and analytical hierarchy process method in hard rock aquifer system, Karur district, Tamil Nadu, India. Earth Sci Informatics 8:827–842. https://doi.org/10.1007/s12145-015-0213-7

Noori AR, Singh SK (2021) Status of groundwater resource potential and its quality at Kabul, Afghanistan : a review. Environ Earth Sci 80:1–13. https://doi.org/10.1007/s12665-021-09954-3

NSIA (2020) Estimated Population of Afghanistan 2020–21. Kabul, Afghanistan

Pell Frischmann (2012) Afghanistan resource corridor development: water strategy final Kabul River basin Report Version 4.0

Pinto D, Shrestha S, Babel MS, Ninsawat S (2017) Delineation of groundwater potential zones in the Comoro watershed, Timor Leste using GIS, remote sensing and analytic hierarchy process (AHP) technique. Appl Water Sci 7:503–519. https://doi.org/10.1007/s13201-015-0270-6

Qureshi AS (2002) Water resources management in Afghanistan. International Water Management Institute

Saffi MH (2007) Ground water resources at risk in Afghanistan. DACAAR

Selvam S, Dar FA, Magesh NS et al (2016) Application of remote sensing and GIS for delineating groundwater recharge potential zones of Kovilpatti Municipality, Tamil Nadu using IF technique. Earth Sci Informatics 9:137–150. https://doi.org/10.1007/s12145-015-0242-2

Siddi Raju R, Sudarsana Raju G, Rajasekhar M (2019) Identification of groundwater potential zones in Mandavi River basin, Andhra Pradesh, India using remote sensing, GIS and MIF techniques. HydroResearch 2:1–11. https://doi.org/10.1016/j.hydres.2019.09.001

Singh SK, Deepika NA (2017) Assessment of water quality parameters of Bhalswa Lake in New Delhi. Int J Environ Eng 9:52. https://doi.org/10.1504/ijee.2017.10008926

Taher MR, Chornack MP, Mack TJ (2013) Groundwater levels in the Kabul Basin, Afghanistan,2004–2013. USGS

Thapa R, Gupta S, Guin S, Kaur H (2017) Assessment of groundwater potential zones using multi-influencing factor (MIF) and GIS: a case study from Birbhum district, West Bengal. Appl Water Sci 7:4117–4131. https://doi.org/10.1007/s13201-017-0571-z

Tünnermeier T, Houben DG (2005) Hydrogeology of the Kabul Basin part I : geology, aquifer characteristics, climate and hydrography. BGR

Uhl WV, Tahiri MQ (2003) An overview of groundwater resources and challenges. Vincent W. Uhl Uhl, Baron, Rana Associates, Inc., Washington Crossing, PA, USA

Wikipedia contributors (2020a) Kabul. In: Wikipedia, Free Encycl. https://en.wikipedia.org/w/index.php?title=Kabul&oldid=976039159. Accessed 31 Aug 2020

Wikipedia contributors (2020b) Analytic hierarchy process. In: Wikipedia, Free Encycl. https://en.wikipedia.org/w/index.php?title=Analytic_hierarchy_process&oldid=978488631. Accessed 22 Sep 2020

Zaryab A, Noori AR, Wegerich K, Kløve B (2017) Assessment of water quality and quantity trends in Kabul aquifers with an outline for future drinking water supplies. Cent Asian J Water Res 3:3–11

DepthNet: A Monocular Depth Estimation Framework

Anunay
Department of Mechanical Engineering
Delhi Technological University (DTU)
Delhi, India
anunay2608@gmail.com

Pankaj
Department of Computer Science and Engineering
Delhi Technological University (DTU)
Delhi, India
Pankajrajput020010@gmail.com

Chhavi Dhiman
Department of Electronics and communication Engineering
Delhi Technological University (DTU)
Delhi, India
chhavi.dhiman@dtu.ac.in

2021 International Conference on Engineering and Emerging Technologies (ICEET) | 978-1-6654-2714-2/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICEET53442.2021.9659770

***Abstract*—Depth estimation is an important technique in computer vision for applications such as pose modelling, activity recognition, etc. Common methods of depth estimation include stereo vision which aims to trace the corresponding feature displacement to reconstruct a geometry as a depth map. To reduce the reliance on stereo vision systems, training-based deep learning methods have been utilized to generate depth maps using RGB images from a single camera. In this paper we propose a DepthNet framework to predict the relative depth of objects placed in the scene of an RGB image with respect to each other, with the help of multiple resizing upskip connections and up-convolutional layers, which further enhance its depth estimation abilities. The results specialize on the edges and the gradients of the objects existing within the scene by introduced as an Edge Loss. The single layered model with limited number of parameters which enables it to be implemented on any platform which has certain limiting factors of space occupancy and computational power. To validate the performance of the proposed architecture, the experiments are conducted on three publically available datasets: NYU depth dataset V2 [1], KITTI depth dataset [2], SUN-RGBD dataset. The projected work exhibits superior depth estimation results victimizing single RGB images, from the opposite state-of-the-arts. The performance of proposed work is evaluated on the following victimizing parameters: REL, RMS, Squared REL, and RMS_log10.**

Keywords- monocular cameras, RGB single image, depth estimation, edge loss, upskip connections

I. INTRODUCTION

Monocular Depth Estimation, known to be read for a long time period with its vast applications both in field of electronics and Computer Vision, concerning Robot Vision [3], Simultaneous Localization and Mapping(SLAM) [4], 3D modeling [5], human computer interaction [6], in Self-Driving Cars [7], etc. However to fulfill these widespread applications, depth estimation techniques are restricted to utilization of Depth Predicting Sensors such as Activity Vision or LIDAR [8] which measures the target distance by illuminating it with a laser light or by using multi-view observations by stereo matching. Application of these methods [9] [10] [11] being extremely costly and space occupying tackles the real life problem, which gave rise to solutions to overcome the problem, using a single sensor (camera), as well as to the present, a haul of obtaining a high-quality dense depth map from a single RGB image has entered the scene. So, during this work we have a tendency to propose a replacement network architecture named DepthNet and a custom Edge Loss to beat this unwell display downside, by rising this progressive paper for Depth Estimation on two publicly obtainable datasets for indoor (NYU depth dataset V2 [1]) and outdoor (KITTI depth dataset [2]) environments. As of now we have perceived that the depth values of indoor environments are distinctly different from that of outdoor ones. This is an intrinsic property of design that is mostly induced by the perspective-effect during the depth acquisition process.

A commonly observed problem during the time of depth estimation is of predicting the objects and differentiating them separately so that the depth-map of objects kept at the same distance, did not get merged with each other. For example, knowing the shape of an object together with its edge detection will be mutually beneficial for predicting the relative depth map of the RGB image. To address the above challenges, our main contributions are as, the proposed a deep neural network, called DepthNet predicts the relative depth of objects placed in the scene of an RGB image with respect to each other. The flow of the architecture includes the resizing upskip connections [12]and up-convolutional layers has further enhanced its depth estimation ability for monocular cameras. The custom Edge Loss [13] while estimating depth maps, helps to discriminate the edges of different objects so that they do not get merged when placed next to or at the same depth with respect to another object.

The rest of paper is organized as follows: The related work of the proposed model is present in Section II. Followed by Section III which describes the detailed proposed work of Depth Estimation Model, Depth Aware and Edge Loss concepts together with the depth map prediction from single RGB images employing a Monocular camera. Section IV represents the Experimental Results obtained by evaluating the performance of proposed work and then comparing them with the current state-of-the-art papers. Finally, conclusion and future work area are stated in Section V.

II. RELATED WORKS

Monocular Depth Estimation is a task of estimating the relative depth of objects placed in a scene with respect to other objects or surfaces present in the same RGB images obtained from a monocular camera. It can be performed by many methods, most common of them are discussed as: CNN based methods [14] that contemplate task of depth estimation a regression model training which outputs a depth map corresponding to a one RGB image. Another work [15] used Encoder-Decoder primarily based structures that have created vital contributions in several vision related problems like image segmentation [16], optical flow estimation [17], and image reconstruction [18]. The encoder downsamples the image into bottleneck and also the decoder upsample bottleneck to the final output size of depth-map, most generally, one or more encoder-decoder networks are used as a subset of a wider network., followed by residual

978-1-6654-3897-1/21/$31.00 ©2021 IEEE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 09:04:46 UTC from IEEE Xplore. Restrictions apply.

connections between I_{th} layer and $(N\text{-}I)_{th}$ layer, where N is total number of layers.

The prototypic architecture for dense depth map estimation is square measured by fully-convolutional networks Fully-convolutional networks [19, 20]. Over the years, some variations on this particular pattern have been proposed. However, in order to learn multi-scale function representations, all current architectures use convolution and subsampling as fundamental components. The works proposed progressively upsample representations which are pooled at different stages, whereas others expanded convolutions recovers fine-grained predictions making certain sufficiently large context. Newer architectures [21] [22] maintain a high-resolution illustration beside multiple lower-resolution representations throughout the network.

It can also be solved by using Multi Task Learning [23], includes a knowledge transfer mechanism which improves the generalization of tasks by sharing the particular dominion of information between complementary tasks. During this sort of technique [13] that shares the corresponding knowledge of depth prediction together with to the linguistics segmentation that helps in classification of objects in several categories and shares the corresponding information in between training processes by using Information Sharing Units. Whereas, alternative strategies search the best descent direction of the gradient by adaptively choosing the coefficient factors throughout the training process. L.L. et al. [24] proposed a Multi-Layered Depth model to perform Depth estimation and semantic segmentation all together efficiently.

Attention-based models [25] and in particular Transformers [26] are set-to-set models depending on Multi-Head Attention mechanism which empowers the models to encode multiple relationships among the features corresponding to each and every layers and have been successful when are trained on very large datasets. These network architectures are vastly used for learning and training of Natural Language Processing [27]problems. It has been invented that the transformer architecture which have been extremely successful in NLP, are functional when applied to images in tokes format, i.e. an image is further divided into certain number of patches and they are numbered accordingly and are passed through the transformer network hence resulting in the feature training of model from created tokens and patches, and can yield competitive performance in image classification. For example, DPT Hybrid [26] is one of the recently proposed works which uses Vision Transformers [28] for the task of Monocular Depth Estimation.

III. PROPOSED WORK

Network architecture: The proposed Depth Estimation Model, DepthNet, as shown in Fig. 1, is capable of predicting the depth-map of any indoor (NYU V2 [1]) and outdoor (KITTI [2]) scene using single RGB image. Depth Estimation Model consists of a backbone made of ResNet50 Convolutional Neural Network which downsamples the feature map of the input image of size (240×320) to final size of (30×40) but as observed in the native ResNet50 [29] model the final feature map size is reduced to (8×10), but according to the requirement feature map's size is not reduced to that extent which makes it difficult for our final model to learn the corresponding features while resizing them back to our output shape of dimensions$(240 \times 320 \times 1)$. The benefit of using the ResNet50 model is the upskip connections which helps the model to remember the information shared in the complete long tailed network. Four up-convolutional layers were introduced to increase the size of the feature maps received form the output of ResNet50 by the ratio of (2: 2) so that each up-convolutional layer outputs doubles the size of the input given to it, i.e. $(30 \times 40) \rightarrow (60 \times 80) \rightarrow (120 \times 160) \rightarrow (240 \times 320) \rightarrow (480 \times 640)$.

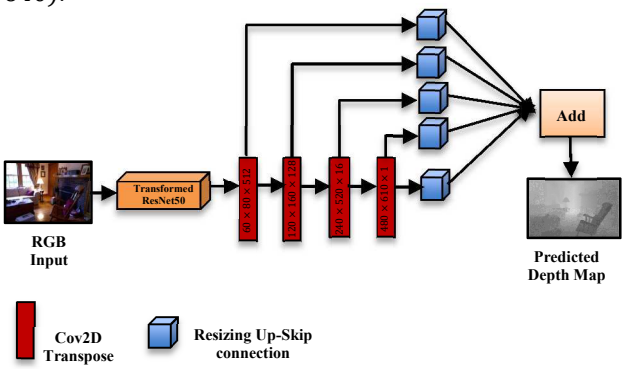


Fig. 1: A brief outline of the proposed specification, a RGB image is given as an input to the ResNet-50 architecture (Orange), deconvolutioned ResNet-50 output to the final size of depth map to be created (Red), besides resizing up-skip connections (Blue) to recollect all the antecedently shared information of the network.

On paying attention towards the output channels obtained from the customized ResNet50 those are 2048, and according to the required depth map to be of a grayscale image, channels of ResNet50 were reduced to 1. To do so, some of the optimum ratios were researched so that output scene gets converged without losing any information while reducing the channel count. Hence, there is a high possibility of the important feature information being lost, to prevent that five resizing up-skip connections are being used to reduce the output feature map size to$(240 \times 320 \times 1)$.

Table I: Layered structure of the proposed DepthNet architecture. The terms 'conv', 'trans-conv', 'relu' stands for Convolution Layer (Conv2D), Convolutional Transpose Layer (Conv2DTranspose) and 'Rectified Linear Unit' activation function respectively.

Serial Number	Layer	kernel	Strides	Input Layer	Output
I.	Input	-	-	-	240x320x3
II.	Transformed ResNet-50	-	-	I.	30x40x2048
III.	trans-conv + relu	(2, 2)	2	II.	60x80x512
IV.	trans-conv + relu	(2, 2)	2	III.	120x160x128
V.	trans-conv + relu	(2, 2)	2	IV.	240x320x64
VI.	trans-conv + relu	(2, 2)	2	V.	480x640x1
VII.	Resize Layer	(240, 320)	-	II.	240x320x2048
VIII.	conv + relu	(1, 1)	1	VII.	240x320x1
IX.	Resize Layer	(240, 320)	-	III.	240x320x512
X.	conv + relu	(1, 1)	1	IX.	240x320x1
XI.	Resize Layer	(240, 320)	-	IV.	240x320x128
XII.	conv + relu	(1, 1)	1	XI.	240x320x1
XIII.	Resize Layer	(240, 320)	-	V.	240x320x64
XIV.	conv + relu	(1, 1)	1	XIII.	240x320x1
XV.	Resize Layer	(240, 320)	-	VI.	240x320x1
XVI.	conv + relu	(1, 1)	1	XV.	240x320x1
XVII.	Add Layer	-	-	VIII, X, XII, XIV, XVI.	240x320x1

The Resizing Up-Skip Connections gets there inputs from ResNet50 and all four up-convolutional layers, resulting in

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 09:04:46 UTC from IEEE Xplore. Restrictions apply.

five outputs having the same size of feature map (240×320) and equal numbers of channels($equal\ to\ 1$), which is done by a resize layer which resizes the input feature map obtained from ResNEt-50 and up-convolutional layers to the output size of (240×320) followed by a convolutional layer which converts the channels of resizing layer output to 1, as seen in Fig. 2. These resizing up-skip connections are in general used for remembering the previously shared information that has the probability to be lost. At the end, these five resized outputs are added and as a result our model

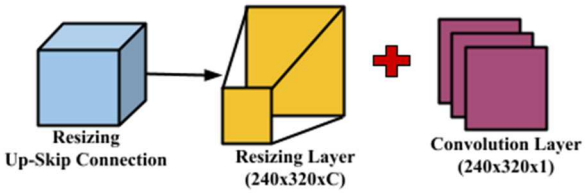


Fig. 2: The Resizing Up-Skip connections includes a resizing layer which resizes the output feature maps obtained from the inputs to (240×320) and then convolutional layer reduces the number of output channels present in the outputs procured from the resizing layers to 1.

will output this added information, and predict the relative depth between the two objects present in RGB images.

Losses: In order to better illustrate the custom losses, we tend to train the architecture for preserving the useful information obtained as features and these losses will help the network to train where the network should focus lots on. Here we use the heat map characteristics to show the final resulting feature maps of the network on predicted monocular depth, where the darker portion specify lower values (closer to zero) or closer objects and the brighter portions specify the higher values (closer to one) or the far away objects.

i) Depth aware loss: Depth Aware Loss is used to minimize the pixel-wise difference between the feature focusing terms (attention term) and regularization terms calculated by the predicted depth values and ground truth scene corresponding to the RGB images on the training datasets, which helps to supervise the monocular depth estimation task, defined in Eq. (1).

$$L_{DA} = 1/N\ (\sum(\phi_{Di}{}^{GT} - \psi_{Di})\ |\phi_{Di}{}^{GT} - \psi_{Di}|) \qquad (1)$$

where $\phi_{Di}{}^{GT}$ is an attention term, equal to normalized ground truth value, ψ_{Di} is a regularization term which is employed for learning of the features of the nearby objects and also tackles the gradient vanishing problems which leads to termination of edges of the distant objects, for example: an object rectangular in shape seems to be elliptical or oval in shape, throughout the complete duration of the training, as defined in Eq. (2):

$$\psi_{Di} = 1 - \frac{(min(log(d_i), log(d_i{}^{GT})))}{(max(log(d_i), log(d_i{}^{GT})))} \qquad (2)$$

Where $d_i$and $d_i{}^{GT}$ are predicted depth values and ground truth values respectively (unnormalized) and N is total numbers of pixel values in predicted depth map.

ii) Edge Loss: For architecture to focus expeditiously on the native gradients and edges of structures and on the surfaces present in the scene and preserve information regarding it, we tend to propose the edges in the output scenes obtained from the projected architecture and introduce the Edge Loss as shown in Eq. (3).

$$L_E = \frac{1}{N}(\sum(|d_{ix} - d_{ix}^{GT}| + |d_{iy} - d_{iy}^{GT}|)) \qquad (3)$$

Where N is total numbers of pixel values in predicted depth map, d_{ix} and are d_{iy} sobel gradients of the predicted depth scene d_i in x and y direction, defined in Eq. (4) and (5) below.

$$d_{ix} = d_i \times \begin{matrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{matrix} \qquad (4)$$

$$d_{iy} = d_i \times \begin{matrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{matrix} \qquad (5)$$

The main reason to use this Edge Loss is to differentiate and maintain a high inter-class separation among the different objects. Hence, it resolves the problem faced during the prediction of grayscale depth map when the objects placed next to each other and at same distance from the monocular camera, their depth map do not get merged with each other.

The final attention loss applied to the proposed Depth Estimation Model can be defined as the summation of the outputs of Depth-Aware Loss and Edge Loss, as given below, Eq. (6) i.e.

$$L_{Depth} = L_{DA} + L_E \qquad (6)$$

IV. Experimental Results

Experimental Details: The proposed DepthNet architecture is implemented in tensorflow. In the training process, an adaptive momentum optimizer (Adam) is used which can handle sparse gradients on noisy problems. The model is trained for a total of 70 epochs. For initial 50 epochs Combined Depth Loss is used and for next 20 epochs depth edge loss is used while training, to further specify the gradients in the predicted depth maps. The learning rate is set to e^{-4} and batch size taken 4 images for each step per epoch. One epoch took approximately 39 minutes with Tesla T4. The entire proposed model consists of 28M parameters, 23.5M parameters in ResNet50 backbone and remaining in CNN Layers. The training loss curves for both NYU Depth Datasets V2 and KITTI Depth Dataset are shown in Fig. 5. *Datasets*: To evaluate the performance of the proposed model, experiments are conducted on three publicly available datasets: *NYU Depth Dataset V-2* [1], *KITTI Depth Dataset* [2], *SUN RGB-D Dataset* [30]. All datasets consisting of RGB images captured from a monocular and stereo camera. *NYU Depth Dataset V-2* [1] provides images and depth charts for various indoor scenes shot at pixel resolution of(640×480), which includes images of Bedrooms, Bathrooms, Work Places, Kitchen, etc. The dataset contains fifty thousand and six hundred and fifty-four training and testing samples respectively. We trained our network on thirty-five thousand subset images. The sample images are shown in Fig. 3. *KITTI Dataset* [2] includes depth maps from projected LiDAR point clouds that were matched against the depth estimation from the stereo cameras. The depth maps have an upper bound of 80 meters. The network was trained on the twelve thousand subset images to generalize the results considering indoor and outdoor environments both. The sample frames are shown in Fig. 4. SUN RGB-D Dataset

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 09:04:46 UTC from IEEE Xplore. Restrictions apply.

[30], is a dataset of 10335 real RGB-D images of room scenes. Each RGB image has a corresponding depth and segmentation map. The training and testing sets contain 5285 and 5050 images, respectively. Subset of this dataset consists of NYU depth dataset, which were already used for training the model on indoor dataset, hence it was further used for testing.



Fig. 3: Some examples showing the monocular depth estimation performed real time on NYU Depth Dataset-V2 [1].

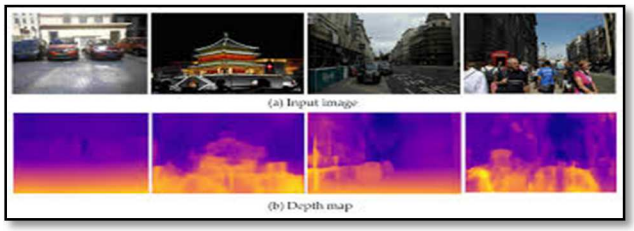


Fig. 4: Some examples showing the monocular depth estimation performed on KITTI Depth Dataset [2].

Evaluation Metrics : To evaluate the prediction quality of the proposed Depth Estimation methods, we used four evaluation metrics: Root Mean Square (RMS) [31] Error, Average Relative Error(REL) [32], Logarithmic Root Mean Square Error(RMSE_log) [31] and Squared Relative Error(Squared REL) [31] to check the similarity measure between predicted depth value and ground truth as defined in Eq. (7)-(10).

$$RMSE = \sqrt{(1/n(\sum(y_p - \hat{y}_p)))} \quad (7)$$

$$REL = 1/n\ (\sum(|y_p - \hat{y}_p|/y) \quad (8)$$

$$\text{Squared } REL = 1/n\ (\sum(|y_p - \hat{y}_p|\text{^}2\ /y) \quad (9)$$

$$RMSE_\log = \sqrt{(1/n\ (\sum(||log\ (y_p) - log\ (\hat{y}_p)||\text{^}2)))} \quad (10)$$

Where y_p, $\hat{y}_p$ and y are normalized ground truth value, predicted depth map pixel values and testing image's pixel value respectively, n is total numbers of pixel values in predicted depth map, $p \in (0,1)$.

The overall performance of the proposed Results work for KITTI depth dataset [2], NYU depth dataset V-2 [1] is given in Table II. The obtained results are also compared with the recent state-of-the-arts in Table III and IV for NYU depth dataset V2 [1] and KITTI depth dataset [2] respectively.

From Table III, it is clearly noticeable that the proposed work is outperforming the present state-of-the-art papers by a large margin. However, DPT-Hybrid [26] achieved 0.045 log10 value. Whereas, the proposed work is able to achieve 88.59 % better performance for RMS value and similarly 31.52 %for average relative error values. These results, as shown in Fig. 6, assure the depth estimation preciseness and accuracy of the proposed work on NYU depth dataset V2 [1].

From Table IV, it is concluded that the proposed work is outperforming the current state-of-the-arts by a large margin. However, AdaBins [33] achieved 0.058 average relative error value. Whereas, the proposed work is able to achieve 31.52 % better performance for Squared Relative Error value. These results, as shown in Fig. 4, assure the depth estimation preciseness and accuracy of the proposed work on KITTI depth dataset [2]. The qualitative results on the proposed work on testing datasets of NYU V2 and KITTI depth dataset and hence conclude that the final results are the standard and comparable to the ground truth images where the darker pixels represent the closer whereas lighter pixels represent the farther objects and boundaries respectively, as shown in Fig.3.

Table II: Performance of the proposed work for KITTI depth dataset [11] and NYU depth dataset V2 [10] datasets

Dataset	**RMSE** [31]	**REL** [32]	**Squared REL** [31]	**RMSE_log** [33]
KITTI [2]	0.102	0.079	0.063	1.218
NYU Depth V2 [1]	0.0407	0.0376	-	-

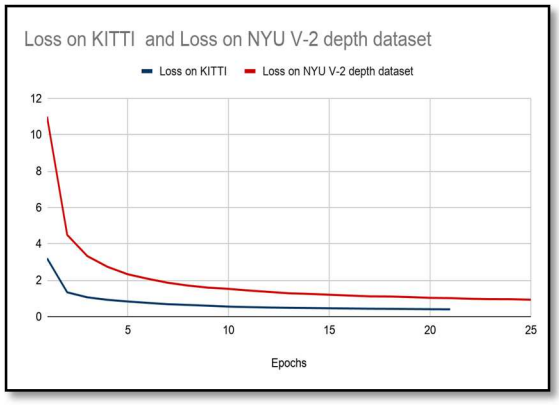


Fig. 5: Representation of continuously degrading loss curves on the KITTI depth dataset [2] and NYU Depth Dataset V2 [1].

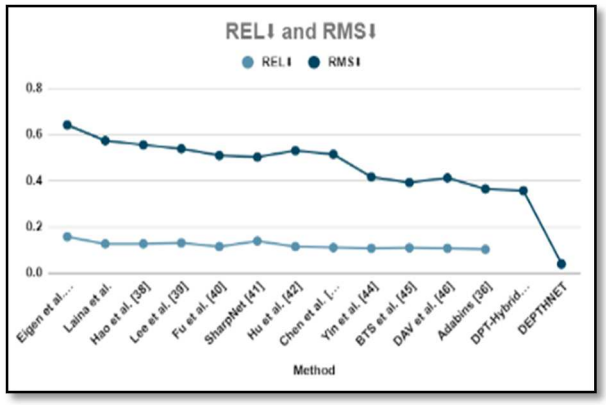


Fig. 6: Average Relative Error(REL) [32] and Root Mean Square(RMS) Error [31], evaluation metrics graph for NYU Depth Dataset V2 [1], showing Ours Model indicating less error values on testing data as compared to corresponding state-of-the-arts.

Table III: Quantitative results of proposed work on the NYU Depth Dataset V2 [1] in comparison with previously stated state-of-the-art papers.

Method	**REL↓**	**RMS↓**	**log10↓**
Eigen et al. [34]	0.158	0.641	-
Hao et al. [35]	0.127	0.555	0.053
Fu et al. [36]	0.115	0.509	0.051
SharpNet [37]	0.139	0.502	0.047
Hu et al. [38]	0.115	0.530	0.050
Chen et al. [39]	0.111	0.514	0.048

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 09:04:46 UTC from IEEE Xplore. Restrictions apply.

Yin et al. [40]	0.108	0.416	0.048
BTS et al. [41]	0.110	0.392	0.047
DAV et al. [42]	0.108	0.412	-
Adabins [33]	0.103	0.364	**0.044**
DPT-Hybrid [26]	-	0.357	0.045
DepthNet (our)	**0.0376**	**0.0407**	1.165

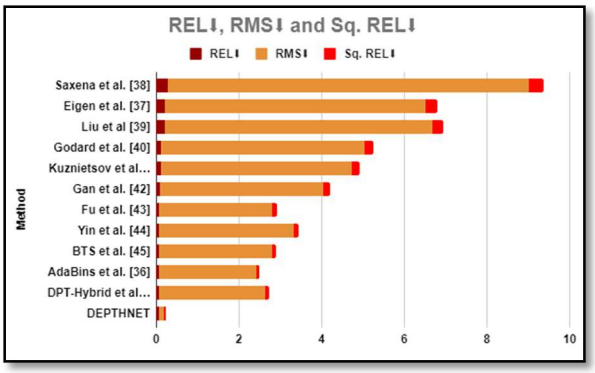


Fig. 7: Graphical comparison of the previously stated and the proposed network architecture on basis of Root Mean Square Error [31], Average Relative Error [32] and Squared Relative Error [31] on KITTI depth dataset [2], indicating that our model outperformed by giving less values for thee error than the corresponding state-of-the-art papers.

Table IV: Quantitative results of the projected work on the KITTI depth dataset [2] and comparing it with previously stated state-of-the-arts.

Method	**REL↓**	**RMS↓**	**Sq. REL↓**	**RMS_log↓**
Eigen *et al.* [34]	0.203	6.307	0.282	0.282
Liu *et al* [43]	0.201	6.471	0.273	0.273
Godard *et al.* [44]	0.114	4.935	0.206	0.206
Kuznietsov *et al.* [45]	0.113	4.621	0.189	0.189
Gan *et al.* [46]	0.098	3.933	0.173	0.173
Fu *et al.* [36]	0.072	2.727	0.120	0.120
Yin *et al.* [40]	0.072	3.258	0.117	0.117
BTS *et al.* [41]	0.059	2.756	0.096	0.096
AdaBins *et al.* [33]	**0.058**	2.360	0.088	**0.088**
DPT-Hybrid *et al.* [26]	0.062	2.573	0.092	0.092
DepthNet (our)	0.079	**0.102**	**0.063**	1.218

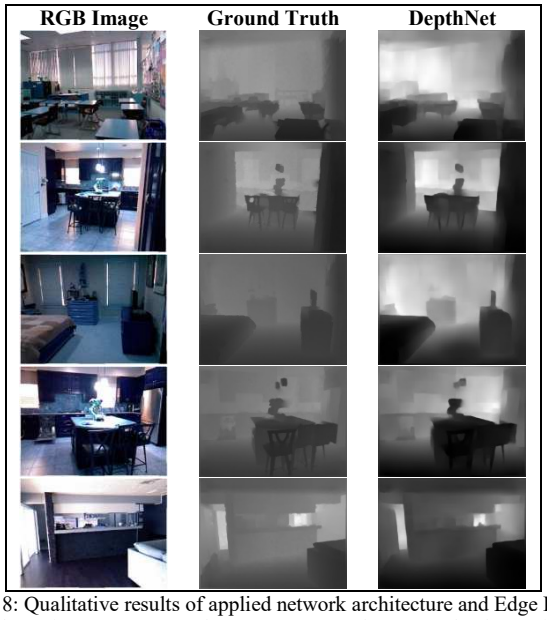


Fig. 8: Qualitative results of applied network architecture and Edge Loss while testing on NYU Depth Dataset V2 [1], the output depth map is at resolution$(240 \times 320 \times 1)$.

In case of testing of KITTI dataset, we represented the results in the plasma color map where, dark bluish color shows closer and as the pixels become reddish and yellowish it shows far away objects respectively, as expressed in Fig. 8.

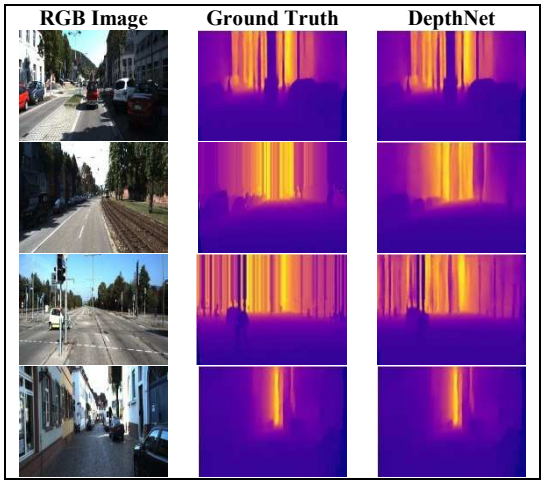


Fig. 9: Qualitative results of applied network architecture and Edge Loss while testing on KITTI Depth Dataset [2], the output depth map is at resolution$(240 \times 320 \times 1)$.

Considering the qualitative and quantitative results of the proposed architecture and loss algorithm, it can be said that the aim, to trace the corresponding features of a RGB frame to reconstruct a geometry as a depth map is fulfilled efficiently. As shown in the Table IV and Table V, the comparison results on the commonly used matrices outperforms that of AdaBins [33], the current state-of-the-art architecture.

V. CONCLUSION

On comparison with the current state-of-the-art papers, our proposed DepthNet architecture predicts the relative depth of an object with respect to another object or surface and the custom Edge Loss thereby differentiates them from their overlapping surfaces, and hence preserves the better dense depth maps for predicting on the two publicly available datasets, for indoor: NYU depth dataset V2 [1], as well as for outdoor: KITTI depth dataset [2], environments. The projected methodology is computationally economical and is more inclined towards the implementation in form of any mobile application or places wherever the weight constraints are applied, as an example, within the field of Autonomous Aerial or Ground Vehicles.

REFERENCES

[1] N. Silberman, D. Hoiem, P. Kohli and R. Fergus, "Indoor segmentation and support inference from rgbd images.," in *European Conference on Computer Vision*, 2012.

[2] A. Geige, P. Lenz and R. Urtasun, "Are we ready for autonomous driving? the kitti vision benchmark suite.," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2012.

[3] L. Nalpantidis and A. Gasteratos , "Stereo Vision Depth Estimation Methods for Robotic Applications.," in *Depth Map and 3D Imaging Applications: Algorithms and Technologies.*, 2012.

[4] K. Tateno, . F. Tombari, . I. Laina and N. , "CNN-SLAM: Real-time dense monocular SLAM with learned depth prediction.," in *Computer Vision and Pattern Recognition (CVPR)*, 2017.

[5] X. Wang, W. Yin, T. Kong, Y. Jiang, L. Li and C. Shen, "Task-Aware Monocular Depth Estimation for 3D Object Detection.," in *AAAI Conference on Artificial Intelligence*, 2020.

[6] Y. Duan, H. Deng and F. Wang, "Depth Camera in Human-Computer Interaction. Intelligent Networks and Intelligent

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 09:04:46 UTC from IEEE Xplore. Restrictions apply.

System," in *Fifth International Conference on Intelligent Networks and Intelligent Systems*, 2012.

[7] N. Appiah and N. Bandaru, "Obstacle detection using stereo vision for self-driving cars".

[8] Y. Wang, W.-L. Chao, D. Garg, B. Hariharan, M. Campbell and K. Q. Weinberger, "Pseudo-LiDAR from Visual Depth Estimation: Bridging the Gap in 3D Object Detection for Autonomous Driving," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019.

[9] V. Harisankar, V. V.V Sajith and K. P. Soman, "Unsupervised Depth Estimation From Monocular Images For Autonomous Vehicles," in *Fourth International Conference on Computing Methodologies and Communication (ICCMC)*, 2020.

[10] A. Sabnis and L. Vachhani, "Single image based depth estimation for robotic applications," in *IEEE Recent Advances in Intelligent Computational Systems*, 2011.

[11] Y. Wang, Z. Lai, G. Huang, B. H. Wang, L. van der Maaten, M. Campbell and K. Q. Weinberger, "Anytime Stereo Image Depth Estimation on Mobile Devices," in *International Conference on Robotics and Automation (ICRA)*, 2019.

[12] N. Zheng, Y. Shi, W. Rong and Y. Kang , "Effects of Skip Connections in CNN-Based Architectures for Speech Enhancement," in *Journal of Signal Processing Systems*, 2020.

[13] J. Jiao, Y. Cao, Y. Song and R. Lau, "Look Deeper into Depth: Monocular Depth Estimation with Semantic Booster and Attention-Driven Loss," in *European Conference on Computer Vision (ECCV)*, 2018.

[14] R. Garg, V. Kumar B.G, G. Carneiro and I. Reid, "Unsupervised CNN for Single View Depth Estimation: Geometry to the Rescue," in *European Conference on Computer Vision(ECCV)*, 2016.

[15] C. Godard, O. M. Aodha, M. Firman and . G. J. Brostow, "Digging Into Self-Supervised Monocular Depth Estimation," in *International Conference on Computer Vision (ICCV)*, 2019.

[16] F. Sultana, A. Sufian and P. Dutta, "Evolution of Image Segmentation using Deep Convolutional Neural Network.," in *arXiv:1907.11111v1*, 2019.

[17] J. Hur and S. Rot, " Optical Flow Estimation in the Deep Learning Age," in *Modelling Human Motion*, 2020.

[18] P.-Y. Liu and E. Y. Lam, "Image Reconstruction Using Deep Learning.," in *arXiv:1809.10410*, 2018.

[19] J. Long, E. Shelhamer and T. Darrel, "Fully Convolutional Networks for Semantic Segmentation," in *Computer Vision and Pattern Recognition (CVPR)*, 2015.

[20] J. Dai, Y. Li, K. He and J. Sun, "R-FCN: Object Detection via Region-based Fully Convolutional Networks," in *arXiv:1605.06409* , 2016.

[21] Y. Arora, I. Patil and T. Nguyen , "Fully Convolutional Network for Depth Estimation and Semantic Segmentation.," in *stanford.edu*.

[22] B. Kang, Y. Lee and T. Q. Nguyen, "Depth-Adaptive Deep Neural Network for Semantic Segmentation," in *IEEE Transactions on Multimedia*, 2018.

[23] Y. Zhang and Q. Yang, "A Survey on Multi-Task Learning," in *arXiv:1707.08114v3*, 2021.

[24] L. Liebel and M. Körner, "MultiDepth: Single-Image Depth Estimation via Multi-Task Regression and Classification," in *IEEE Intelligent Transportation Systems Conference (ITSC)*, 2019.

[25] D. Bahdanau, K. Cho and Y. Bengio, "Neural Machine Translation by Jointly Learning to Align and Translate," in *arXiv:1409.0473*, 2016.

[26] R. Ranftl, A. Bochkovskiy and A. Vladlen Koltun, "Vision Transformers for Dense Prediction.," in *arXiv:2103.13413v1*, 2021.

[27] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser and I. Polosukhin, "Attention Is All You Need.," in *arXiv:1706.03762v5* , 2017.

[28] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. ,. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit and N. Houlsby, "AN IMAGE IS WORTH 16X16 WORDS: TRANSFORMERS FOR IMAGE RECOGNITION AT SCALE," in *arXiv:2010.11929*, 2020.

[29] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," in *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.

[30] S. Song, S. P. Lichtenberg and J. Xiao, "SUN RGB-D: A RGB-D Scene Understanding Benchmark Suite.," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.

[31] A. Botchkarev, "Performance Metrics (Error Measures) in Machine Learning Regression, Forecasting and Prognostics: Properties and Typology," in *arXiv:1809.03006* , 2018.

[32] C. Chen, J. Twycross and J. M. Garibaldi, "A new accuracy measure based on bounded relative error for time series forecasting," in *PloS one.*, 2017.

[33] S. Farooq Bhat,, . P. Wonka and I. Alhashim, "AdaBins: Depth Estimation using Adaptive Bins," in *arXiv:2011.14141.*, 2020.

[34] D. Eigen, C. Puhrsch and R. Fergus, "Depth map prediction from a single image using a multi-scale deep network," in *ICLR*, 2014.

[35] Z. Hao, y. li, S. You and F. Lu, "Detail preserving depth estimation from a single image using attention guided networks.," in *International Conference on 3D Vision (3DV)*, 2018.

[36] H. Fu, . M. Gong, C. Wang, K. Batmanghelich and D. Tao, "Deep Ordinal Regression Network for Monocular Depth Estimation," in *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018.

[37] M. Ramamonjisoa and V. Lepetit, "Fast and accurate recovery of occluding contours in monocular depth estimation.," in *International Conference on Computer Vision*, 2019.

[38] J. Hu, M. Ozay, Y. Zhang and T. Okatani, "Revisiting single image depth estimation: Toward higher resolution maps with accurate object boundaries," in *IEEE Winter Conference on Applications of Computer Vision*, 2019.

[39] X. Chen, Z. -Jun Zha and . X. Chen, "Structure Aware residual pyramid network for monocular depth estimation," *arXiv:1907.06023.,* 2019.

[40] W. Yin, Y. Liu, C. Shen and Y. Yan, "Enforcing geometric constraints of virtual normal for depth prediction," in *IEEE/CVF International Conference on Computer Vision*, 2019.

[41] J. H. Lee, M.-K. Han, I. H. Suh and D. Wook Ko, "From big to small: Multi-scale local planar guidance for monocular depth estimation," in *arXiv:1907.10326*, 2019.

[42] L. Huynh, P. Nguyen-Ha, J. Matas, E. Rahtu and J. Heikkilä, "Guiding monocular depth estimation using depth-attention volume," in *European Conference on Computer Vision*, 2020.

[43] F. Liu, C. Shen, G. Lin and I. Reid, "Learning depth from single monocular images using deep convolutional neural fields," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016.

[44] C. Godard, O. M. Aodha and G. J. Brostow, "Unsupervised Monocular Depth Estimation With Left-Right Consistency," in *IEEE Conference on Computer Vision and Pattern Recognition* , 2017.

[45] Y. Kuznietsov, J. Stuckler and B. Leibe, "Semi-supervised deep learning for monocular depth map prediction," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.

[46] Y. Gan, X. Xu, W. Sun and L. Lin, "Monocular depth estimation with affinity, vertical pooling, and label enhancement," in *European Conference on Computer Vision (ECCV)*, 2018.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 09:04:46 UTC from IEEE Xplore. Restrictions apply.

J Mater Sci: Mater Electron

Down-shifting photoluminescence studies of thermally stable Dy^{3+} ions doped borosilicate glasses for optoelectronic device applications

Rajat Bajaj[1], Aman Prasad[2], A. V. S. Yeswanth[3], Pooja Rohilla[1], Sumandeep Kaur[1], and A. S. Rao[1,*]

[1] *Department of Applied Physics, Delhi Technological University, Bawana Road, New Delhi 110042, India*
[2] *Department of Physics and Computer Science, Dayalbagh Educational Institute, Agra 282005, India*
[3] *Department of Electronics, Bhagwan Parushuram Institute of Technology (BPIT), Dr. K. N. Katju Marg, Rohini Sector 17, New Delhi 110089, India*

Received: 7 October 2021
Accepted: 25 December 2021

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

ABSTRACT

In this research work, Dy^{3+} doped alkali zinc alumino borosilicate (AZABS) glasses have been prepared via melt quenching technique. A series of AZABS glasses of varying concentrations of dysprosium (Dy^{3+}) (0.1–2.5 mol%) was prepared. It was found that under UV excitation, 0.5 mol% Dy^{3+} doped glass exhibited maximum luminescence intensity. Subsequent photoluminescence studies like emission/excitation spectra, temperature-dependent photoluminescence and decay kinetics were also performed. The value of direct bandgap as calculated from Tauc Plot for 0.5 mol% of Dy^{3+} ions doped AZABS glass is 2.85 eV. Yellow to blue ratio and colorimetry analysis have showed that the as-prepared glasses are capable of producing cool white light. Application of Dexter theory showed that the energy transfer mechanism between the dopant ions in the glass matrix was of dipole–dipole nature. It is proposed that these Dy^{3+} doped AZABS glasses can be used as prospective materials in optoelectronic device applications such as solid-state lighting and w-LEDs.

1 Introduction

It is well known that rare earth (RE) ions doped materials have elicited immense interest in researchers across the globe because of their superior and far-reaching applications in the areas of photovoltaic devices, lasers, displays, photodynamic therapies, drug deliveries, etc. [1–11]. The most persistent research of RE doped materials has taken place in the areas of w-LEDs and solid-state lighting (SSL) applications. The energy requirements of humankind have far exceeded than what we can regenerate. This is putting tremendous pressure on mother earth, thereby leading to many serious environmental hazards and repercussions for our future generations. Therefore, it is urgently required to develop such

Address correspondence to E-mail: drsrallam@gmail.com

https://doi.org/10.1007/s10854-021-07667-8
Published online: 23 January 2022

environment friendly technologies that help to reduce the ecological impact of human advances so that we leave a cleaner and greener earth for our upcoming generations. SSL devices have shown the way forward in this area. The energy consumption of artificial light sources account for 1/6 to 1/5 of the total global electricity production. Today, the average artificial light quantity per capita is 600 times higher than what it was in the nineteenth century. In the USA itself, expenditures for lighting can be cut down drastically by $100 billion dollars over the period 2000–2020 [12, 13]. Therefore, development of cost effective and environment friendly lighting devices can indeed help reduce carbon emissions by around 28 million metric tons annually [14]. Developments in SSL technologies have paved the way for research in w-LEDs that are far superior to conventional lighting sources like incandescent lamps and tube lights. This is due to the fact that w-LEDs have high efficiency, longer life span and above all, consume much less energy [15].

Now the conventional approach to construct a w-LED was to use an amalgamation of YAG:Ce^{3+} phosphor with a blue InGaN LED chip. But the major disadvantage of using this approach is the lack of red component in the emission which leads to a low colour rendering index (CRI), low thermal quenching temperature and halo effect [6–8]. Another approach is to fabricate the w-LEDs by mixing RGB phosphors that can be excited by UV LED. This phosphor needs to be encapsulated in an epoxy region made up of a polymer material. The luminescence efficiency of the phosphor depends on the thickness of material and the concentration of the phosphor. But it has been observed that this epoxy region deteriorates under high temperature, high excitation and high energy sources [16]. To counter the aforementioned disadvantages, RE doped glassy systems are being developed. This is due to the fact that there is no epoxy region involved in the manufacturing process. Also, the sheer fact that glasses have a low production cost, simpler manufacturing procedure, homogenous light emission, higher susceptibility to heat and a homogenous light emission makes them more efficient than phosphors for w-LEDs/SSL applications [7, 17].

Boron Trioxide (B_2O_3) is known to be the most proficient glass former as it leads to the formation of glasses with high thermal stability, high transparency, lower melting points and high RE ion solubility [18–22]. But the addition of B_2O_3 in the glass matrix leads to high phonon energies of the host lattice, resulting in more non-radiative energy losses, thereby leading to loss of luminescence efficiencies [14, 23]. Therefore, metal oxides such as SiO_2 are mixed with the borate glasses to lower this high phonon energy of the lattice thereby leading to increase its efficiency as a luminescent device [10, 24, 25]. Al_2O_3 in the host matrix helps by acting as a good network modifier and imparts excellent chemical and mechanical strength and also leads to emission property enhancement of the glass [23, 26]. Na_2O helps in enhancing the glassy nature characterization as it alters the functional group of boron, thereby changing its coordination [27]. ZnO helps in enhancing the optical and mechanical properties by increasing the thermal stability, chemical durability and reducing the crystallization of the glassy system [27, 28]. All the above highlighted characteristic features of the constituent chemicals such as B_2O_3, SiO_2, Al_2O_3, ZnO and Na_2O motivated us to prepare a glass system namely alkali zinc alumino borosilicate (AZABS) glasses to understand its practical utility in optoelectronic device applications.

RE doped glasses have found immense use in luminescence applications due to the fact that RE ions have a 4f configuration, emitting in the visible/NIR region under UV excitation. This has helped RE doped glasses find applications in various optoelectronic, laser, w-LED and SSL applications. Furthermore, RE doped glasses have better thermal and chemical stability, high doping capacity and low manufacturing cost which makes them superior to conventional phosphors [26, 29–32]. Dy^{3+} ions have a $4f^9$ configuration and give visible emission between f and f transitions. They are capable of acting as potential single phase white phosphors [29]. Under UV excitation, Dy^{3+} ions give out intense emissions in the blue and yellow region corresponding to $^4F_{9/2} \rightarrow {}^6H_{15/2}$ and $^4F_{9/2} \rightarrow {}^6H_{13/2}$ transitions, respectively. White light emission from such glassy materials doped with Dy^{3+} can be tuned by varying the yellow to blue (Y/B) intensity ratio [29].

In this paper, Dy^{3+} doped AZABS glasses have been prepared. The aim of this research work is to synthesize such glassy materials that have high luminescence efficiency along with high thermal and chemical stabilities. This is the fundamental requirement in the development of various lighting technologies that are cost effective and less carbon

emitting as has been mentioned earlier. Various morphological and photoluminescence studies have been carried out on the as-prepared glass samples. Dexter Theory has been applied to study the energy transfer mechanisms between the dopant ions in the glass matrix. The conclusive results from these studies assure us that Dy^{3+} doped AZABS glasses can be utilized for optoelectronic device applications such as w-LED and SSL devices.

2 Experimental

Quick melt quenching technique [32] was used to synthesis the Dy^{3+} ions doped AZABS glasses with the composition details as given below:

$$35B_2O_3 - 20SiO_2 - (15 - x)Al_2O_3 - 15ZnO - 15Na_2CO_3 - xDy_2O_3$$
where $(x = 0.0, 0.1, 0.5, 1.0, 1.5, 2.0, 2.5\ \text{mol}\%)$.

Starting powders like boric acid (99.5%), silicon dioxide (99.9%), aluminium oxide (99.9%), zinc oxide (99%), sodium carbonate (99.9%) and dysprosium oxide (99.99%) were purchased from Sigma-Aldrich and in total, 8 g of sample was weighed in stoichiometric ratio. The powder was mixed in an agate mortar until a homogeneous mixture was formed. Further the powder was placed in an alumina crucible and heated by keeping it in an electric furnace at 1250 °C for 2 h. Thereafter, the melt was poured in between two pre-heated brass plates and further annealing was done at ~ 350 °C for another 2 h to avoid the flaws caused by thermal and mechanical stress. All these steps of glass preparation are summarized in form of flow chart in Fig. 1. Finally, the glass obtained was transparent and circular in shape with uniform thickness 2.2 mm and diameter about 1.5–2.0 cm. These glasses are labelled as AZABS 0.1 to AZABS 2.5 on the basis of mol% of Dy^{3+} ion concentration ranging from 0.1 to 2.5 mol%, respectively. The X-ray diffraction (XRD) of an undoped AZABS glass sample was recorded by using Bruker D8 Advance Diffractometer with nickel filtered CuKa radiation (1.5406 Å) (under operating conditions of tube voltage of 40 kV and tube current of 40 mA) with a scanning speed of 0.02 s^{-1}, and angle of diffraction ranges from 10° to 80°. The absorption spectral recordings were made on a Jasco V-770 Spectrophotometer. Photoluminescence (PL) emission and excitation spectra were recorded on a Jasco made FP-8300 spectrofluorometer having

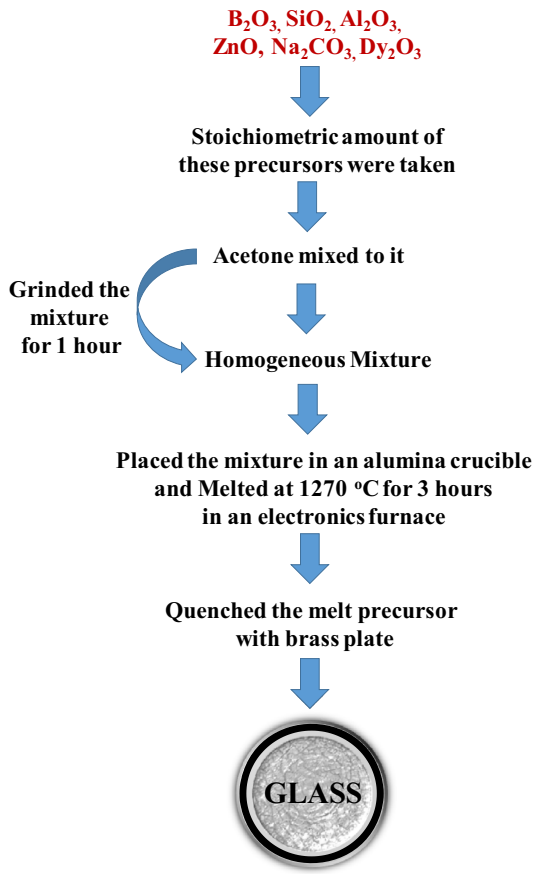


Fig. 1 Flow chart of all the steps involved in glass preparation

1.0 nm high resolution under Xenon excitation having an excitation range from 200 to 600 nm and lamp power of 150 W. The temperature-dependent PL studies were carried out by Ocean optics Flame-S-XR1-ES. The PL decay spectral studies were conducted on an Edinburgh FL920 Fluorescence lifetime spectrometer having a spectral resolution of 0.1 nm under Xenon excitation. All the measurements were recorded at room temperature except temperature-dependent studies.

3 Results and discussion

3.1 X-ray diffraction studies and thermal parameters

Figure 2 shows the X-ray diffraction patterns of (a) an undoped AZABS glass and (b) Dy^{3+} doped glasses with variable concentration of Dy^{3+} ions. It can be easily seen from the figure that a single broad hump exists in the observed patterns. This broadness

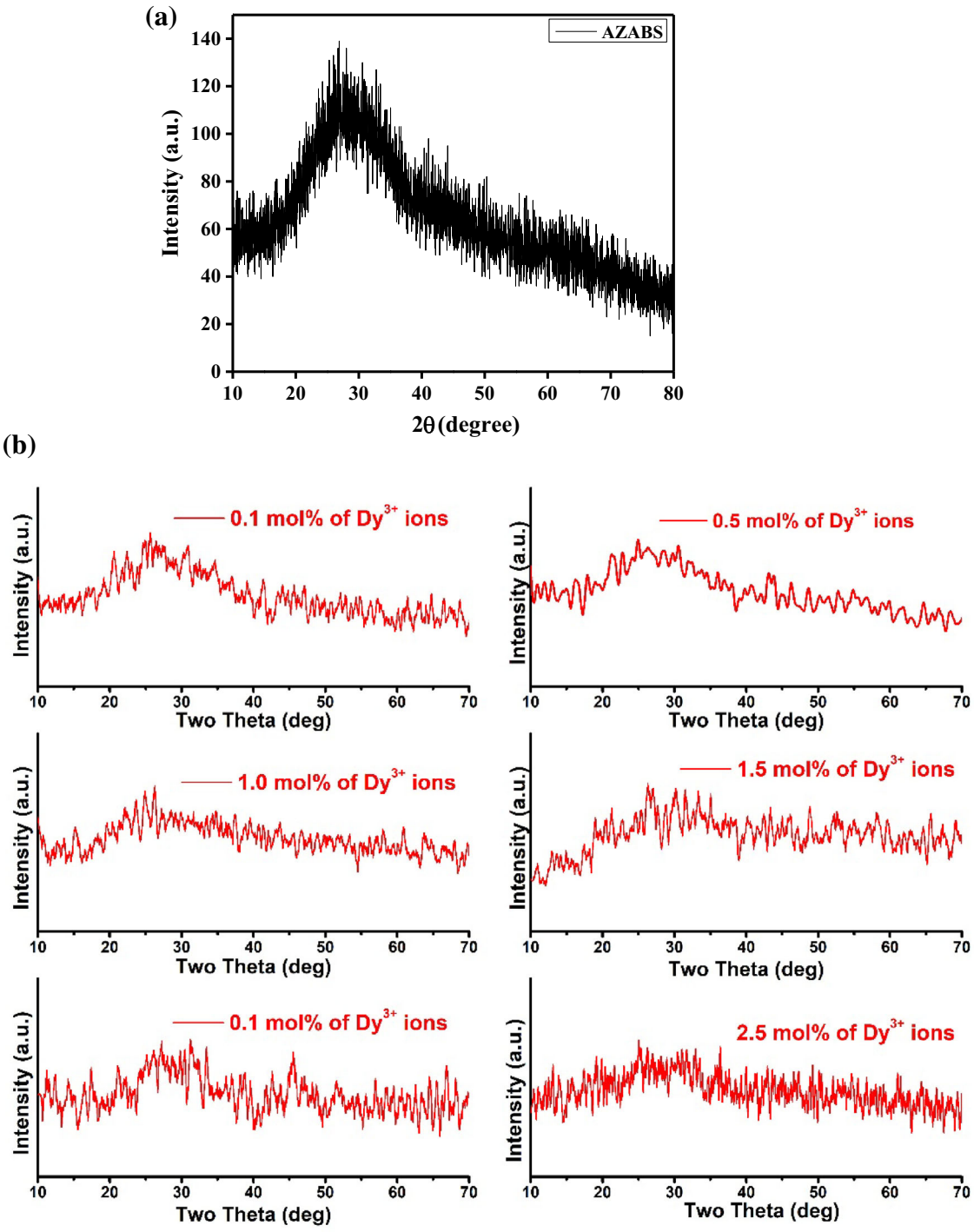


Fig. 2 XRD pattern of **a** an undoped AZABS glass and **b** Dy^{3+} doped glasses with variable concentration of dysprosium ions

actually confirms the amorphous nature of the host material signifying absence of any long range structural order in the glassy matrix [33].

Thermal parameters for an undoped AZABS glass have already been reported by us in our previous research work [32]. No observable change in TGA/DSC spectra and corresponding numerical values of peak crystallization temperature (T_c), melting temperature (T_m), glass transition temperature (T_g), onset crystallization temperature (T_x) and Hurby's thermal stability parameter (K_H) was observed. As has been reported earlier by us, the AZABS glass is very much

thermally stable and can be used in lighting applications.

3.2 Absorption spectral analysis

The absorption spectrum of 0.5 mol% Dy^{3+} doped AZABS glass has been provided in Fig. 3. The spectrum has been recorded in the UV–VIS and NIR region and shows the bands originating from the ground $^6H_{15/2}$ state. As is evident from the graph, lot many bands are present denoting transitions from $^6H_{15/2} \rightarrow {}^7F_{11/2} + {}^6H_{9/2}$, $^6F_{9/2} + {}^6H_{7/2}$, $^6F_{7/2} + {}^6H_{5/2}$, $^6F_{5/2}$, $^6F_{3/2}$, $^4F_{9/2}$, $^4I_{15/2}$, $^4G_{11/2}$, $^4F_{7/2} + {}^4I_{13/2}$, $^6P_{5/2} + {}^4M_{19/2} + {}^4$ (P, D)$_{3/2}$, $^6P_{7/2}$ and $^4M_{17/2} + {}^6P_{3/2}$. These band values match with the ones reported by Carnall et al. [34]. Also, it can be easily seen that bands corresponding to transitions $^6H_{15/2} \rightarrow {}^6P_{5/2} + {}^4M_{19/2} + {}^4$ (P, D)$_{3/2}$ and $^6H_{15/2} \rightarrow {}^4G_{11/2}$ are very weak. This is due to the fact that the host lattice of the glass possesses strong absorption in this region. On the contrary, the transition from $^6H_{15/2} \rightarrow {}^7F_{11/2} + {}^6H_{9/2}$ is highly intense. This is an hypersensitive transition obeying selection rules $|S| = 0$, $|L| \leq 2$, $|J| \leq 2$ [14, 35, 36].

Optical bandgap of 0.5 mol% of Dy^{3+} ions doped AZABS glass was obtained via Tauc plot. The Tauc equation is given as below [37]:

$$\alpha h\nu = C(h\nu - E_g)^n \quad (1)$$

where α is the coefficient of absorption and E_g is the energy band gap, $h\nu$ is the photon energy, C is energy independent constant and n denotes values specific

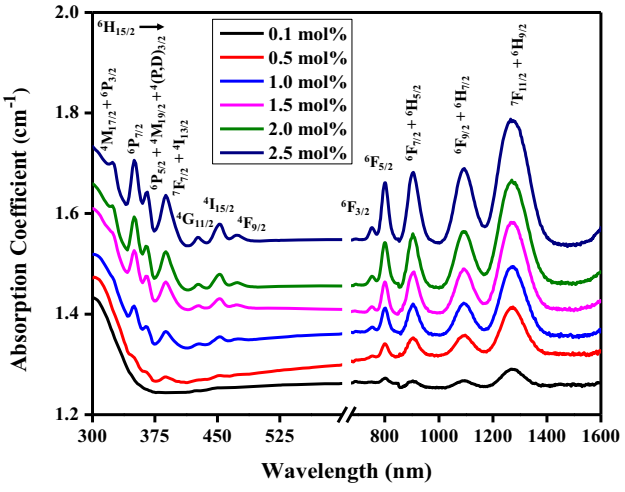


Fig. 3 Absorption of x mol% Dy^{3+} ions (x = 0.1 to 2.5 mol%) doped AZABS glasses

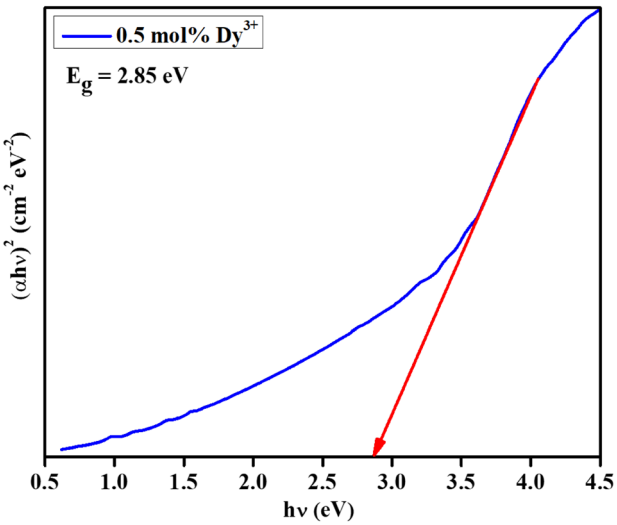


Fig. 4 Tauc plot for 0.5 mol% of Dy^{3+} ions doped AZABS glass

to various transitions, i.e. n = 1/2, 2, 1/3 and 3 for direct allowed, indirect allowed, direct forbidden and indirect forbidden transitions [38]. Figure 4. shows the Tauc plot for 0.5 mol% of Dy^{3+} ions doped AZABS glass. The value of direct bandgap obtained from the plot by linear fit in the linear region of Tauc plot is 2.85 eV.

3.3 Photoluminescence (PL) spectral studies

3.3.1 *Optimization of Dy^{3+} ions concentration in AZABS glass*

A series of Dy^{3+} doped AZABS (Dy^{3+} = 0.1 mol%, 0.5 mol%, 1 mol%, 1.5 mol%, 2.0 mol% and 2.5 mol%) glasses was prepared to look for the optimum concentration of Dy^{3+} in the glass sample which gives the highest luminescence peak. The emission spectra of this series under 350 nm excitation is given in Fig. 5. Two sharp bands are seen corresponding to $^7F_{9/2} \rightarrow {}^6H_{15/2}$ (482 nm) and $^7F_{9/2} \rightarrow {}^6H_{13/2}$ (575 nm) transitions, the intensity of which increases till 0.5 mol% concentration of Dy^{3+} and then decreases, i.e. concentration quenching occurs after 0.5 mol%. This implies that beyond 0.5 mol%, most of the energy from the host lattice is lost non-radiatively i.e., via collisions amongst the dopant ions. So we have chosen 0.5 mol% Dy^{3+} doped AZABS glasses for our PL studies. The inset of Fig. 5 shows the intensity versus concentration graph for the as-prepared glasses.

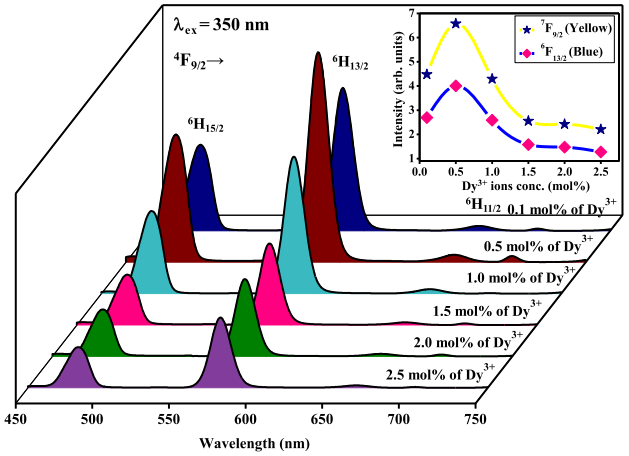


Fig. 5 PL emission spectra of varying Dy^{3+} activator ions concentration (x = 0.1 to 2.5 mol%) doped AZABS glasses at λ_{ex} = 350 nm. The inset diagram shows intensity vs concentration graph

3.3.2 *PL excitation & emission spectral analysis*

Figure 6 shows the PL excitation spectra of 0.5 mol% Dy^{3+} doped AZABS glass under emission wavelength of 575 nm. Six bands are seen in the spectra corresponding to the transitions from $^6H_{15/2} \rightarrow {}^6P_{3/2}$ (325 nm), $^6P_{7/2}$ (350 nm), $^6P_{5/2} + {}^4M_{19/2}$ (365 nm), $^4F_{7/2} + {}^4I_{13/2}$ (385 nm), $^4G_{11/2}$ (425 nm) and $^4I_{15/2}$ (453 nm) [39, 40]. Amongst all the excitation transitions, the most prominent one is $^6H_{15/2} \rightarrow {}^6P_{7/2}$ corresponding to 350 nm wavelength. So this wavelength was chosen to record the emission spectra. The fact that the AZABS glass shows sharp excitation bands from the n-UV to blue region proves that these glass samples can be used for w-LEDs applications under UV/n-UV excitation.

The PL emission spectra is given in Fig. 6 under 350 nm excitation show three excitation bands corresponding to transitions from the energy level $^4F_{9/2}$ to the levels $^6H_{15/2}$, $^6H_{13/2}$ and $^6H_{11/2}$ attributed to the 482 nm, 575 nm and 665 nm wavelength, respectively. The intensity of $^4F_{9/2} \rightarrow {}^6H_{13/2}$ transition is the highest. This transition is an electric dipole in nature and the transition from $^4F_{9/2} \rightarrow {}^6H_{15/2}$ is a magnetic dipole in nature following the selection rules $S = 0$, $L = 2$ and $J = 2$ [14]. The Y/B ratio between this electric dipole transition and magnetic dipole transition for the Dy^{3+} 0.1 to 2.5 mol% doped AZABS glass was calculated to be 1.61 to 1.75. The values of Y/B ratio for all the glasses is greater than unity which is the desired value to obtain white light. Relatively strong yellow peak over blue is due to low symmetry sites with no inversion centre around the Dy^{3+} ions and high covalence activity between Dy^{3+} ions and oxygen ligands. Therefore, the Y/B ratio for Dy^{3+} is used as an effective tool to probe the structural nature of the luminescent glass. [41]

The partial energy level diagram showing emission and excitation processes along with cross-relaxation mechanisms is shown in Fig. 7. As is evident from the diagram after excitation under 350 nm, the radiative transitions take place from $^4F_{9/2}$ level giving rise to the observed three bands. The self-quenching

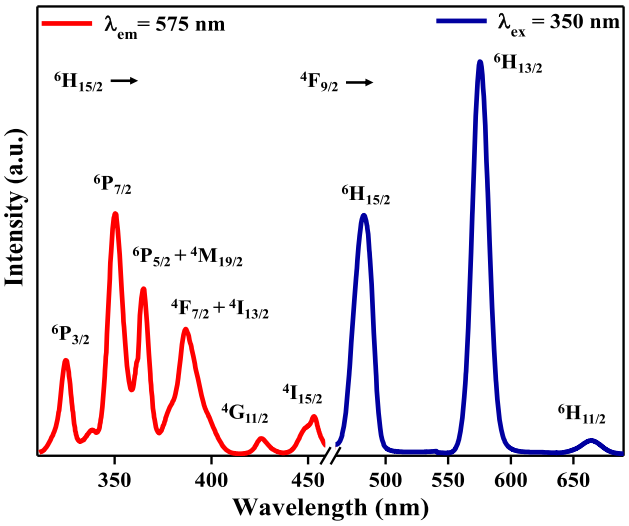


Fig. 6 PL excitation and PL emission spectrum of 0.5 mol% of Dy^{3+} ion doped AZABS glasses recorded at λ_{em} = 575 nm and λ_{ex} = 350 nm

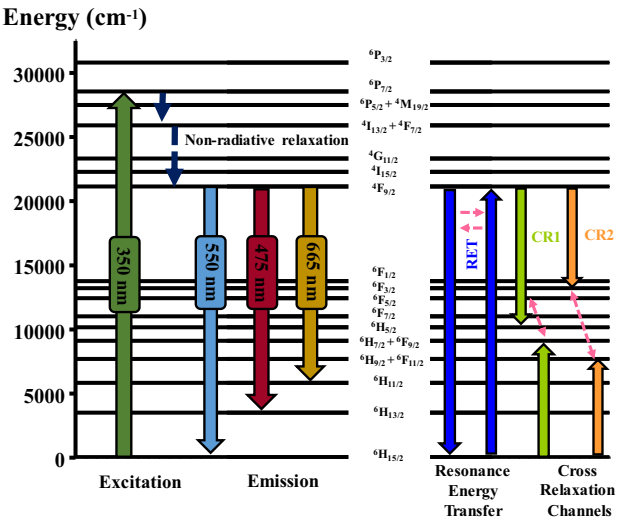


Fig. 7 Schematic energy level diagram of Dy^{3+} ions doped AZABS glasses

observed in Fig. 5. can be explained via resonant energy transfer mechanism (RET) [42, 43]. The possible non-radiative, cross-relaxation channels along with RET are given as [14, 44]:

$$\text{CR1}: {}^4F_{9/2} + {}^6H_{15/2} \rightarrow {}^6H_{5/2} + ({}^6H_{7/2}, {}^6F_{9/2})$$

$$\text{CR2}: {}^4F_{9/2} + {}^6H_{15/2} \rightarrow {}^6F_{3/2} + ({}^6H_{9/2}, {}^6F_{11/2})$$

$$\text{RET}: {}^4F_{9/2} \rightarrow {}^6H_{15/2}$$

3.3.3 Dexter theory analysis

To understand the mechanism of energy transfer, Dexter theory was applied to the emission data of the as prepared Dy-doped glasses by using the equation [45]:

$$\log(I/x) = K' - Q\log(x)/3 \quad (2)$$

$$\text{And } K' = \log k - \log \beta \quad (3)$$

Here K' is an independent constant, I is intensity and x is mol% concentration of activator ion. The value of Q determines the type of interaction responsible for the energy transfer between the neighbouring ions. For $Q = 6$, dipole–dipole interaction is responsible for energy transfer and for $Q = 8$ and 10, dipole–quadrupole and quadrupole–quadrupole interaction is responsible, respectively. Figure 8. shows the plot between $\log(I/x)$ and $\log(x)$. The value of Q was calculated by using the slope value of the given graph and it came out 5.37, which is close to 6. So it can be concluded that dipole–dipole interaction was responsible for the energy transfer due to which concentration quenching took place.

3.3.4 PL decay spectral analysis

The PL decay profiles of the AZABS glass samples for 575 nm emission wavelength at 350 nm excitation are given in Fig. 9. The PL decay profiles are used to calculate the lifetimes of ${}^4F_{9/2}$ level. It is evident from the figure that the decay curves show bi-exponential behaviour for all the concentrations of Dy^{3+}. The bi-exponential function signifies that ion-ion interactions are present in the system along with energy transfer giving rise to two components of emission: a slow component and a fast one. There exists a possibility of a different decay for the lanthanide ion. These ions might be at the surface of the glass or within the core of the glassy matrix [46, 47]. The experimental lifetimes calculated for the as-prepared Dy^{3+}-doped AZABS glass (x = 0.1 mol% to 2.5 mol%) are 7.87 ms, 6.84 ms, 6.72 ms, 5.57 ms, 5.38 ms and 5.03 ms, respectively, as shown in Table 1. It is seen that the values of the lifetimes decrease with the increase in the concentration of the Dy^{3+} ions. This happens due to the loss of energy via energy transfer between the Dy^{3+} ions themselves. This is also known as quenching. The high lifetime of these glasses indicates that they are well suited for lighting and other display applications.

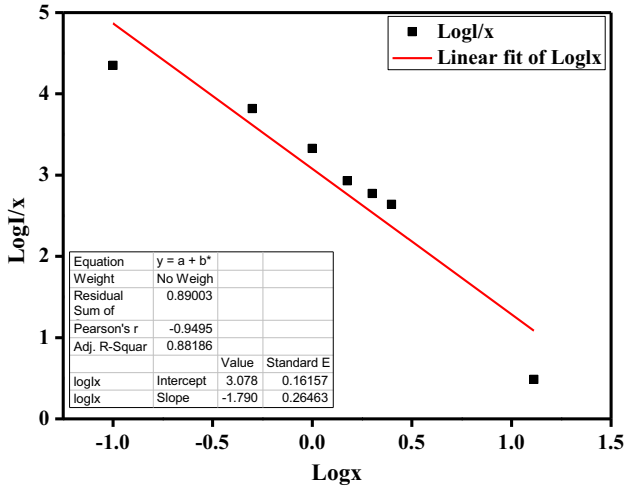


Fig. 8 Plot between $\log(x)$ vs $\log(I/x)$ for Dy^{3+} ions doped AZABS glasses

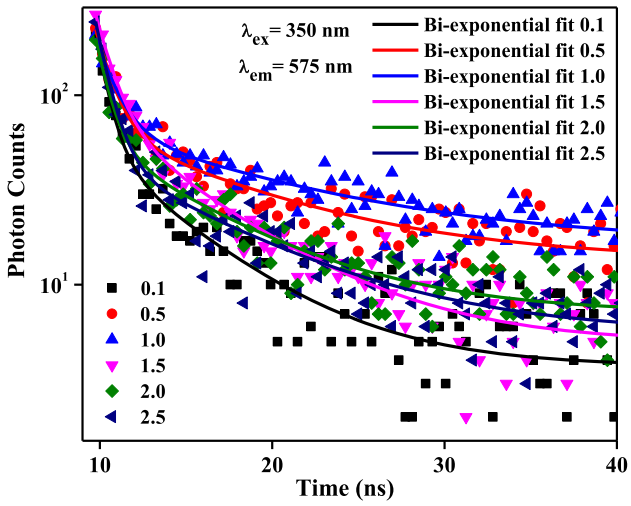


Fig. 9 Luminescence decay curves of x mol% of Dy^{3+} ions (x = 0.1 to 2.5 mol%) doped AZABS glasses recorded under 350 nm excitation wavelength

Table 1 CIE chromaticity coordinates (x, y), CCT and average decay time of Dy^{3+} ions doped AZABS glasses under 350 nm excitation

Sample ID AZABS: x mol%Dy^{3+}	CIE (x, y) at λ_{ex} = 350 nm	CCT (K) at λ_{ex} = 350 nm	t_{avg} (ms)
x = 0.1 mol%	(0.3759, 0.4098)	4314.72	7.87
x = 0.5 mol%	(0.3776, 0.4117)	4282.46	6.84
x = 1.0 mol%	(0.3764, 0.4117)	4310.69	6.72
x = 1.5 mol%	(0.3716, 0.4074)	4409.00	5.57
x = 2.0 mol%	(0.3724, 0.4059)	4383.76	5.38
x = 2.5 mol%	(0.3714, 0.4041)	4399.63	5.03

3.3.5 *Colorimetry analysis*

The PL emission spectra of AZABS glass doped with 0.1 to 2.5 mol% of Dy^{3+} ions was studied for its PL properties under 350 nm wavelength excitation. For the recorded spectra, CIE colour coordinates were calculated and have been shown in Table 1 and Fig. 10. It is evident from the CIE diagram that these coordinates fall in the white light region of the diagram.

Correlated colour temperature (CCT) can further be used to check the luminescence features of the material. It is simply defined as the temperature of a Planckian black body whose colour coordinates closely resemble the standard white light source. CCT expression as given by McCamy is [48]:

$$\mathrm{CCT} = -449n^3 + 352n^2 - 6823.2n + 5520.3 \qquad (4)$$

Here the value of n is given by $[(x - x_e)/(y - y_e)]$ where x_e = 0.332 and y_e = 0.186. The CCT values for the Dy-doped AZABS glass is given in Table 1. It is known that lamps with CCT < 3200 K give warm white light, whereas lamps with CCT > 4000 K give cool white light. The as-prepared glasses have CCT values greater than 4000 K [40]. This fact along with the CIE data proves that Dy^{3+} doped AZABS glasses can indeed be used for w-LED and SSL applications.

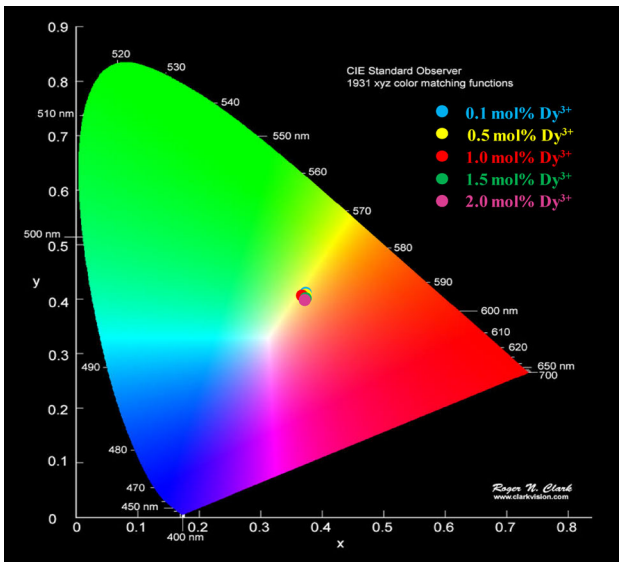


Fig. 10 CIE diagram of x mol% Dy^{3+} activator ion concentration (x = 0.1 to 2.5 mol%) doped AZABS glasses at λ_{ex} = 350 nm

3.3.6 *Temperature-dependent PL studies*

In applications of SSL and w-LEDs, where operating temperature can increase beyond 470 K, it is important to know the thermal stability of the material that is to be deployed [49]. The temperature-dependent emission spectra of 0.5 mol% of Dy^{3+} ions doped AZABS glass has been shown in Fig. 11. The Fig. 12 shows the tendency of the $^4F_{9/2} \rightarrow {}^6H_{13/2}$ transition to lose intensity with increase in temperature. It is clear from the spectra that the luminescence intensity decreases with increase in temperature from 30 °C till 200 °C due to the thermal quenching effect. Till 473 K, the intensity reduces to 75% of its value at 298 K.

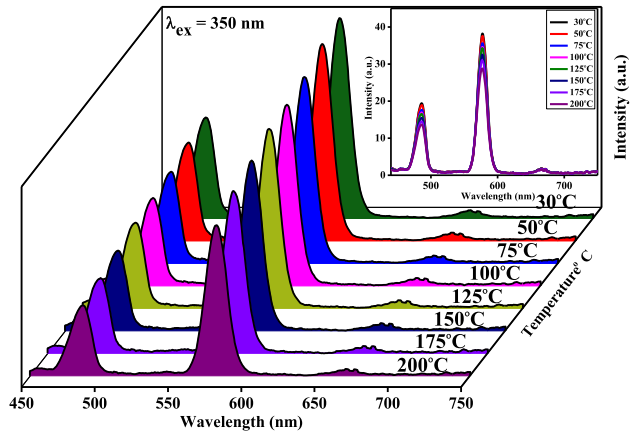


Fig. 11 Temperature-dependent emission intensity variation of 0.5 mol% of Dy^{3+} ions doped AZABS glass under 350 nm excitation with rise in temperature from RT to 200 °C

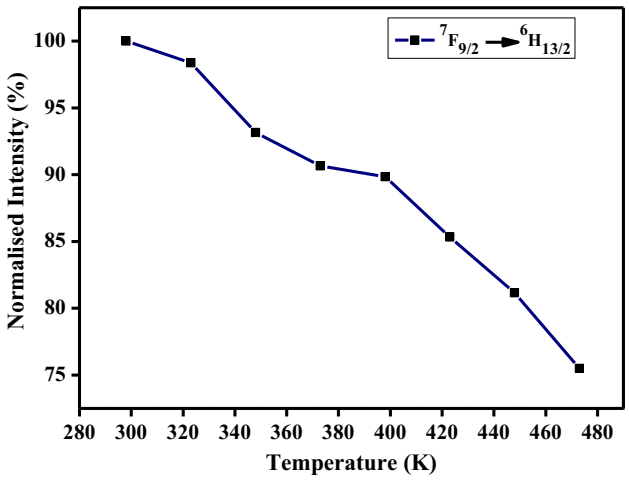


Fig. 12 Plot shows the normalized PL intensity with rise in temperature range of 298–473 K

The thermal quenching effect in luminescence intensity is observed due to phonon interaction in which the excited luminescence centre is thermally activated through a crossing point between the ground and excited states [50]. The activation energy required for thermal quenching can be calculated via Arrhenius Equation [51, 52]:

$$I(T) = \frac{I_0}{1 + c\exp\left[-\left(\frac{\Delta E}{KT}\right)\right]} \quad (5)$$

where I_0 is the initial intensity, $I(T)$ is the intensity at given temperature T, ΔE is the activation energy for thermal quenching and K is the Boltzmann constant (8.629×10^{-5} eV). This equation was used to fit the temperature-dependent intensities and the corresponding graph is shown in Fig. 13. The activation energy was found out to be 0.226 eV, which is higher than similar reported materials [49–51]. This proves that Dy^{3+} doped AZABS glass shows excellent thermal stability and is suitable for various lighting and w-LEDs applications.

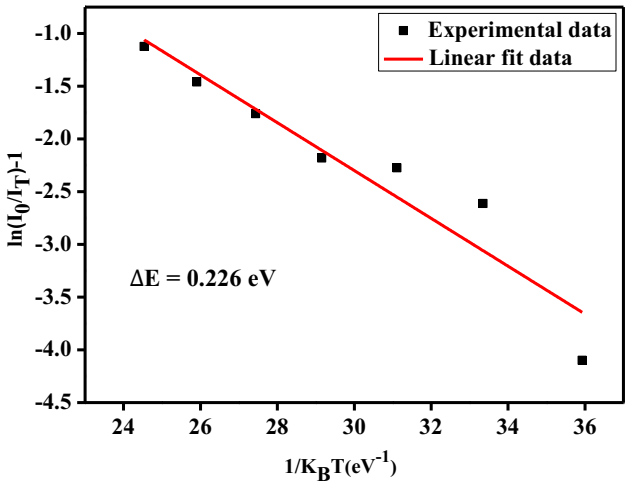


Fig. 13 $\ln[(I_0/I_T) - 1]$ versus $1/k_BT$ plot to calculate activation energy of 0.5 mol% Dy^{3+} doped AZABS glass

4 Conclusion

Dy^{3+} doped AZABS glasses were prepared using the quick melt quench technique. Various morphological and photoluminescence studies were conducted on the as-prepared glass samples. 0.5 mol % of Dy^{3+} ions doped AZABS glass was found to exhibit the highest luminescence intensity under UV excitation. Absorption spectral analysis of the same showed many bands denoting transitions from $^6H_{15/2} \rightarrow {}^7F_{11/2} + {}^6H_{9/2}$, $^6F_{9/2} + {}^6H_{7/2}$, $^6F_{7/2} + {}^6H_{5/2}$, $^6F_{5/2}$, $^6F_{3/2}$, $^4F_{9/2}$, $^4I_{15/2}$, $^4G_{11/2}$, $^4F_{7/2} + {}^4I_{13/2}$, $^6P_{5/2} + {}^4M_{19/2} + {}^4$ (P, D)$_{3/2}$, $^6P_{7/2}$ and $^4M_{17/2} + {}^6P_{3/2}$. Using this data, the optical bandgap of 0.5 mol % of Dy^{3+} ions doped AZABS glass was found to be 2.85 eV. Under 350 nm excitation, the glass samples show sharp bands related to transitions from the energy level $^4F_{9/2}$ to the levels $^6H_{15/2}$ (482 nm), $^6H_{13/2}$ (575 nm) and $^6H_{11/2}$ (665 nm). Dexter theory was applied to the PL data and its analysis showed that the energy transfer between the dopant ions was of dipole–dipole in nature. Y/B ratio calculations and colorimetry analysis proved that the as-prepared glasses are capable of producing cool white light. Temperature-dependent PL studies show that these AZABS glass samples have excellent thermal stabilities. Decay kinetics of the as-prepared glass samples indicate high lifetimes. On the basis of aforementioned characteristic results of our research work, we recommend Dy^{3+} doped AZABS glasses as suitable candidates for applications in w-LEDs and SSL applications.

Acknowledgements

The work reported in this paper has been partially supported by a Major Research Project (EMR/2016/007766) sanctioned by SERB, Department of Science and Technology, Govt. of India, New Delhi. Dr. Sumandeep Kaur is grateful to CSIR, New Delhi for the award of Research Associate fellowship (08/133(0053) 2020-EMR-I).

Author contributions

RB: Work plan, experimental work, manuscript writing. AP: Experimental Help, calculations. AVSY: Experimental Help, calculations. PR: Experimental Help, calculations. SK: Validation of results, Manuscript correction. ASR: Work plan, validation of results, manuscript correction.

Data availability

The data that support the findings of this study are available on request from the corresponding author, [Prof. A.S. Rao]. The data are not publicly available due to restrictions, e.g. their containing information that could compromise the privacy of research participants.

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. K. Swapna, S. Mahamuda, A.S. Rao, T. Sasikala, P. Packiyaraj, L.R. Moorthy, G.V. Prakash, Luminescence characterization of Eu^{3+} doped zinc alumino bismuth borate glasses for visible red emission applications. J. Lumin. **156**, 80–86 (2014). https://doi.org/10.1016/j.jlumin.2014.07.022
2. K. Annapoorani, K. Marimuthu, Spectroscopic properties of Eu^{3+} ions doped Barium telluroborate glasses for red laser applications. J. Non-Cryst. Solids **463**, 148–157 (2017). https://doi.org/10.1016/j.jnoncrysol.2017.03.004
3. S. Arunkumar, K. Marimuthu, Structural and luminescence studies on Eu^{3+}: B_2O_3-Li_2O-MO-LiF (M=Ba, Bi, Pb and Zn) glasses. AIP Conf. Proc. **1447**, 573–574 (2012). https://doi.org/10.1063/1.4710133
4. S. Mahamuda, K. Swapna, M. Venkateswarlu, A. Srinivasa Rao, S. Shakya, G. Vijaya Prakash, Spectral characterisation of Sm^{3+} ions doped Oxy-fluoroborate glasses for visible orange luminescent applications. J. Lumin. **154**, 410–424 (2014). https://doi.org/10.1016/j.jlumin.2014.05.017
5. S. Mahamuda, K. Swapna, A. Srinivasa Rao, T. Sasikala, L. Rama Moorthy, Reddish-orange emission from Pr^{3+} doped zinc alumino bismuth borate glasses. Phys. B Condens. Matter **428**, 36–42 (2013). https://doi.org/10.1016/j.physb.2013.07.010
6. A.K. Vishwakarma, M. Jayasimhadri, Pure orange color emitting Sm^{3+} doped $BaNb_2O_6$ phosphor for solid—state lighting applications. J. Lumin. **176**, 112–117 (2016). https://doi.org/10.1016/j.jlumin.2016.03.025
7. C.B. Annapurna Devi, S. Mahamuda, M. Venkateswarlu, K. Swapna, A. Srinivasa Rao, G. Vijaya Prakash, Dy^{3+} ions doped single and mixed alkali fluoro tungsten tellurite glasses for LASER and white LED applications. Opt. Mater. (Amst.) **62**, 569–577 (2016). https://doi.org/10.1016/j.optmat.2016.11.016
8. N. Deopa, A.S. Rao, S. Mahamuda, M. Gupta, M. Jayasimhadri, D. Haranath, G.V. Prakash, Spectroscopic studies of Pr^{3+} doped lithium lead alumino borate glasses for visible reddish orange luminescent device applications. J. Alloys Compd. **708**, 911–921 (2017). https://doi.org/10.1016/j.jallcom.2017.03.020
9. R. Bajaj, M. Gupta, R. Nagarajan, A.S. Rao, G.V. Prakash, Strong structural phase sensitive rare-earth photoluminescence color flips in $KLaF_4$:RE^{3+} (RE^{3+} = Eu^{3+}, Er^{3+}/Yb^{3+}) nanocrystals. Dalton Trans. (2020). https://doi.org/10.1039/d0dt01179g
10. Y. Qiao, P. Chen, Luminescence, energy transfer, and color adjustment of CaO-CaF_2-Al_2O_3 glass co-doped with CeO_2 and Sm_2O_3. J. Non-Cryst. Solids **552**, 120461 (2021). https://doi.org/10.1016/j.jnoncrysol.2020.120461
11. J. Dahiya, A. Hooda, A. Agarwal, S. Khasa, Tuneable colour flexibility in Dy^{3+} & Eu^{3+} co-doped lithium fluoride bismuth borate glass system for solid state lighting applications. J. Non-Cryst. Solids **576**, 121237 (2022). https://doi.org/10.1016/j.jnoncrysol.2021.121237
12. R.H.A. Bergh, G. Craford, A. Duggal, The promise and challenge of solid-state lighting. Phys. Today **54**, 42–47 (2001). https://doi.org/10.1063/1.1445547
13. G. Zissis, Energy consumption and environmental and economic impact of lighting: the current situation. Handb. Adv. Light. Technol. (2017). https://doi.org/10.1007/978-3-319-00176-0
14. N. Deopa, A.S. Rao, Photoluminescence and energy transfer studies of Dy^{3+} ions doped lithium lead alumino borate glasses for w-LED and laser applications. J. Lumin. **192**, 832–841 (2017). https://doi.org/10.1016/j.jlumin.2017.07.052
15. K. Jha, M. Jayasimhadri, Spectroscopic investigation on thermally stable Dy^{3+} doped zinc phosphate glass for white light emitting diodes. J. Alloys Compd. **688**, 833–840 (2016). https://doi.org/10.1016/j.jallcom.2016.07.024
16. K. Jha, A.K. Vishwakarma, M. Jayasimhadri, D. Haranath, Multicolor and white light emitting Tb^{3+}/Sm^{3+} co-doped zinc phosphate barium titanate glasses via energy transfer for

optoelectronic device applications. J. Alloys Compd. **719**, 116–124 (2017). https://doi.org/10.1016/j.jallcom.2017.05.076

17. H. Masai, Y. Yamada, Y. Suzuki, K. Teramura, Y. Kanemitsu, T. Yoko, Narrow energy gap between triplet and singlet excited states of SN^{2+} in borate glass. Sci. Rep. **3**, 1–7 (2013). https://doi.org/10.1038/srep03541
18. Y. Zhang, C. Lu, L. Sun, Z. Xu, Y. Ni, Influence of Sm_2O_3 on the crystallization and luminescence properties of boroaluminosilicate glasses. Mater. Res. Bull. **44**, 179–183 (2009). https://doi.org/10.1016/j.materresbull.2008.03.004
19. H.M. Gomaa, H.A. Saudi, I.S. Yahia, M.A. Ibrahim, H.Y. Zahran, Optik Influence of exchanging CeO_2 with Cu_2O_3 on structural matrix, shielding, and linear / nonlinear optical parameters of the cerium-sodium borate glass. Optik (Stuttg.) **249**, 168267 (2022). https://doi.org/10.1016/j.ijleo.2021.168267
20. N.S. Prabhu, V. Hegde, A. Wagh, M.I. Sayyed, O. Agar, S.D. Kamath, Physical, structural and optical properties of Sm^{3+} doped lithium zinc alumino borate glasses. J. Non-Cryst. Solids **515**, 116–124 (2019). https://doi.org/10.1016/j.jnoncrysol.2019.04.015
21. J. Rajagukguk, J. Kaewkhao, M. Djamal, R. Hidayat, Suprijadi, Y. Ruangtaweep, Structural and optical characteristics of Eu^{3+} ions in sodium-lead-zinc-lithium-borate glass system. J. Mol. Struct. **1121**, 180–187 (2016). https://doi.org/10.1016/j.molstruc.2016.05.048
22. W. Bao, X. Yu, T. Wang, H. Zhang, C. Su, Tb3+ / Eu3+ co-doped Al2O3 – B2O3 – SrO glass ceramics: preparation, structure and luminescence properties. Opt. Mater. (Amst.) **122**, 111772 (2021). https://doi.org/10.1016/j.optmat.2021.111772
23. N. Deopa, A.S. Rao, Spectroscopic studies of Sm^{3+} ions activated lithium lead alumino borate glasses for visible luminescent device applications. Opt. Mater. (Amst) **72**, 31–39 (2017). https://doi.org/10.1016/j.optmat.2017.04.067
24. H. Gui, C. Li, C. Lin, Q. Zhang, Z. Luo, L. Han, J. Liu, T. Liu, A. Lu, Glass forming, crystallization, and physical properties of MgO-Al_2O_3-SiO_2-B_2O_3 glass-ceramics modified by ZnO replacing MgO. J. Eur. Ceram. Soc. **39**, 1397–1410 (2019). https://doi.org/10.1016/j.jeurceramsoc.2018.10.002
25. N.F. Yusof, K.H.K. Bulat, N.A. Badarulzaman, M.A.A.M. Nor, The effect of boron oxide on concentration of iron ion released from SiO_2-B_2O_3-Na_2CO_3-Fe_2O_3 Glass System. Int. J. Curr. Res. Sci. Eng. Technol. **1**, 130 (2018). https://doi.org/10.30967/ijcrset.1.s1.2018.130-134
26. N. Deopa, S. Saini, S. Kaur, A. Prasad, A.S. Rao, Spectroscopic investigations on Dy^{3+} ions doped zinc lead alumino borate glasses for photonic device applications. J. Rare Earths **37**, 52–59 (2019). https://doi.org/10.1016/j.jre.2018.04.013
27. M.H.A. Mhareb, M. Alqahtani, F. Alshahri, Y.S.M. Alajerami, N. Saleh, N. Alonizan, M.I. Sayyed, M.G.B. Ashiq, T. Ghrib, S.I. Al-Dhafar, T. Alayed, M.A. Morsy, The impact of barium oxide on physical, structural, optical, and shielding features of sodium zinc borate glass. J. Non-Cryst. Solids. **541**, 120090 (2020). https://doi.org/10.1016/j.jnoncrysol.2020.120090
28. J. Rajagukguk, B. Sinaga, J. Kaewkhao, Structural and spectroscopic properties of Er^{3+} doped sodium lithium borate glasses, Spectrochim. Acta Part A Mol. Biomol. Spectrosc. **223**, 117342 (2019). https://doi.org/10.1016/j.saa.2019.117342
29. K. Swapna, S. Mahamuda, A. Srinivasa-Rao, M. Jayasimhadri, T. Sasikala, L. Rama-Moorthy, Visible fluorescence characteristics of Dy^{3+} doped zinc alumino bismuth borate glasses for optoelectronic devices. Ceram. Int. **39**, 8459–8465 (2013). https://doi.org/10.1016/j.ceramint.2013.04.028
30. N. Kiran, A. Suresh Kumar, White light emission from Dy^{3+} doped sodium-lead borophosphate glasses under UV light excitation. J. Mol. Struct. **1054–1055**, 6–11 (2013). https://doi.org/10.1016/j.molstruc.2013.09.023
31. A.M. Babu, B.C. Jamalaiah, J.S. Kumar, T. Sasikala, L.R. Moorthy, Spectroscopic and photoluminescence properties of Dy^{3+}-doped lead tungsten tellurite glasses for laser materials. J. Alloys Compd. **509**, 457–462 (2011). https://doi.org/10.1016/j.jallcom.2010.09.058
32. R. Bajaj, A.S. Rao, G.V. Prakash, Photoluminescence down-shifting studies of thermally stable Eu^{3+} ions doped borosilicate glasses for visible red photonic device applications. J. Non-Cryst. Solids. **575**, 121184 (2022). https://doi.org/10.1016/j.jnoncrysol.2021.121184
33. S. Hemalatha, M. Nagaraja, A. Madhu, K. Suresh, N. Srinatha, Spectroscopic studies of transition metal ion-doped borate glasses for optical applications. Mater. Today Proc. (2021). https://doi.org/10.1016/j.matpr.2021.08.075
34. W.T. Carnall, P.R. Fields, K. Rajnak, Electronic energy levels of the trivalent lanthanide aquo ions. IV. Eu^{8+}. J. Chem. Phys. **49**, 4424–4442 (1968). https://doi.org/10.1063/1.1669893
35. P. Rekha Rani, M. Venkateswarlu, S. Mahamuda, K. Swapna, N. Deopa, A.S. Rao, Spectroscopic studies of Dy^{3+} ions doped barium lead alumino fluoro borate glasses. J. Alloys Compd. **787**, 503–518 (2019). https://doi.org/10.1016/j.jallcom.2019.02.088
36. H. George, N. Deopa, S. Kaur, A. Prasad, M. Sreenivasulu, M. Jayasimhadri, A.S. Rao, Judd-Ofelt parametrization and radiative analysis of Dy^{3+} ions doped sodium bismuth strontium phosphate glasses. J. Lumin. (2019). https://doi.org/10.1016/j.jlumin.2019.116693

37. S. Kaur, M. Jayasimhadri, A.S. Rao, A novel red emitting Eu^{3+}doped calcium aluminozincate phosphor for applications in w-LEDs. J. Alloys Compd. **697**, 367–373 (2017). https://doi.org/10.1016/j.jallcom.2016.12.150
38. A.K. Vishwakarma, M. Jayasimhadri, Significant enhancement in photoluminescent properties via flux assisted Eu3+ doped $BaNb_2O_6$ phosphor for white LEDs. J. Alloys Compd. **683**, 379–386 (2016). https://doi.org/10.1016/j.jallcom.2016.05.052
39. M. Reddi Babu, N. Madhusudhana Rao, A. Mohan Babu, N. Jaidass, C. Krishna Moorthy, L. Rama Moorthy, Effect of Dy^{3+} ions concentration on optical properties of lead borosilicate glasses for white light emission. Optik (Stuttg) **127**, 3121–3126 (2016). https://doi.org/10.1016/j.ijleo.2015.12.018
40. C.R. Kesavulu, H.J. Kim, S.W. Lee, J. Kaewkhao, N. Chanthima, Y. Tariwong, Physical, vibrational, optical and luminescence investigations of Dy^{3+}-doped yttrium calcium silicoborate glasses for cool white LED applications. J. Alloys Compd. **726**, 1062–1071 (2017). https://doi.org/10.1016/j.jallcom.2017.08.091
41. L. Mishra, A. Sharma, A.K. Vishwakarma, K. Jha, M. Jayasimhadri, B.V. Ratnam, K. Jang, A.S. Rao, R.K. Sinha, White light emission and color tunability of dysprosium doped barium silicate glasses. J. Lumin. **169**, 121–127 (2016). https://doi.org/10.1016/j.jlumin.2015.08.063
42. D.L. Dexter, A theory of sensitized luminescence in solids. J. Chem. Phys. **21**, 836–850 (1953). https://doi.org/10.1063/1.1699044
43. D.L. Dexter, J.H. Schulman, Theory of concentration quenching in inorganic phosphors. J. Chem. Phys. **22**, 1063–1070 (1954). https://doi.org/10.1063/1.1740265
44. K. Linganna, C.S. Rao, C.K. Jayasankar, Optical properties and generation of white light in Dy^{3+}-doped lead phosphate glasses. J. Quant. Spectrosc. Radiat. Transf. **118**, 40–48 (2013). https://doi.org/10.1016/j.jqsrt.2012.12.002
45. M.K. Sahu, M. Jayasimhadri, Conversion of blue emitting thermally stable $Ca_3Bi(PO_4)_3$ host as a color tunable phosphor via energy transfer for luminescent devices. J. Lumin. **227**, 117570 (2020). https://doi.org/10.1016/j.jlumin.2020.117570
46. A. Prasad, A.S. Rao, M. Gupta, G.V. Prakash, Morphological and luminescence studies on $KGdF_4$:Yb^{3+}/Tb^{3+} up-conversion nanophosphors. Mater. Chem. Phys. **219**, 13–21 (2018). https://doi.org/10.1016/j.matchemphys.2018.07.056
47. A. Prasad, A.S. Rao, G.V. Prakash, A study on up-conversion and energy transfer kinetics of KGdF4: Yb3+ / Er3+ nanophosphors. J. Mol. Struct. **1205**, 127647 (2020). https://doi.org/10.1016/j.molstruc.2019.127647
48. C.S. McCamy, Correlated color temperature as an explicit function of chromaticity coordinates. Color Res. Appl. **17**, 142–144 (1992). https://doi.org/10.1002/col.5080170211
49. L. Wang, M. Xu, H. Zhao, D. Jia, Luminescence, energy transfer and tunable color of Ce^{3+}, $D^{y3}+/T^{b3}+$ doped $BaZ_n2(P_O4)2$ phosphors. New J. Chem. **40**, 3086–3093 (2016). https://doi.org/10.1039/c5nj03148f
50. D. Xu, Z. Yang, J. Sun, X. Gao, J. Du, Synthesis and luminescence properties of double-perovskite white emitting phosphor Ca_3WO_6:Dy^{3+}. J. Mater. Sci. Mater. Electron. **27**, 8370–8377 (2016). https://doi.org/10.1007/s10854-016-4848-z
51. M.K. Sahu, J. Mula, White light emitting thermally stable bismuth phosphate phosphor $Ca_3Bi(PO_4)_3$:Dy^{3+} for solid-state lighting applications. J. Am. Ceram. Soc. **102**, 6087–6099 (2019). https://doi.org/10.1111/jace.16479
52. R. Bajaj, A.S. Rao, G.V. Prakash, Linear and nonlinear photoluminescence from thermally stable KYF_4:Eu^{3+} cubic nanocrystals. J. Alloys Compd. **885**, 160893 (2021). https://doi.org/10.1016/j.jallcom.2021.160893

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Effect of Gate Oxide Material Variability on The Analog Performance of T-Gate GaN-MOS-HEMT with Graded Buffer

Megha Sharma
Department of Applied Physics
Delhi Technological University
Delhi, India
meghasharma_phd2k18@dtu.ac.in

Bhavya Kumar
Department of Applied Physics
Delhi Technological University
Delhi, India
bhavyakumar_phd2k18@dtu.ac.in

Rishu Chaujar
Department of Applied Physics
Delhi Technological University
Delhi, India
chaujar.rishu@dtu.ac.in

2021 7th International Conference on Signal Processing and Communication (ICSC) | 978-1-6654-2739-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICSC53193.2021.9673466

***Abstract* — This paper investigates the effect of gate oxide material variability on the analog performance of the T-gate GaN-MOS-HEMT with AlGaN gradient buffer. The different oxide materials such as HfO_2 (hafnium oxide), SiO_2 (silicon oxide), and composite HfO_2-SiO_2 are investigated. The DC and analog performance for different gate oxides have been studied by the Atlas-Silvaco simulator. The simulated study confirmed that using the HfO_2 as a gate oxide significantly reduced the leakage current, a positive shift of threshold voltage (-1.5V), the higher peak of transconductance (0.11S), transconductance generation factor (75V^{-1}), early voltage (V_{EA}) and intrinsic gain (A_v) is observed. These tremendous results confirm that the proposed structure is applicable for analog applications.**

Keywords — Analog, dielectric, oxide, T-gate, graded.

I. Introduction

In the last many years, silicon-based MOSFETs have dominated semiconductor technology. Nowadays, silicon-based devices approach their theoretical limit of operation as they suffer from undesirable effects. Therefore, there is an extensive need for new materials and technology that can overcome silicon technology limits. In recent years, much research has been done on GaN-based devices [1-8]. The GaN material has tremendous physical properties such as larger bandgap, high saturation velocity, higher electron mobility [1-3], making the GaN devices more prominent for the RF application. The RF performance of the GaN HEMTs can be improved by scaling the dimension of the device. As the dimension of GaN HEMT reduces, then some undesirable effects are induced, such as strain at the AlGaN/GaN layer, gate leakage current. The strain effect is eliminated by the InAlN layer with Al composition 0.83 [9]. Several advantages have been shown by the InAlN/GaN HEMT over AlGaN/GaN HEMTs. However, these devices suffer from two major problems, i.e., poor electron confinement in the channel and gate leakage current. These two issues drastically degrade the device's performance. The two engineering schemes have been imposed in the proposed device, which effectively overcomes these issues. The gate oxide layer has been inserted, which reduced the leakage current [10], and carrier confinement has been enhanced by the graded buffer [11]. The graded buffer generates negative polarization charges, which uplifts the conduction band discontinuity at GaN/buffer interface [12].

In this paper, we have investigated the analog performance of the proposed device with variable gate oxide materials. The DC and analog parameters for different gate oxide materials are theoretically compared. It has been observed that using the high k material gate oxide like HfO_2 effectively reduced the leakage current and enhanced the analog parameters of the device.

II. Device Structure

The proposed device structure is shown in Fig. 1. The value of the device parameters is given in Table I. The parasitic capacitance is reduced by the Al_2O_3 passivated layer, which further enhances the RF parameters of the device [13]. The T-shaped gate structure is formed with a gate length of 70 nm, which further improves the RF performance of the device. The 5 nm gate oxide is used, InAlN barrier layer is used with Al composition 0.83. The spacer layer has been inserted in between the channel and barrier layer. The insertion of the spacer layer enhances the conduction band discontinuity and further enhances the 2DEG concentration. [14], to increase the 2DEG confinement in the channel region, a 1µm AlGaN gradient buffer is inserted below the channel layers. The contact resistance of source/ drain regions is reduced by heavy doping [15].

III. Physical Models

The Atlas-Silvaco simulator has been used to simulate the proposed device [16]. Many physical models such as

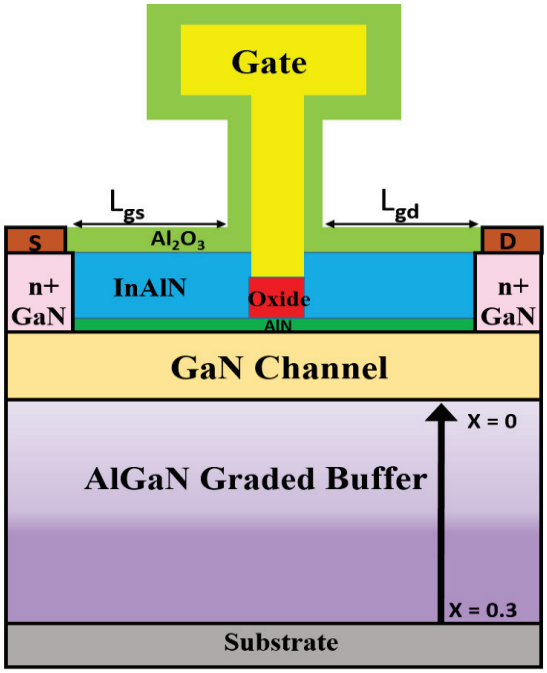


Fig. 1. Proposed device structure.

978-1-6654-2739-5/21/$31.00 ©2021 IEEE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:36:29 UTC from IEEE Xplore. Restrictions apply.

high field mobility, Fermi-Dirac, low field mobility, Shockley-Read-Hall, and polarization models are used in the simulation. The electron trapping mechanism has been calculated by the SRH model. The mobility models have investigated the scattering mechanism, i.e., low field and high field [16].

1. Low Field Mobility Model

$$\frac{1}{\mu(T,\ N)} = a\left(\frac{N}{N^{ref}}\right)\left(\frac{T}{300}\right)^{-15} \times \ln\left[1 + 3\left(\frac{T}{300}\right)^{2}\left(\frac{N}{N^{ref}}\right)^{-2/3}\right] + b\left(\frac{T}{300}\right)^{1.5} + \frac{c}{\left(e^{1065/T}\right) - 1} \quad (1)$$

where μ(T, N) is the mobility as a function of the doping and ambient temperature.

2. High Field Mobility Model

$$\mu(E) = \frac{\mu(N,T) + V_{sat}\dfrac{E(N_1 - 1)}{E_C{}^{N_1}}}{1 + a_n\left(\dfrac{E}{E_C}\right)^{N_2} + \left(\dfrac{E}{E_C}\right)^{N_1}} \quad (2)$$

where E and vsat are the electric fields, and saturated velocity of electrons and E_C, a_n, N_1, N_2 are coefficients.

3. Total Polarization Model

The two types of polarization induced in GaN-based devices, Spontaneous and piezoelectric polarization. The total polarization model is defined in Eq. (15) [17]:

$$P_{total} = [(P_{PE} + P_{SP})(bottom)] - [(P_{PE} + P_{SP})(top)] \quad (3)$$

For the lattice-matched InAlN/GaN, only spontaneous polarization is considered. The complete lattice relaxation is considered at the interface of the GaN channel and graded buffer. Hence total polarization is spontaneous polarization. The polarization gradient induces the space polarization charges in buffer, the density of these charges is given below [12].

$$\rho P = -\nabla . P = -\frac{\partial P}{\partial Z} \quad (4)$$

TABLE I. DESIGN PARAMETERS OF THE PROPOSED DEVICE

S. No.	Device Parameters	Proposed Structure
1	Gate length (L_g)	70 nm
2	Al_2O_3 passivation layer	10 nm
3	Source to gate length (L_{gs})	400 nm
4	Gate to drain length (L_{gd})	1.5 μm
5	Barrier layer (InAlN)	10 nm
6	AlN layer	1 nm
7	Gate Oxide	5 nm
8	GaN buffer	1.5 μm
9	Source/drain doping	1×10^{20} cm^{-3}
10	Gate width	2×50 μm
11	Gate head	400 nm
12	Gate stem height	100 nm

IV. RESULTS AND DISCUSSION

The substantial analysis of the impact of different gate oxide materials on analog parameters of the proposed device has been investigated in this section. Fig. (2, 3) shows the linear and log scale drain current density at fixed drain voltage V_{ds} = 2 V for different gate oxides. The drain current density of HfO_2 gate oxide is 0.23 A observed, which is about 4.64% smaller than the SiO_2 and 4.21% lower than the composite HfO_2-SiO_2 gate. The drain current density in Si-based MOSFETs is proportional to the dielectric constant of oxides [18], but in MOS-HEMT, the two-dimensional electron well is formed at the InAlN/AlN heterojunction, which facilitates the high speed of electrons. Hence in MOS-HEMT, drain current increases on reducing the dielectric constant of the gate oxide. Fig. 4 exhibits the threshold voltage (V_{th}) for variable gate oxides. The threshold voltage of HfO_2 (-1.5) gate oxide is shifted toward the positive side than the SiO_2 and stack HfO_2-SiO_2 gate oxides, showing that HfO_2 has better control over a channel. The HfO_2 has a higher dielectric constant and larger barrier height than the SiO_2, significantly reducing the gate leakage current.

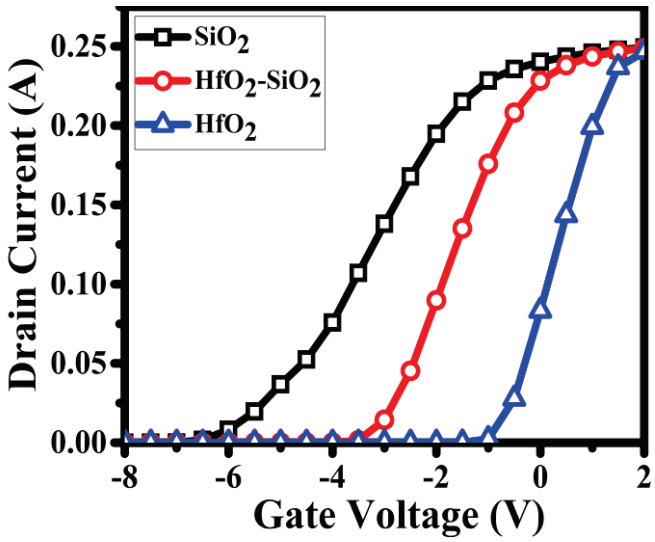


Fig. 2. Drain current characteristics for different gate oxide material.

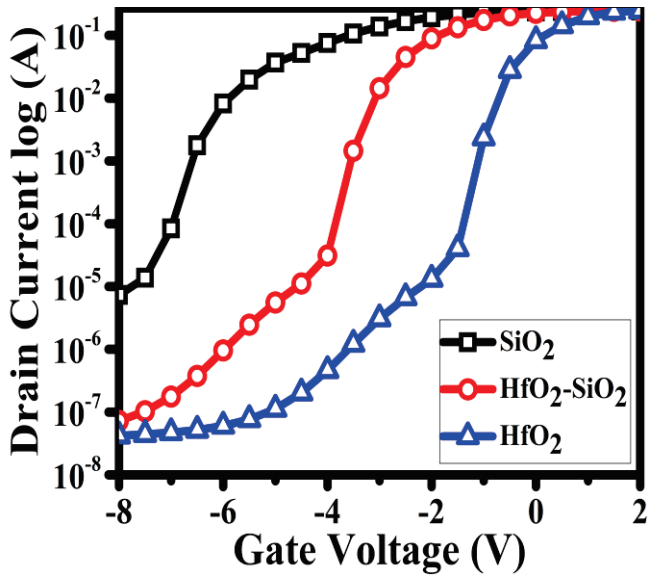


Fig. 3. Log scale drain characteristics for different gate oxide material.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:36:29 UTC from IEEE Xplore. Restrictions apply.

The mathematical formulation of analog parameters such as g_m, TGF, g_d, V_{EA}, and A_v are expressed by the equations (5-9) [19-20]:

$$g_m = \frac{\partial I_d}{\partial V_{gs}} \quad (5)$$

$$TGF = \frac{g_m}{I_d} \quad (6)$$

$$g_d = \frac{\partial I_d}{\partial V_{ds}} \quad (7)$$

$$A_v = \frac{g_m}{g_d} \quad (8)$$

$$V_{EA} = \frac{I_d}{g_d} \quad (9)$$

The device's transconductance (g_m) is expressed as the change in drain current over a change in gate voltage at constant drain voltage. Fig. 5 shows the transconductance characteristics of the device with different gate oxides. It has been observed that the gate controllability and transconductance change vigorously with the different gate oxides material. The higher peak of transconductance is observed for HfO_2 (0.11 S) gate oxide, which is 20 times higher than the SiO_2 (0.006 S) and 10 times higher than the composite HfO_2-SiO_2 (0.015 S) gate oxides. The HfO_2 oxide has a higher dielectric permittivity, which effectively improves the gate modulation and reduces the transconductance loss of the device.

Transconductance generation factor (TGF) is the ratio of g_m and I_d, and it is an important parameter in the analog application. TGF defines the gain available/power dissipate [21], where g_m shows the gain and I_d expresses the dissipated power. The impact of different gate oxides on TGF value is shown in Fig. 6. The higher peak of TGF is obtained for the HfO_2 gate oxide at V_{gs} = -1.5, which is 4 times higher than the SiO_2 gate oxide and 66.3% higher than the composite HfO_2-SiO_2. The enhancement of TGF value for HfO_2 is due to the higher g_m value for the other two gate oxides. The plot of output conductance for different gate oxides is shown in Fig. 7. Output conductance is the derivative of the drain current and drain voltage. The output conductance for the HfO_2 gate oxide is lower than the SiO_2 and stack HfO_2-SiO_2 gate oxides. Early voltage is expressed as the ratio of drain current and g_d [22]. In Fig. 8, the effect of different gate oxides on early voltage is shown. The high peak of early voltage is observed for the HfO_2 gate oxide at V_{gs} = 2 V, which is 2 times higher than the other two oxides. Intrinsic gain is expressed as the ratio of g_m and g_d [23]. The impact of gate oxide material on the intrinsic gain of the proposed device is depicted in Fig. 9. The higher value of intrinsic gain is observed for the HfO_2 oxide at V_{gs} = 0.5, which is 4 times higher than the SiO_2 and 3 times higher than the composite HfO_2-SiO_2. The enhancement of intrinsic value for HfO_2 gate oxide is basically due to the higher transconductance value and lower output conductance of HfO_2 gate oxide. The performance parameters for all the three gate oxide layers are mentioned in Table II.

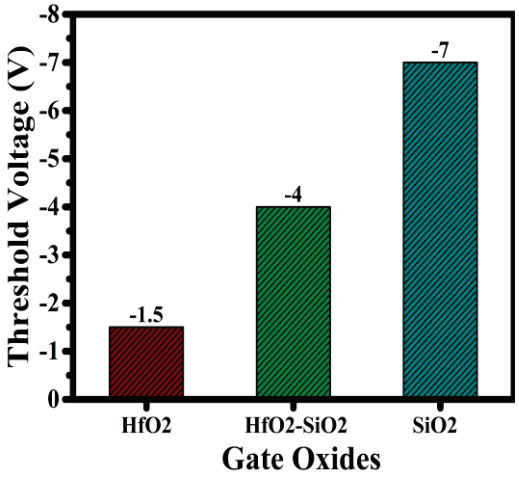


Fig. 4. Threshold voltage bar graph of variable gate oxide material.

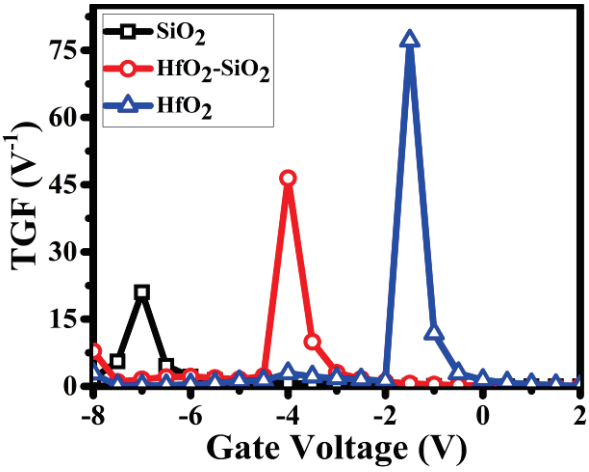


Fig. 6. Transconductance generation factor with variable gate oxides.

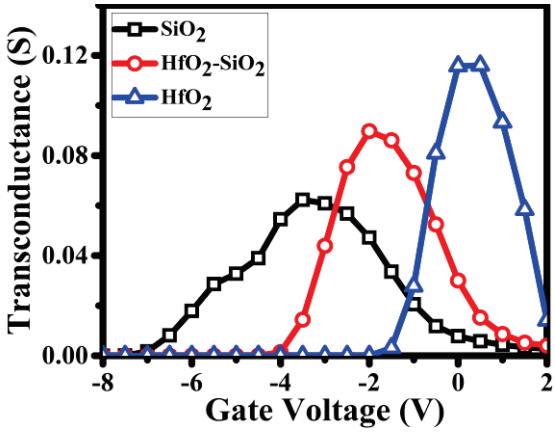


Fig. 5. Transconductance characteristics for variable gate oxide materials.

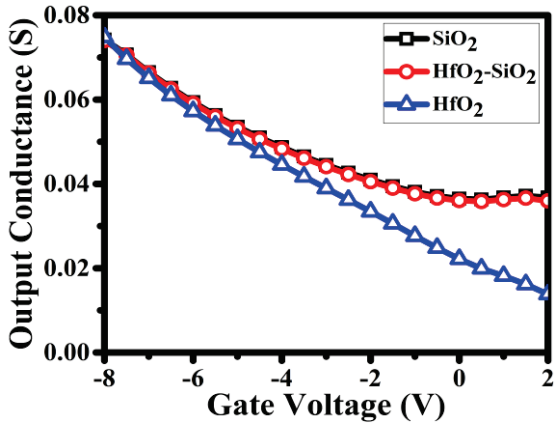


Fig. 7. Output conductance with gate bias for different gate oxides material.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:36:29 UTC from IEEE Xplore. Restrictions apply.

TABLE II
PERFORMANCE PARAMETERS FOR THE DIFFERENT GATE OXIDES

Gate Oxides	V_{th} (V)	g_m (S)	TGF (V^{-1})	g_d (S)	A_v (dB)	V_{EA} (V)
HfO_2	-1.5	0.11	77	1.3×10^{-2}	5.8	17.8
HfO_2-SiO_2	-4	0.015	46	3.6×10^{-2}	2.2	6.92
SiO_2	-7	0.006	21	3.68×10^{-2}	1.3	6.7

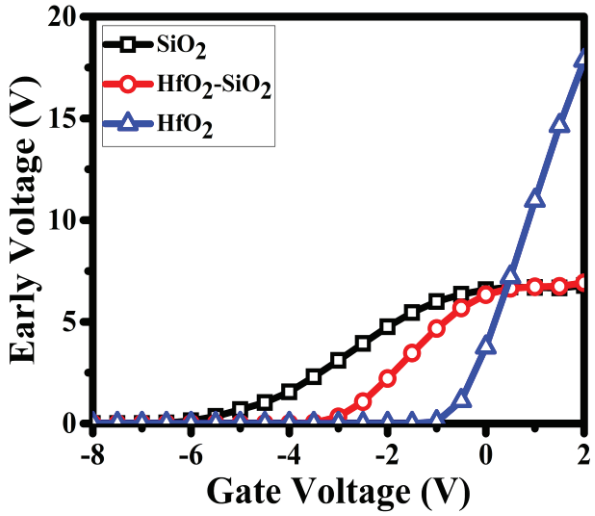


Fig. 8. Early voltage for different gate oxides.

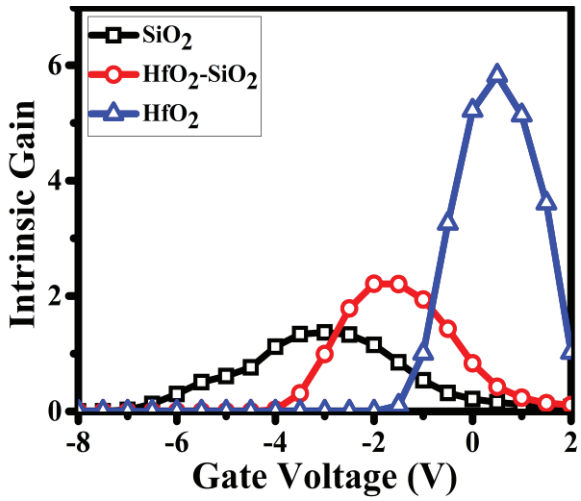


Fig. 9. Intrinsic gain for variable gate oxide material.

V. CONCLUSION

The effect of different gate oxides such as SiO_2, HfO_2, and composite HfO_2-SiO_2 on InAlN/GaN MOS-HEMT with graded buffer has been studied in this paper. The insertion of HfO_2 oxide below the gate significantly reduced the leakage current and enhanced the various analog parameters. The value of transconductance increases by 20 times, four times increment of transconductance generation factor, intrinsic gain value increase by 346%, and 165% increment in early voltage compared with SiO_2.

ACKNOWLEDGMENT

The authors are grateful to Micro-electronics Lab, Applied Physics Department, DTU for their support.

REFERENCES

[1] P. Cui and Y. Zeng, "Electrical properties of 90-nm InAlN/GaN HEMT on silicon substrate," Phys. E Low-Dimensional Syst. Nanostructures, vol. 134, pp. 114-821, 2021, DOI: 10.1016/j.physe.2021.114821.

[2] P. Cui and Y. Zeng, "Technology of sub-100 nm InAlN/GaN HEMTs on silicon with suppressed leakage current," Solid. State. Electron., vol. 185, pp. 108-137, 2021, DOI: 10.1016/j.sse.2021.108137.

[3] N. Chand, S. K. Swain, S. M. Biswal, A. Sarkar, and S. Adak, "Comparative study on Analog & RF Parameter of InAlN/AlN/GaN Normally off HEMTs with and without AlGaN Back Barrier," Devices for Integrated Circuit (DevIC), 2021 pp. 616–620, 2021, DOI: 10.1109/devic50843.2021.9455877.

[4] M. Sharma and R. Chaujar, "Impact of Graded Back-Barrier on Linearity of Recessed Gate InAlN/GaN HEMT," Proc. 2nd Int. Conf. VLSI Device, Circuit Syst. VLSI DCS 2020, pp. 18–19, 2020, DOI: 10.1109/VLSIDCS47293.2020.9179897.

[5] Ramkumar N, E. Parthasarathy, "In AlN / GaN High Electron Mobility Transistor with In GaN Back barrier for Future High Frequency Applications," Elementary Education Online vol. 20, no. 1, pp. 2186–2193, 2021, DOI: 10.17051/ilkonline.2021.01.241.

[6] M. Sharma and R. Chaujar, "The Performance Analysis of 70nm T-gate InAlN/AlN MOS-HEMT using Graded Buffer," Devices for Integrated Circuit (DevIC), 2021 vol. 1, pp. 466–470, 2021, DOI: 10.1109/devic50843.2021.9455887.

[7] Shaveta, H. M. M. Ahmed, and R. Chaujar, "Rapid detection of biomolecules in a dielectric modulated GaN MOSHEMT," J. Mater. Sci. Mater. Electron., 2020, DOI: 10.1007/s10854-020-04216-7.

[8] S. Sano, K. Ebihara, T. Yamamoto, T. Sato, and N. Miyazawa, "GaN HEMTs for wireless communication," SEI Tech. Rev., no. 86, pp. 65–70, 2018.

[9] M. Neuburger, T. Zimmermann, and E. Kohn, "Unstrained InAlN/ GaN HEMT Structure," Int. J. High Speed Electron. Syst vol. 14, no. 3, pp. 785–790, 2004 DOI:10.1142/S0129156404002831.

[10] J. J. Freedsman, T. Egawa, Y. Yamaoka, Y. Yano, A. Ubukata, T. Tabuchi, and K. Matsumoto. "Normally-OFF Al_2O_3/AlGaN/GaN MOS-HEMT on 8 in. Si with low leakage current and high breakdown voltage (825 V)," Appl. Phys. Express, vol. 7, no. 4, 2014, DOI: 10.7567/APEX.7.041003.

[11] D. Jena, S. Heikman, D. Green, D. Buttari, R. Coffie, H. Xing, S. Keller, S. DenBaars, J. S. Speck, and U. K. Mishra., "Realization of wide electron slabs by polarization bulk doping in graded III-V nitride semiconductor alloys," Appl Phys. Lett., vol. 81, no. 23, pp. 4395–4397, 2002, DOI: 10.1063/1.1526161.

[12] T. Inoue, T. Nakayama, Y. Ando, M. Kosaki, H. Miwa, K. Hirata, T. Uemura, and H. Miyamoto, "Polarization engineering on buffer layer in GaN-based heterojunction FETs," IEEE Trans. Electron Devices, vol. 55, no. 2, pp. 483–488, 2008, DOI: 10.1109/TED.2007.912367.

[13] D. S. Lee, X. Gao, S. Guo, D. Kopp, P. Fay, and T. Palacios, "300-GHz InAlN/GaN HEMTs With InGaN Back Barrier," IEEE Electron Device Lett., vol. 32, no. 11, pp. 1525–1527, 2011 DOI:10.1109/LED.2011.2164613.

[14] Y. L. and E. Y. C. Niraj Man Shrestha1, "Simulation study on electrical characteristic of AlGaN/GaN high electron mobility transistors with AlN spacer layer," Jpn. J. Appl. Phys., vol. 53, 2014, DOI:10.7567/JJAP.53.04EF08.

[15] P. Murugapandiyan, S. Ravimaran, and J. William, "DC and microwave characteristics of Lg 50 nm T-gate InAlN/AlN/GaN

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:36:29 UTC from IEEE Xplore. Restrictions apply.

HEMT for future high power RF applications," AEU - Int. J. Electron. Commun., vol. 77, pp. 163–168, 2017, DOI: 10.1016/j.aeue.2017.05.004.

[16] ATLAS User's Manual, SILVACO International, CA, USA, 2016.

[17] S. Li, M. Ware, J. Wu, P. Minor, Z. Wang, Z. Wu, Y. Jiang, and G. J. Salamo. "Polarization induced pn-junction without dopant in graded AlGaN coherently strained on GaN," Appl. Phys. Lett., vol. 101, no. 12, 2012, DOI: 10.1063/1.4753993.

[18] N. Gupta and R. Chaujar, "Investigation of temperature variations on analog/RF and linearity performance of stacked gate GEWE-SiNW MOSFET for improved device reliability," Microelectron. Reliab., vol. 64, pp. 235–241, 2016, DOI: 10.1016/j.microrel.2016.07.095.

[19] B. Kumar and R. Chaujar, "Analog and RF Performance Evaluation of Junctionless Accumulation Mode (JAM) Gate Stack Gate All Around (GS-GAA) FinFET," Silicon, vol. 13, pp. 919–927, 2021, DOI: 10.1007/s12633-020-00910-7.

[20] A. Kumar, M. M. Tripathi, and R. Chaujar, "Comprehensive analysis of sub-20 nm black phosphorus based junctionless-recessed channel MOSFET for analog/RF applications," Superlattices Microstruct., vol. 116, pp. 171–180, 2018, DOI: 10.1016/j.spmi.2018.02.018.

[21] H. Pardeshi, "Analog / RF performance of AlInN / GaN underlap DG MOS- HEMT," Superlattices Microstruct., vol. 88, pp. 508–517, 2015, DOI:10.1016/j.spmi.2015.10.009.

[22] R. Chaujar, R. Kaur, M. Saxena, M. Gupta, and R. S. Gupta, "Laterally amalgamated DUal Material GAte Concave (L-DUMGAC) MOSFET for ULSI," Microelectron. Eng., vol. 85, no. 3, pp. 566–576, 2008, DOI: 10.1016/j.mee.2007.10.002.

[23] J. Madan, R. S. Gupta, and R. Chaujar, "Mathematical modeling insight of hetero gate dielectric-dual material gate-GAA-tunnel FET for VLSI/analog applications," Microsyst. Technol., vol. 23, no. 9, pp. 4091–4098, 2017, DOI: 10.1007/s00542-016-2872-9.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:36:29 UTC from IEEE Xplore. Restrictions apply.

Plasma Chemistry and Plasma Processing
https://doi.org/10.1007/s11090-022-10229-3

ORIGINAL PAPER

Effect of Plasma Control Parameters on the Growth of Nitrogen-Doped Nanocone-Vertical Graphene Hybrid: Theoretical Investigations

Shruti Sharma[1] · **Suresh C. Sharma**[1]

Received: 9 May 2021 / Accepted: 20 December 2021
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

A theoretical model on the growth of nitrogen-doped nanocone-vertical graphene hybrid (N-NCN-VG) is developed. The model contains the plasma sheath equations, which are solved via boundary conditions. Using the plasma sheath equations variation of ion number density, ion energy correlated with potential distribution with distance into the plasma sheath is explored. The model is combined with the growth of N-NCN-VG, in C_2H_2/NH_3 plasma to examine the effect of plasma control parameters i.e., total gas pressure (250 mTorr–5 Torr) and input plasma power (300–700 W) on the kinetics of plasma species (neutrals, positive ions, and electrons) and the growth attributes like height, thickness and number density of carbon of the N-NCN-VG hybrid. In addition, the model interprets the growth of VG sheet due to defect formation on nanocone. From our theoretical investigations, it is found that as the total gas pressure increases and input plasma power decreases the height of the nanocone, graphene sheet increases along with an increase in carbon number density on the nanocone surface. In addition, the field enhancement factor of the N-NCNVG hybrid decreases with an increase in input power. The obtained results are in good agreement with the existing experimental observations.

Keywords Nitrogen doped nanocone-vertical graphene hybrid · PECVD · Plasma control parameters · Total gas pressure · Input plasma power

Introduction

Intensive efforts have been made over the last few decades for designing electron field emitters owing to its promising application in X-ray generators, flat panel display, field emission (FE) guns, etc. [1–6] using nanocarbon allotrope because of its low voltage requirement for the extraction of electrons from its surface. The unique properties of carbon nanotubes (CNTs) and graphene have allowed these nanocarbon allotropes suitable for use as cold cathode emitters.

✉ Suresh C. Sharma
suresh321sharma@gmail.com

[1] Department of Applied Physics, Delhi Technological University, Shahbad Daulatpur, Bawana Road, Delhi 110042, India

The electron FE occurs mainly from the tip of CNT while its side remains less sensitive to FE and is less stable towards FE whereas for FE from graphene, the sharp edges of the graphene should be vertical to the substrate. Furthermore, except for the FE stability, the graphene FE properties are low in comparison to CNT. Considering both CNT and graphene, efforts have been made to design an improved and stable field emitter [7–9] which is a hybrid of CNT and graphene (Carbon Micronymphaea), as a 3-D field emitter. This hybrid structure consists of graphene branches from the tip and walls of the CNTs.

In material engineering, Plasma Enhanced Chemical Vapor Deposition (PECVD) is a well-established technique for growing nanostructures in which, the source gas is ionized via electrical energy to produce plasma comprising energetic electrons, positive, negative charged ions, neutrals, and radicals. The strength of the electric field formed within the plasma sheath (area of uncompensated charges) effectively controls the flux of energy and ionized species on the catalyst surface. Plasma control parameters influence sheath dynamics which in turn affect the electric field strength. Owing to the flexibility of PECVD technique, it is adopted for polymer treatment [10], in medicine for anti-cancer therapy [11]. Graphene, CNT, CNT-Graphene Hybrid grown using PECVD technique have found a wide range of applications such as in the fabrication of electron field emitters [12, 13], transparent conductors [14], electrodes for the lithium-ion battery [15, 16], supercapacitors [17, 18], coating of target material for ion acceleration [19], biomolecule sensors[20]. In space electric propulsion systems, multiwall CNTs are promising techniques to enhance channel wear resistance [21], as field emitter operating without or reduced gas supply installed on a hall effect thruster [4, 6].

Nitrogen doping of these structures lowers the work function, enhances the number of defects, active sites for oxygen reduction reactions, electrical conductivity, specific capacitance, and lowers resistance to charge transfer thus making them efficient electron field emitters, oxygen electrocatalyst for rechargeable Zn-Air batteries, electrochemical sensors, supercapacitors, visible-light-driven photocatalyst for organic dye degradation [22–26].

Studies shows (i) Regulating the growth time, concentration of C, microwave power, growth temperature, length of the composite structure can be effectively controlled, which in turn influence its field emission characteristics [27, 28]. (ii) Growth of graphene sheet occurs from the defects (split C–C bonds) on CNT via high-energy bombardment of ions in plasma and lengthening by epitaxial mechanism [8]. (iii) Growth of vertical graphene (VG) sheet over CNT via defect formation by varying PECVD process parameters which influence growth of VG sheet through mathematical modeling [29].

Motivated by the above mentioned works, author have tried to model and numerically analyzed the growth of graphene over nanocone surface and linked chemical sputtering/etching mechanism with the growth of N-NCN-VG Hybrid by varying the PECVD process control parameters i.e., total gas pressure, input plasma power, and growth time. Varying ion number density, etchant number density, carbon surface concentration are the key factors affecting growth of nanocone, defects, and VG Sheet over nanocone surface. In this model, the VG sheet nucleates over the defect generated on the nanocone surface.

Model

In this section, the growth kinetics of the N-NCN-VG hybrid is modeled via the PECVD route accompanying by key presumptions with equations to detail its growth and nucleation. In the present work chemical sputtering/etching assisted growth of N-NCN-VG

Hybrid is modeled wherein, plasma made of electrons, ion plus neutral of C_2H_2, ion plus neutral of H_2, ion plus neutral of NH_3 has been considered and are elaborated in Table 1 which are key reactants in plasma chemistry of the C_2H_2 /NH_3 gas combination [30]. Figure 1 shows the schematic of N-NCN-VG Hybrid growth on the Nickel (Ni) catalyst over Silicon (Si) substrate. Several stages are involved in the growth: creation of ultra-small metal catalyst nanoparticle as a beginning step; followed by its saturation with carbon species; and ultimately, the nanocone growth covering saturated nanoparticle. The incoming carbon species are responsible for the growth of carbon nanocone, while hydrogen ions and nitrogenous ions which build up the electric field near the nano-hillock surface take part in the role of etching of carbon nanocone [31]. So, there exist both deposition-etching during the growth of carbon nanocone. Initially, saturation of catalyst nanoparticles takes place by the adsorption, diffusion, and sticking of hydrocarbon species. Once, the saturation of catalyst nanoparticle is reached, nucleation of nano-hillock on its top takes place by additional carbon atoms. Here we have assumed that, along with the nucleation of nano-hillock at the center of catalyst nanoparticle, it also grows radially until the catalyst nanoparticle is fully covered [32]. So a nano-hillock of increased height and base radius with a small slope angle is obtained, which leads to weaker chemical sputtering/etching yield, consequently, carbon species deposition takes over during the preliminary growth stage of carbon nanocone [33]. With increasing deposition of carbon species, slope angle of nano-hillock increases, and chemical sputtering/etching becomes more effective and the rate of deposition decreases. When deposition of carbon reaches equilibrium with the etching of carbon so stable nanocone will be formed [33]. The presence of hydrocarbon, hydrogen, and nitrogenous ion in plasma lead to formation of defects on nanocone surface. Under applied bias and electrostatic field, the energetic ions (energy above the threshold energy needed for bond breaking) continuously bombard the nanocone plane to split several C–C bonds (bond energy 17–20 eV) [34], thereby creating cracks and dangling bonds (local defects) that are well-known nucleation sites for two dimensional VG sheet growth on nanocone surface[27, 35]. Then VG sheet proceeds the way as discussed in previous studies [36–38]. Later these defects capture hydrocarbon species and form carbon monomers after several processes to attain stability. Carbon monomers then diffuse and aggregate to make graphitic carbon clusters [37, 38]. As time elapses, the as formed carbon clusters diffuse and aggregate to form graphene nuclei promoting the formation of graphene islands which in turn diffuse and align to produce larger islands or flat graphene sheet [37–39]. Alignment of graphene layers takes place normally to nanocone surface under the effect of internal tensile force and inbuilt electric field (induced between bulk plasma and the negatively biased substrate surface as a result of charge separation region known as plasma sheath) along with regular dissociation and diffusion of hydrocarbon and carbon species, respectively, yielding graphene sheet growth [38]. For the growth of the graphene sheet, a nanocone template of base radius r and height h has been considered in the plasma constituting the ionized gas mixture of C_2H_2/NH_3. It is known that a typical PECVD, is described by the formation of plasma sheath. Within the plasma sheath, the medium shall be regarded as

Table 1 Neutrals and ions considered in model

Neutrals			Ions		
Type a	Type b	Type c	Type a	Type b	Type c
C_2H_2, CH_4, CH_3	H, H_2	NH_3,N_2, NH_4, HCN	$C_2H_2^+$, CH_4^+, CH_3^+	H^+, H_2^+	NH_3^+,N_2^+, NH_4^+, HCN^+

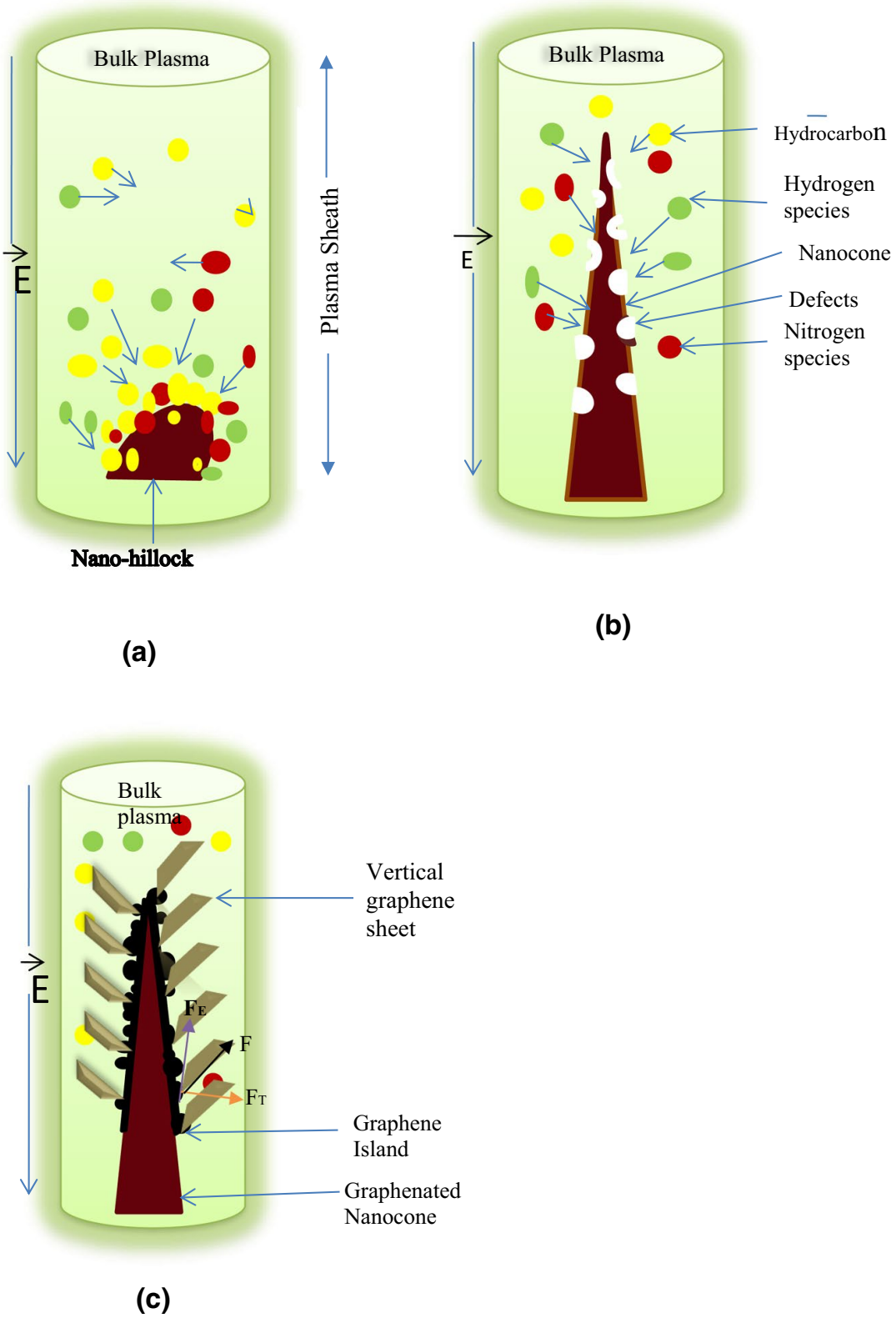


Fig. 1 Schematic showing the **a** Incident ionic species sputter the nano-hillock followed by the deposition of hydrocarbon species. **b** Simultaneously the height of the spike increases to result in an increase of slope angle and enhancement of chemical sputtering followed by formation of defects on the nanocone surface. **c** Carbon atoms are captured by the defects, which diffuse and agglomerate to form graphene island. The resultant (F) of these forces (internal tensile force (FT) and plasma- induced electric field force (FE)) tend to align the graphitic layer in the direction normal to the nanocone surface

two interpenetrating fluids constituting neutral molecules and ions, respectively [40] where the microscopic plasma sheath electrostatic force (from space charge separation) accelerates the ions heading to the surface [41]. Considering the induced sheath electrostatic field to be along the z-axis i.e., in the vertically downward direction. Presently in the plasma sheath model the ion velocities, ion energies, the number density of ions, and neutral species in the plasma is found by coupling the equation of continuity and momentum balance equation to the Poisson's equation which relates potential distribution in the plasma sheath and number densities of charged species [42, 43]

$$\frac{\partial\left(n_{ik}\vartheta_{ik}\right)}{\partial z} = \upsilon n_{ik,} \tag{1}$$

$$\frac{\partial\left(n_{e}\vartheta_{e}\right)}{\partial z} = \upsilon n_{e,} \tag{2}$$

$$n_{ik}\vartheta_{ik}\frac{\partial\vartheta_{ik}}{\partial z} = \frac{-en_{ik}}{M_{ik}}\frac{d\zeta}{dz} - n_{ik}\eta_{ik}\vartheta_{ik} - \frac{T_{ik}}{M_{ik}}\frac{dn_{ik}}{dz}, \tag{3}$$

$$en_{e}\frac{d\zeta}{dz} = T_{e}\frac{dn_{e}}{dz} + M_{e}n_{e}\eta_{e}\vartheta_{e,} \tag{4}$$

$$\frac{d^{2}\zeta}{dz^{2}} = -4\pi e\left(n_{e} - \sum_{k}\gamma_{ik}n_{ik}\right), \tag{5}$$

where M_{ik}; n_{ik}; T_{ik}; and ϑ_{ik} are the masses, number densities, temperature, and fluid velocities of the positively charged species, respectively [k corresponds to the a,b,c type of positively charged ions (see Table 1). M_e; n_e; T_e; and ϑ_e are the mass, number densities, temperature, and velocities of electrons, respectively.

Equations (1) and (2) are the continuity equations and Eqs. (3) and (4) are momentum balance equations for ions and electrons, respectively, and take into account ionization of neutrals by electrons with ionization frequency $\upsilon = \mu_{ik}\gamma_{ik}N$, where μ_{ik} is the ionization potential,$\gamma_{ik} = \frac{n_{ik}}{n_e}$ or $\frac{p_k}{p}$,where $p = NT_i$ and p_k are the total gas pressure and pressure of kth gas and $N = \sum_{k}\gamma_k n_k$ is the total neutral species number density, n_k is the number density of kth neutral. $\eta_\beta = \sigma_\beta N U_\beta$ is the collision frequency accounting the collision of ions and electrons with the neutrals, β refer to electrons and k ions, σ_β is the cross-section of the collision, U_β is the mean speed of the charged species β, ionization rate $\left(\upsilon\alpha n_k \exp\left(\frac{-E_{ik}}{T_e}\right)\right)$ with E_{ik} as first ionization potential energy of the kth neutral. The plasma sheath Eqs. (1)–(4) along with Poisson's Eq. (5) are solved by applying boundary conditions as follows.

$\zeta = 0, \frac{d\zeta}{dz} = -\frac{T_e}{e\lambda}, \vartheta_e = 0, \vartheta_{ik} = \vartheta_{iko}, \vartheta_{iko} = \sqrt{\frac{T_e}{M_{ik}}}$ *where* $\lambda = \frac{T_n}{\sigma_\beta p}$ is the ion mean free path.

Presumptions considered for simplified modelling [44–46]

(a) Maxwellian distribution of electrons and ions is considered
(b) Ion – surface interaction is assumed coulomb interaction

(c) For the sheath to perform its function and repel electrons, the potential must be monotonically decreasing with increase in z. This will occur if ni(z) > ne(z) for all z in the sheath.
(d) Plasma-surface interaction heats the nanocone surface, thus gaining the constant temperature Ts and is assumed uniform along nanocone
(e) The collisions of the neutrals, ions with nanocone surface are assumed elastic.

Chemical Sputtering/Etching

Low Ion Energy Sputtering

The incoming hydrogen, nitrogenous, hydrocarbon ions project and transfer their momentum to the atoms on the growing N-NCN –VG hybrid. It takes place above the threshold energy of incident ions. As a result, physically sputtered carbon atoms originate predominantly from the growing hybrid, thereby removing amorphous carbon. The threshold is specified as minimum incident energy needed to split a C–C bond which for physical sputtering of amorphous carbon by NH_3^+ and H^+ ions has been calculated to about 31.08 eV, and 22.44 eV for carbon surface binding energy of $E_B(\approx 4.5$ eV) [47].

For the sputtering of low energy ions, the threshold energy is given as [47, 48].

$$E_{thr} = \left\{ \begin{array}{l} \dfrac{E_B}{\Upsilon(1-\Upsilon)}, \left(M_{ik} \le M_c\right) \\ \dfrac{6.7E_B}{\Upsilon}, \left(M_{ik} > M_c\right) \\ \Upsilon = \dfrac{4M_{ik}M_c}{\left(M_{ik}+M_c\right)^2} \end{array} \right\}$$

where M_{ik} and M_c are the incident ion mass which can be hydrogen, nitrogenous, hydrocarbon ion and mass of target, respectively.

Ion Assisted Etching

Incident hydrogen, nitrogenous ions generate split bonds in the cascaded collision which are immediately stabilized by the rich atomic hydrogen and nitrogen flux. Repeated bond (sp^3-C) splitting events along with stabilization by H and N generate stable C–H and C–N bonds on the surface, which then get thermally desorbs from the surface, thereby removing amorphous carbon via the formation of volatile hydrocarbons, HCN, CN, C_2N_2 [49, 50]. Below are the reactions (A) & (B) showing ion-assisted etching, where the ion (I^+) can be hydrogen, nitrogenous.

$$C_{(ads)}H_{(plasma)} \xrightarrow{K_I I^+} CH_{(plasma)} \quad \text{(A)}$$

$$C_{(ads)}N_{(plasma)} \xrightarrow{K_I I^+} CN_{(plasma)} \quad \text{(B)}$$

where K_I is the rate coefficient for ions incident over surface.

Carbon species flux(Γ_C) leaving the surface is expressed by equation (C)

$$\Gamma_C = Y_I K_I n_{ik} \upsilon_o \theta_C + \frac{J_a \sigma_{ads} J_b}{\omega} + S_C \sigma_{ads} J_b + K_d \upsilon_o \theta_{CH}, \tag{C}$$

where the first, second, third term on the right side of equation (C) denotes the flux of carbon atom and carbon-bearing species leaving the N-NCN-VG Hybrid surface due to ion assisted sputtering and fourth terms flux of carbon-bearing species leaving the N-NCN-VG Hybrid surface due to thermal desorption of CH species. $\boldsymbol{Y_I}$ is the chemical sputtering/etching yield based on ion energy as $\mathbf{Y_I} \propto \sqrt{\boldsymbol{E_i}} - \sqrt{\boldsymbol{E_{thr}}}$ [51], where $\boldsymbol{E_i}$ *i*s the ion energy in eV and $\boldsymbol{E_{thr}}$ is the threshold energy in eV. $K_I = \sqrt{\frac{eT_e}{M_i}} \Big/ \upsilon_o$ denotes the rate coefficient for ions incident on the surface. $\boldsymbol{\theta_C}, \boldsymbol{\theta_{CH}}$ is the fraction of surface sites covered with C atom, CH species. $\mathbf{K_d} = \boldsymbol{K_0 exp}\left(\frac{-\boldsymbol{E_{dh}}}{\boldsymbol{K_B T_S}}\right)$ is the desorption rate constant. $\mathbf{K}_0$ is the number of attempted escapes per second from the adsorption well, hence $\mathbf{K}_0 = \frac{\omega}{2\boldsymbol{\pi}}$. $E_{dh} (\approx 1.8$ eV) is the hydrocarbon desorption energy. $Y_I = \frac{1}{\pi^2}\left(\frac{M_c M_{ik}}{(M_c + M_{ik})^2}\right)\frac{E_i}{E_B}(\cos\theta)^{-f}$, where f is constant for certain M_c/M_i in this ion energy range [52]. θ is the angle of incidence of ions on growing N-NCN-VG hybrid, k_B is the Boltzmann's constant.

Growth of Nitrogen-Doped Nanocone

The growth of nitrogen-doped nanocone, shown in terms of volumetric change (base radius r, tip radius r_1 and height h) with time, with the number density of carbon and hydrogen developed over the catalyst nanoparticle. Due to plasma-induced stress and/or localized electrostatic field, the nanocone growth takes place in the vertical direction. Nanocone growth process traced in Eq. (6) includes surface diffusion and incorporation of carbon monomer, along with surface diffusion of carbon cluster with the various energy barriers. The fourth term is the etching term which decreases the growth rate of nanocone. The final two terms add up to the sticking of neutrals of type 1 and 3, respectively over nanocone surface [53, 54].

$$\begin{aligned}\frac{1}{3}\pi \frac{d\left(\frac{r^3 - r_1^3}{r - r_1}\right)h}{dt} = &\left[\begin{pmatrix}\frac{D_1}{\omega}\exp\left(\frac{-E_{SDc}}{k_B T_s}\right) + \frac{D_2}{\omega}\exp\left(\frac{-E_{SDcl}}{k_B T_s}\right)\\ +A_k \exp\left(\frac{-\delta E_{INc}}{k_B T_s}\right)\omega r\end{pmatrix} \times S_C \times j_{ia} \times V\right]\\ &- \left[\pi r h \times S_H \times j_{ib} \times V\right] + \left[\left(\chi_a j_a + \chi_c j_c\right) \times V\right]\end{aligned} \tag{6}$$

where $V = \mathrm{M_{nacn}}/_{nacn}$, $M_{nacn} (\approx 12g)$, and ρ_{nacn} are the mass and density of the nanocone, D_1 and D_2 are surface diffusion coefficients, $\mathrm{A_k}$ is the carbon monomer incorporation speed into nanocone, $E_{INc} \approx (0.4$ eV) is the energy barrier for carbon monomer to diffuse along the nanocone-catalyst interface, respectively. Activation energy for surface diffusion of carbon monomer and of carbon clusters are $E_{SDc} \approx 0.3$ eV, $E_{SDcl} \approx 0.48$ eV along the nanocone-catalyst interface. $(c_a = c_c) \approx 1$ are the sticking coefficient of neutral atom of type a and c [39, 54, 55]. j_{ia}, j_{ib} are the ion collection current of type a and b over nanocone surface, j_a, j_c are the neutral collection current of type a and c over nanocone surface.

Defects Formation on Nanocone Surface

Time variation of energy balance at the nanocone surface in C_2H_2/NH_3 plasma is described by Eq. (7). The R.H.S of Eq. (7) describes the power transferred to the nanocone surface,

due to incidence of plasma species i.e., positive charged ions, neutrals, and electrons (first term), due to neutrals generation at nanocone surface (second term), due to hydrogen assisted etching of nanocone surface (third term), due to sputtering by incident ions (fourth term) give rise to defect on nanocone surface, due to surface diffusion along with the incorporation of created carbon species over nanocone surface (fifth term). Time variation of linear defect density on nanocone surface is shown by L.H.S. of Eq. (7)

$$P = M_{nacn} C_c T_s \frac{1}{Nd}\frac{\partial Nd}{\partial t} = \left[\sum_{k}^{a,b,c}\left[\left[\left(j_{ik}^{c_{cur}} + j_{ik}^{c_{tip}}\right)\in_{ij}^{c}\right] + \left[\left(j_{k}^{c_{cur}} + j_{k}^{c_{tip}}\right)\in_{k}^{c}\right]\right] + \left(j_{e}^{c_{cur}} + j_{e}^{c_{tip}}\right)\in_{e}^{c}\right] + \left[\frac{3}{2}k_B T_s \sum_{k}^{a,b,c}\left(1-\chi_{ik}\right)\left(j_{ik}^{c_{cur}} + j_{ik}^{c_{tip}}\right)\right] + \left[\left(\upsilon_o \sigma_{ads} J_b E_B\right)A_c\right] + \left[\left(\sum_{k}^{a,b,c} J_{ik} Y_s\left(E_i\right)\left(1-\theta_t\right)A_c\right)\right] + \left[S_C \omega \exp\left(\frac{-E_{SDc}}{k_B T}\right)A_c E_{INc}\right] \tag{7}$$

where M_{nacn}, C_c, and $A_c\left(=\pi r^2 + \pi r h\right)$ are the mass, specific heat, and area of nanocone. *Nd* is the linear (micron length) defect density produced on nanocone. $J_k = \left(\frac{n_k u_{thk}}{4}\right)$ and $J_{ik} = \left(n_{ik}\left(\frac{k_B T_e}{M_{ik}}\right)^{0.5}\right)$ denoting impinging flux composing neutrals and positive charged ions of type k, respectively [56]. $\in_{ik}^{c} = \left[\left(\left(\frac{2-Z\Psi_\beta}{1-Z\Psi_\beta}\right) - Z\Psi_\beta\right)k_B T_\beta\right]$ denotes the average energy accumulated over nanocone surface owing to β species, β signifies electrons, neutrals, and ions of type k, where $\Psi_\beta = \left(\frac{E_B}{k_B T_\beta}\right)$ and E_B is the energy barrier at the nanocone surface [45]. $j_{ik}^{c_{tip}} = \left[\pi r_1^2\left(\frac{8k_B T_{ik}}{\pi M_{ik}}\right)^{0.5} n_{ik}(z)\left(1 - Z\chi_V\right)\exp\left(\frac{-eU_V}{k_B T_s}\right)\right]$ denotes ion collection current due to positively charged species at the nanocone tip. $j_{ik}^{c_{cur}} = \left[n_{ik}(z) r_1 h\left(\frac{2\pi k_B T_{ik}}{M_{ik}}\right)^{0.5}\left(2\left(\frac{eV_V}{\pi k_B T_{ik}}\right)^{0.5} + \exp\left(\frac{eV_V}{k_B T_{ik}}\right) erfc\left(\frac{eV_V}{k_B T_{ik}}\right)^{0.5}\right)\exp\left(\frac{-eU_V}{k_B T_s}\right)\right]$ is the ion collection current over curved surface of nanocone due to positively charged species. The charge number over the nanocone surface is denoted by *Z*. $n_{ik}(z) = n_{iko}\left(1 - \frac{2e\varphi(z)}{M_{ik}\vartheta_{iko}}\right)$ is the positive ion density in plasma sheath as a function of structure coordinate. u_{thk} is thermal velocity of type k neutral, respectively. ϑ_{iko} is velocity of positive ion, $\varphi(z) = \varphi_o \exp\left(\frac{-|z|}{\lambda_d}\right)$ denoting static potential inside the plasma sheath, φ_o is the negative potential at the substrate surface. $\chi_V = eV_{ctip}(z), V_{ctip}(z) = \left(\frac{e}{4\pi\varepsilon_o r_1}\right)$ is the potential at the tip of nanocone. M_{ik}, T_{ik}, V_V are the mass, temperature of positive ion of type k, potential at curved surface of nanocone [57].

$$V_V = \sigma\frac{\left(3\sqrt{h^2 - 2hz + 2z^2} + (-2h + 4z)\log\left[\frac{-2h+4z+2\sqrt{2}\sqrt{h^2-2hz+2z^2}}{\sqrt{2}}\right]\right)}{4\sqrt{2}\varepsilon_o}$$

$j_e^{c_{tip}} = \left[n_e(z)\pi r_1^2\sqrt{\left(\frac{8k_B T_e}{\pi m_e}\right)}\exp\left[\frac{Z\chi_V}{k_B T_e} + \frac{eU_V}{k_B T_s}\right]\right]$ and $j_e^{c_{cur}} = \left[n_e(z) r_1 h\sqrt{\left(\frac{2\pi k_B T_e}{m_e}\right)}\exp\left[\frac{eV_V}{k_B T_e} + \frac{eU_V}{k_B T_s}\right]\right]$ denotes the collection current due to electrons over the tip and lateral surface of the nanocone. We will herein call these collections as OML collections. For OML theory to apply $r << \lambda_d$, where r is the base radius of nanocone and λ_d is the Debye length [58–60]. $n_e(z) = n_{eo}\exp\left(\frac{|e|\varphi(z)}{k_B T_e}\right)$ is the electron density inside the plasma sheath in response to

structure coordinate. $j_k^{c_{tip}} = n_k(z)r_1^2\left(\frac{8\pi k_B T_k}{M_k}\right)^{0.5}$ and $j_k^{c_{cur}} = n_k(z)\pi r_1 h\left(\frac{2\pi k_B T_k}{M_k}\right)^{0.5}$ are the neutral collection current over tip and lateral surface of nanocone. $\upsilon_o \approx \left(10^{15}\mathrm{cm}^{-2}\right), \sigma_{ads} \approx \left(10^{-16}\mathrm{cm}^2\right)$ is the area density of adsorption sites, reaction cross-section with atomic hydrogen [41]. $Y_s\left(E_i\right)$ is the sputtering yield function of ion energy [61]. θ_t is the total surface coverage. Using mass conservation, area of the resulting defect is estimated to be $Rd^2 = \left(\frac{r_1^2 + r_1 h}{hNd}\right)$, where Rd is the defect radius.

Rate of Growth of Hydrogen Radical and Carbon Species Over Nanocone Surface

Interaction between neutrals and ions of type a, b, c with nanocone surface serves formation of hydrogen radical and carbon species over nanocone surface through surface process that are described by Eqs. (8) and (9)

$$\begin{aligned} S_H^{\cdot} = &\sum_b J_b(1-\theta_t) + \sum_a n_{sa}\omega \exp\left(\frac{-\delta E_{td}}{k_B T_s}\right) + \sum_{ia}\left(\sum_a \frac{n_{sa}Y_d}{\upsilon_o}\right)J_{ia} + \sum_{ia} J_{ia} - \sum_b n_{sb}\omega \exp\left(\frac{-\delta E_{ah}}{k_B T_s}\right) \\ &- \sum_b n_{sb}\sigma_{ads}J_b - \sum_{ib}\left(\sum_b n_{sb}\sigma_{ads}\right)J_{ib} + \sum_{ic}\left(\sum_c \frac{n_{sc}Y_d}{\upsilon_o}\right)J_{ic} - \sum_{ic} J_{ic} \end{aligned} \tag{8}$$

$$\begin{aligned} S_C = &\sum_a J_a(1-\theta_t) + \sum_a n_{sa}\omega \exp\left(\frac{-\delta E_{td}}{k_B T_s}\right) + \sum_{ia}\left(\sum_a \frac{n_{sa}Y_d}{\upsilon_o}\right)J_{ia} + \sum_{ia} J_{ia} + \sum_{ib}\left(\sum_{ia} \frac{J_{ia}\sigma_{ads}}{\omega}\right)J_{ib} \\ &- S_C\omega \exp\left(\frac{-Eevp}{k_B T_S}\right) - \Gamma_C \end{aligned} \tag{9}$$

where $S_H^{\cdot}$ and $S_C^{\cdot}$ are the instantaneous surface concentrations of hydrogen radical and carbon species generated on the nanocone surface, $\delta E_{td}(= 2.1\,\mathrm{eV})$ hydrocarbons thermal dissociation energy, $\delta E_{ah}(= 1.8\,\mathrm{eV})$ hydrogen species adsorption energy, $E_{evp}(= 1.8\,\mathrm{eV})$ carbon evaporation energy, $n_{sk}\left(= \theta_k \upsilon_o\right)$ is surface concentration of neutral species of type k. θ_k is the surface coverage by species of type k, $\omega\left(\approx 10^{13}\mathrm{Hz}\right)$ is thermal vibration frequency, $Y_d\left(\approx 2.49 \times 10^{-2} + 3.29 \times 10^{-2} \times E_i\right)$ is stitching probability [43, 61], respectively. The surface processes are explained in Table 2.

Growth of Graphene Island

When energetic ions strike the surface of nanocone, defects are produced on the nanocone surface. This ion irradiation produced defects on nanocone surface are energetically unstable. Later these defects capture hydrocarbon species and form carbon monomers after several processes to attain stability. Carbon monomers afterward diffuse and accumulate to make carbon clusters. As time elapses the as formed carbon clusters along with diffusion, accumulate to make graphene nuclei which upon further diffusion and accumulation form graphene island that in turn diffuse and stitch together to produce larger islands [37, 39].

Further details regarding the growth rate equation of graphene island and explanation of growth of Nitrogen-doped Vertical-Graphene (N-VG) Sheet over nanocone, can be found in supplementary information.

Table 2 Explanation of surface process expressions used in Eq. (8) and (9)

Expressions	Explanation and reaction involved
$\sum_b J_b(1-\theta_t)$	Hydrogen Adsorption on nanocone surface $H_{(plasma)} \rightarrow H_{(ads)}$ $H_{2(plasma)} \rightarrow H_{2(ads)}$
$\sum_a n_{sa}\omega \exp\left(\frac{-\delta E_{td}}{k_B T_s}\right)$	Thermal dissociation of adsorbed hydrocarbons $C_2H_{2(ads)} \rightarrow 2C_{(ads)} + 2H_{(ads)}\, CH_{(ads)} \rightarrow C_{(ads)} + H_{(ads)}$
$\sum_{ia}\left(\sum_a \frac{n_{sa}Y_d}{\upsilon_o}\right)J_{ia}$	Ion-induced dissociation of adsorbed hydrocarbons $C_2H_2^+ + C_2H_{2(ads)} \rightarrow 2C_{(ads)} + H_{2(plasma)} + C_2H_2^+$ $CH^+ + CH_{(ads)} \rightarrow C_{(ads)} + H_{(ads)} + CH^+$ $C_2H_2^+ + CH_{(ads)} \rightarrow C_{(ads)} + H_{(ads)} + C_2H_2^+$ $CH^+ + C_2H_{2(ads)} \rightarrow 2C_{(ads)} + H_{2(plasma)} + C\,H^+$
$\sum_{ia} J_{ia}$	Hydrocarbons ion decomposition $C_2H_2^+ \rightarrow 2C_{(ads)} + H_{2(plasma)}$ $CH^+ \rightarrow C_{(ads)} + H_{(ads)}$
$\sum_b n_{sb}\omega \exp\left(\frac{-\delta E_{ah}}{k_B T_s}\right)$	Hydrogen desorption $H_{(ads)} \rightarrow H_{(des)}$
$\sum_b n_{sb}\sigma_{ads} J_b$	Adsorbed hydrogen species loss due to interaction with incoming hydrogen atoms from plasma
$\sum_{ib}\left(\sum_b n_{sb}\sigma_{ads}\right)J_{ib}$	Adsorbed hydrogen species loss due to interaction with incoming hydrogen ions from plasma
$\sum_{ic}\left(\sum_c \frac{n_{sc}Y_d}{\upsilon_o}\right)J_{ic}$	Ion induced dissociation of adsorbed ammonia $NH_3^+ + NH_{(ads)} \rightarrow N_{(ads)} + H_{(ads)} + NH_3^+$ $NH_2^+ + NH_{2(ads)} \rightarrow N_{(ads)} + 2H_{(ads)} + NH_2^+$
$\sum_{ic} J_{ic}$	Loss of hydrogen due to Ammonia ion decomposition $NH_2^+ \rightarrow N_{(ads)} + H_{2(plasma)}$ $NH_4^+ \rightarrow N_{(ads)} + 2H_{2(plasma)}$
$\sum_a J_a(1-\theta_t)$	Adsorption of hydrocarbon on nanocone surface $C_2H_{2(plasma)} \rightarrow C_2H_{2(ads)}$ $CH_{(plasma)} \rightarrow CH_{(ads)}$
$S_c\omega \exp\left(\frac{-Eevp}{k_B T_S}\right)$	Evaporation of Carbon $C_{(ads)} \rightarrow C_{(ev)}$
$\sum_{ib}\left(\sum_{ia} \frac{J_{ia}\sigma_{ads}}{\omega}\right)J_{ib}$	Carbon generation on N-NCN under hydrocarbon and hydrogen ions interaction $CH^+_{(plasma)} + H^+_{(plasma)} \rightarrow C_{(ads)} + H_{2(plasma)}$

Charging of N-NCN-VG Hybrid:

Charging of N-NCN-VG Hybrid takes place due to accretion of positive ions and electrons on N-NCN-VG Hybrid surface.

$$Z^{\cdot} = \sum_k^{a,b,c} j_{ik}^{c_{tip}} + \sum_k^{a,b,c} j_{ik}^{c_{cur}} + \sum_k^{a,b,c} j_{ik}^{gr} - \gamma_e\left(j_e^{c_{tip}} + j_e^{c_{cur}} + j_e^{gr}\right) \tag{10}$$

where Z is the charge number over N-NCN-VG Hybrid.

$$j_{ik}^{gr} = n_{ik}(z)\left(l_{gr}t_{gr} + t_{gr}h_{gr} + h_{gr}l_{gr}\right)\left(\frac{k_B T_{ik}}{2\pi^2 m_{ik}}\right)^{0.5} \times \left(\begin{array}{c}\frac{2}{\sqrt{\pi}}\left(\frac{eV_V}{k_B T_{ik}}\right) \\ +\exp\left(\frac{eV_V}{k_B T_{ik}}\right) erfc\left(\frac{eV_V}{k_B T_{ik}}\right)\end{array}\right) \times \exp\left(\frac{-eU_V}{k_B T_s}\right)$$

$j_e^{gr} = n_e(z)\left(l_{gr}t_{gr} + t_{gr}h_{gr} + h_{gr}l_{gr}\right)\left(\frac{k_B T_e}{2\pi^2 m_e}\right)^{0.5}\left(\exp\left(\frac{eV_V}{k_B T_e} + \frac{eU_V}{k_B T_s}\right)\right)$ denotes the ion and electron collection current at the VG surface. $\gamma_e(=1)$ is sticking coefficient of electron. In Eq. (10),

the first, second, and third term on the right-hand side describes the charge developed on nanocone tip, nanocone curved surface, and VG surface owing to positive charged ions accretion of type a, b and c. The last term indicating the reduction in charge due to electron accretion at the nanotip, nanocone curved surface, and VG sheet.

Number Density Equalization of Electrons in Plasma

Equation (11) represents number density equalization of electrons in plasma

$$n_e^{\cdot} = \sum_{k}^{a,b,c} \delta_k n_k - \sum_{k}^{a,b,c} \alpha_k n_e n_{ik} - \chi_e n_{hyb}\left(j_e^{c_{tip}} + j_e^{c_{cur}} + j_e^{gr}\right) \tag{11}$$

where δ_k and $\alpha_k\left(=\alpha_{ko}\left(\frac{300}{T_e}\right)^k\right)\mathrm{cm}^3\ \mathrm{sec}^{-1}$ are the ionization coefficient of constituting neutral atoms through external means and coefficient of recombination of electron and positively charged ions with $k=-1.2$ being a constant. n_e and T_e are the electron number density in plasma and electron temperature, respectively.

The first-term on RHS of Eq. (11) embodies per unit time gain in electron density through neutral atom (type a, b, c) ionization. Second, along with the third term represents electron -ion recombinational loss in electron density per unit time along with loss because of electron collection current over N-NCN-VG Hybrid surface.

Number Density Equalization of Plasma Species

Equation (12a)–(12c) represents the number density equalization of positively charged ions, in plasma

$$n_{ia}^{\cdot} = \delta_a n_a - \alpha_a n_e n_{ia} - n_{hyb}\left(j_{ia}^{c_{tip}} + j_{ia}^{c_{cur}} + j_{ia}^{gr}\right) - J_{adia} + J_{desia} + \frac{P}{E_{disa}V} + \sum_i k_{ia} n_a n_{ic} \tag{12a}$$

$$n_{ib}^{\cdot} = \delta_b n_b - \alpha_b n_e n_{ib} - n_{hyb}\left(j_{ib}^{c_{tip}} + j_{ib}^{c_{cur}} + j_{ib}^{c_{gr}}\right) - J_{adib} + J_{desib} + J_{th} + \frac{P}{E_{disb}V} + \sum_i k_{ib} n_b n_{ic} \tag{12b}$$

$$n_{ic}^{\cdot} = \delta_c n_c - \alpha_c n_e n_{ic} - n_{hyb}\left(j_{ic}^{c_{tip}} + j_{ic}^{c_{cur}} + j_{ic}^{c_{gr}}\right) - J_{adic} + J_{desic} + \frac{P}{E_{disc}V} + \sum_i k_{ia} n_a n_{ic} + \sum_i k_{ib} n_b n_{ic} \tag{12c}$$

$$n_a^{\cdot} = \alpha_a n_e n_{ia} - \delta_a n_a + n_{hyb}(1-\chi_{ia})\left(j_{ia}^{c_{tip}} + j_{ia}^{c_{cur}} + j_{ia}^{gr}\right) - n_{hyb}(\chi_a)\left(j_{ia}^{c_{tip}} + j_{ia}^{c_{cur}} + j_{ia}^{gr}\right) - \sum_i k_{ia} n_a n_{ic} \tag{13a}$$

$$n_b^{\cdot} = \alpha_b n_e n_{ib} - \delta_b n_b + n_{hyb}\left(1-\chi_{ib}\right)\left(j_{ib}^{c_{tip}} + j_{ib}^{c_{cur}} + j_{ib}^{gr}\right) - n_{hyb}\left(\chi_b\right)\left(j_{ib}^{c_{tip}} + j_{ib}^{c_{cur}} + j_{ib}^{gr}\right) - \sum_i k_{ib} n_b n_{ic} \tag{13b}$$

$$n_c^{\cdot} = \alpha_c n_e n_{ic} - \delta_c n_c + n_{hyb}\left(1-\chi_{ic}\right) \; - n_{hyb}\left(\chi_c\right)\left(j_{ic}^{c_{tip}} + j_{ic}^{c_{cur}} + j_{ic}^{gr}\right) - \sum_i k_{ib} n_b n_{ic} - \sum_i k_{ia} n_a n_{ic} \tag{13c}$$

where $J_{adik} = \frac{P_k}{\left(2\pi M_{ik} k_B T_{ik}\right)^{0.5}} \times \frac{n_{ik}}{J_{ik}}$ is the adsorption flux on the N-NCN-VG surface. P_k denotes adsorbing species partial pressure. $J_{desik} = J_{ik}\omega \exp\left(-\frac{E_{ads}}{k_B T_{ik}}\right)$ is the desorption flux(type k) from the N-NCN-VG surface. E_{ads} denotes adsorption energy. J_{th} denotes type b ions flux due to thermal dehydrogenation [56]. E_{disk} is the dissociation energy of neutrals of type k, $k_{ia} = 1.4 \times 10^{-9}$ cm^3/s, $k_{ib} = 7.8 \times 10^{-9} \exp\left(-14.8/T_n\right)/T_n^{0.39}$ cm^3/s are the ion –neural reaction rate coefficient [62, 63]. n_{hyb} is the N-NCN-VG number density.

The first term within the right-hand side of Eqs. (12a)–(12c) denotes the neutral atom ionization assisted gain in positive ion density per unit time, second, third, and fourth term denotes the decrease in positive ion density per unit time due to electron–ion recombination, ion collection current at the N-NCN-VG Hybrid surface and adsorption of ion at the surface of N-NCN-VG Hybrid. The fifth term denotes positive ion density gain per unit time through ions desorption out of the surface of N-NCN-VG Hybrid. The term J_{th} describes the number density increase of hydrogen ion in plasma due to thermal assisted dehydrogenation. The term $\frac{P}{E_{disk} V}$ describes the rise in positive ion density owing to applied power-assisted dissociation. E_{disk} is the dissociation energy of type k neutral. The last term of the Eq. (12a)–(12c) represents ion density gain with ion–neutral reaction.

Equations (13a)–(13c) represent neutral atom number density equalization in plasma, $\chi_{ik} (\approx 1)$ denote coefficients of sticking of ion of type k.

The first term on RHS of Eq. (13a)–(13c) represents the gain in neutral atom density per unit time through recombination of electron–ion, the second term represents ionization assisted decrease in neutral atom density, third and fourth term represents ion neutralization assisted, neutral atom density gain per unit time over N-NCN-VG Hybrid and loss in ion density per time through neutral atoms (type a, b, c) accretion on the N-NCN-VG Hybrid surface. The last term represents neutral atom density decrease through ion–neutral reactions.

Results and Discussion

An analytical model used in the present paper, explains the plasma sheath variations and growth of N-NCN-VG Hybrid (specifically the effect of plasma control parameters: total gas pressure and plasma power) including the effect on field emission characteristics of the hybrid. The analytical equations conferred above are worked out in parallel via MATHEMATICA SOFTWARE to study the effect of total gas pressure and input plasma power upon plasma constituents with properties that consequently regulate the dimensions of the N-NCN-VG, respectively. Appropriate boundary conditions, used to solve analytical equations namely, at t = 0 are number density of electron n_{eo} ($= 10^{10}$ cm^{-3}), electron temperature T_{eo} (= 2 eV), ion temperature T_{io} (= 0.16 eV), neutral temperature T_{no} (= 0.15 eV). Substrate potential $U_v = -300$ V, Substrate temperature $T_s = (650$ oC) Neutral atom mass

type a $M_a = \left[(12 \times x) + (1 \times y)\ a.m.u.\ for\ C_xH_y\right]$, Ion mass type a $M_{ia} = \left[(12 \times x) + (1 \times y)\ a.m.u.\ for\ C_xH_y^+\right]$, Neutral atom mass type b $M_b = \left[(1 \times x)\ a.m.u.\ for\ H_x\right]$, Ion mass type b $M_{ib} = \left[(1 \times x)\ a.m.u.\ for\ H_x^+\right]$, Neutral atom mass type c $M_c = \left[(14 \times x) + (1 \times y)\ a.m.u.\ for\ N_xH_y\right]$, Ion mass type c $M_{ic} = \left[(14 \times x) + (1 \times y)\ a.m.u.\ for\ N_xH_y^+\right]$ and volume of chamber $V_c = 1.05 \times 10^4$ cm^{-3}. $M_{HCN} = 27 a.m.u.$, $M_{HCN^+} = 27 a.m.u.$ Initial ion number density of $NH_3{}^+ = 10^7$ cm^{-3}, $NH_4{}^+ = 10^{10}$ cm^{-3}, $H_2{}^+ = 10^8$ cm^{-3}. Neutral atom density of type k $n_{ko} = 5 \times 10^{12}$ cm^{-3}, where k = a, c. Initial neutral density of $H = 10^{13}$ cm^{-3}, $HCN = 10^{11}$ cm^{-3}. Density of catalyst nickel $\left(\rho_{ct}\right) = 8.96$ gcm^{-3}, mass of catalyst nickel $\left(M_{cat}\right) = 58.96 a.m.u.$, diameter of catalyst nickel $\left(D^{\circ}\right) = 60$ nm.. First effect of change of total gas pressure $\left(P_o\right)$ upon growth of N-NCN-VG hybrid (e.g. height and radius of the nanocone, thickness, and height of the graphene sheet protruding from nanocone) due to variation in ion energy, electron temperature, and plasma composition with varying total gas pressure at fixed plasma power has been presented.

Figure 2a showing ion energy in eV variations with increasing total gas pressure. As gas pressure increases, electron mean free path [30] and sheath width decreases [64], resulting in increased collisions among electrons, neutrals, and ions, thus making them speedily lose their energy.

Figure 2b showing an increase in chemical sputtering/etching yield Y_I with increasing energy of bombarding hydrogen ions at normal angle of incidence on the growing N-NCN-VG Hybrid, following study conducted by Roth et al. [65]. Chemical sputtering/etching, affect the growth rate of Hybrid, as etching is proportional to the (i) flux of incoming hydrogen and nitrogen-based ions [66], (ii) the amount of C–C bond breaking occurrence for every ion (iii) the passivation probability of broken bond by atomic hydrogen and nitrogen.

Following this, Fig. 3a–c, shows decreasing etchant ion ($NH_3{}^+$, $H_2{}^+$,$NH_4{}^+$) number density with increasing gas pressure due to increase in collisional recombination reactions, following the study conducted by Mao et al. [30]. Along with it, relative hydrogen etchant density also decreases with increase of total gas pressure, [cf. Figure 3d], which can be explained from the fact that ionization, excitation, dissociation of constituent species i.e., H_2, NH_3, C_2H_2 due to electron impact decreases due to decrease in electron temperature and electron number density with increasing gas pressure [67], following study conducted by Mao et al. [30].

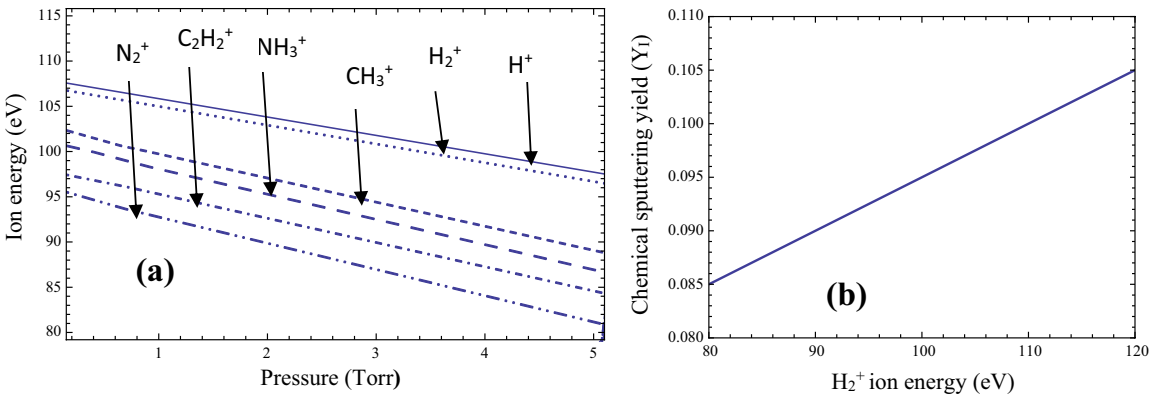


Fig. 2 **a** Ion energy (eV) variation with total gas pressure (Torr) at 300 W, **b** Variation of chemical sputtering yield with $H_2{}^+$ ion energy (eV)

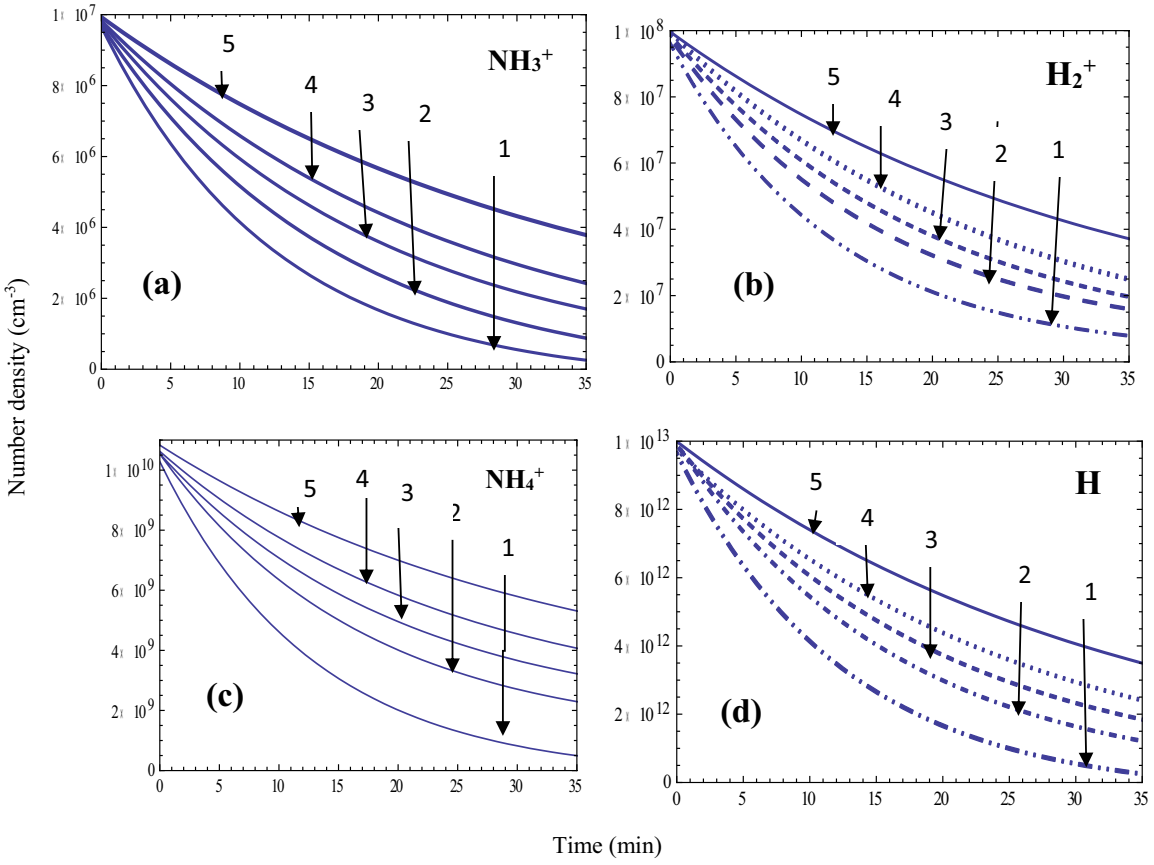


Fig. 3 Time variation of number density of ions **a** NH_3^+, **b** H_2^+, **c** NH_4^+ and **d** Time variation of number density of H atoms for different total gas pressure, (1 = 5 Torr), (2 = 1.2 Torr), (3 = 950mT orr), (4 = 700mTo rr) (5 = 200 mTorr)

Figure 4c shows surface concentration of carbon (S_C) increase with total gas pressure, ascribed to (1) reduced interaction of adsorbed hydrocarbon species on nanocone surface with etchant species from sheath. (2) Less nitrogen ion mediated reaction with hydrocarbon species takes place in plasma due to decreasing nitrogenous ion concentration, thus increasing the availability of carbon dimer (C_2) and carbon atoms for growth. (3) As C_2H_2 is the main growth precursor [68], so decreasing electron impact reaction with increasing pressure makes sufficient availability of C_2H_2. Consequently, growth rate of nanocone increases with increasing gas pressure [cf. Figure 4a], in compliance with studies of Parinov et al. [69], Chhowalla et al. [70], Ganjipour et al. [71]. Also extremely reactive H induced stabilization at relatively low pressure (250mTorr) limits the formation of sp^2 (highly crystalline structure) structure, thus slowing down nanocone growth at relatively low pressure.

Ion-induced dissociation of adsorbed hydrocarbons and ion decomposition reaction on growing nanocone decreases with an increase of plasma pressure [41] so limiting the carbon available for growth thereby preventing the formation of amorphous carbon.

In case of low pressure etching is more, resulting in nanocone of smaller tip radius [cf. Fig. 4b] following study conducted by Parinov et al. [69], Kato et al. [72]

Increasing the time of exposure of carbon nanocone under plasma environment gives rise to defect formation on the nanocone surface. Figure 4d illustrating the linear density of defects evolving with time over nanocone surface with varying total gas pressure. Defect generation on nanocone surface is relatively low at 200 mTorr, which further increases

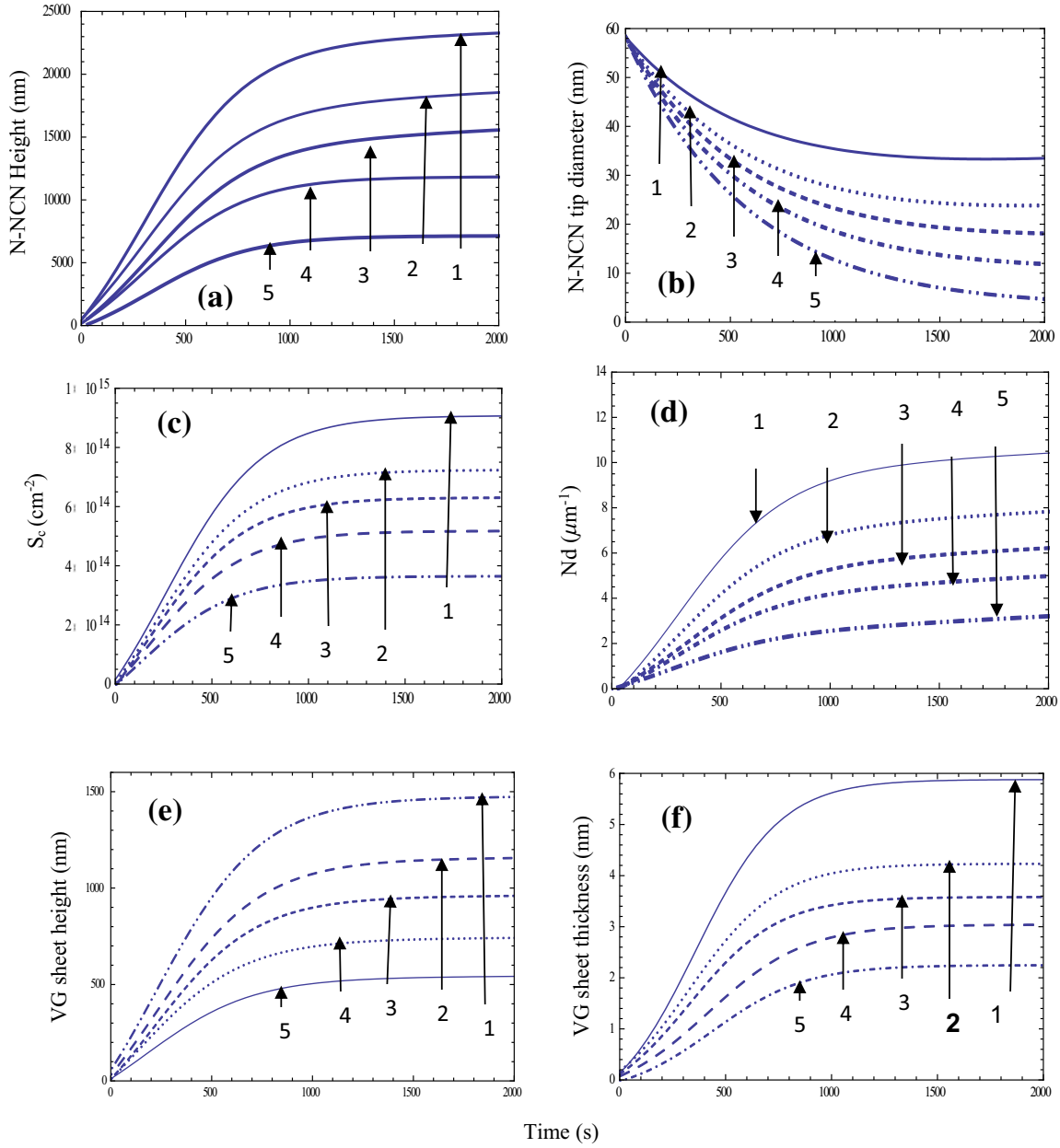


Fig. 4 Time variation of **a** N-NCN height **b** N-NCN tip diameter **c** carbon surface concentration (Sc), **d** linear density of defects (Nd), **e** VG sheet height and **f** VG sheet thickness for different total gas pressure. (1 = 5 Torr), (2 = 1.2 Torr), (3 = 950mTorr), (4 = 700mTorr) (5 = 200mTorr)

with gas pressure. This can be explained by the fact that defective structure is a function of carbon concentration on nanocone surface away from the catalyst nanoparticle [73]. At low-pressure etchant number density is high, so the number of carbon on the nanocone surface is low, whereas at high-pressure sufficient number of carbon are away from the catalyst nanoparticle surface which are available for the formation of defective structure. Also, the number of defects is proportional to the height of the nanocone [74], at low-pressure nanocone of small height is obtained, so number of defect on nanocone surface is low at relatively low pressure. An increase in the number of defects with gas pressure is following the study conducted by Garg et al. [75] where defect on edge of graphene increases with an increase in CH_4 gas pressure.

As discussed previously, nucleation of graphene sheet over nanocone surface is defect directed. Comparatively low pressure (200 mTorr) results in a drop in cluster formation on

decreased defect sites due to the availability of fewer carbon atoms. These clusters further diffuse and agglomerate to form a small graphene island. As a consequence VG sheet number density decreases with a decrease in total gas pressure following the study conducted by Garg et al. [75]. Moreover, low availability of carbon Sc and relatively more of etchant number density in reactive plasma at relatively low gas pressure, results in decrease in VG sheet growth rate as shown in Fig. 4e, following study conducted by Wei et al. [76].

Moreover, reasonably high etchant number density at relatively low pressure leads to the formation of VG sheet of comparatively less thickness as shown in Fig. 4f, following study conducted by Hao et al. [77], Wang et al. [78]

Figure 5a showing ion energy in eV variation with growth time at different input plasma power at fixed total gas pressure. An increase in plasma power at constant total gas pressure leads to an increase in plasma region volume accompanied by an increased electric field in plasma; as a consequence number density of electron, electron temperature, ion temperature, ionization of gases in the plasma increases.

Figure 6a shows reduced surface concentration of carbon (Sc) with increasing plasma power. This is attributed to (1) increase in dissociation of neutrals C_2H_2, NH_3 in plasma with increasing power makes less carbon available for growth, because it is mainly drawn from C_2H_2 reaction at nickel catalyst surface, [cf. Fig. 5b], showing the time evolution of dissociation of C_2H_2 at different plasma power, in accordance to Bell et al. [68]. (2) CH_4 production in the plasma increase with increase in power [79] and as C_2H_2 dissociates more readily on catalyst surface as compared to CH_4, so Sc available for growth decreases [28].

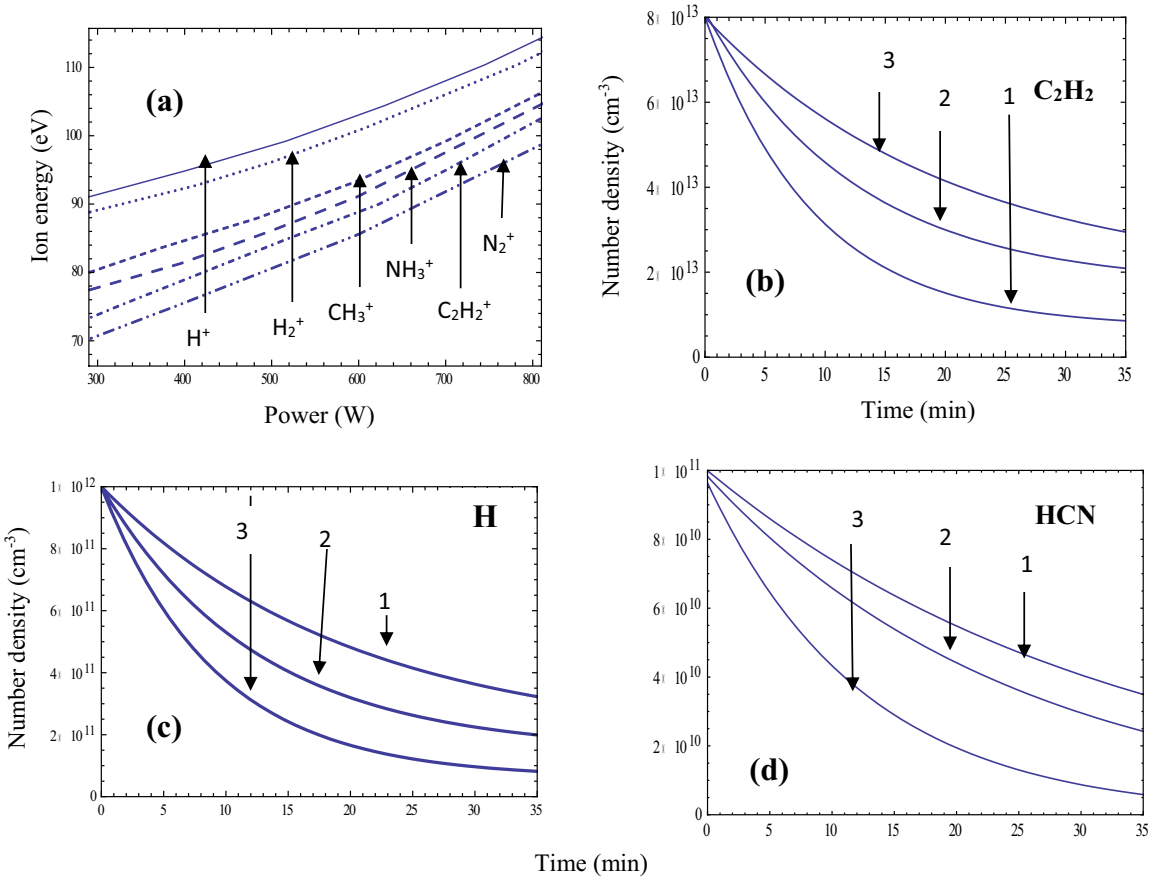


Fig. 5 **a** Variation of Ion energy (eV) with Input Power (W) at 5 Torr and Time variation of number density of **b** C_2H_2 **c** H **d** HCN for different input power, (1 = 700 W), (2 = 500 W), (3 = 300 W) at 5 Torr

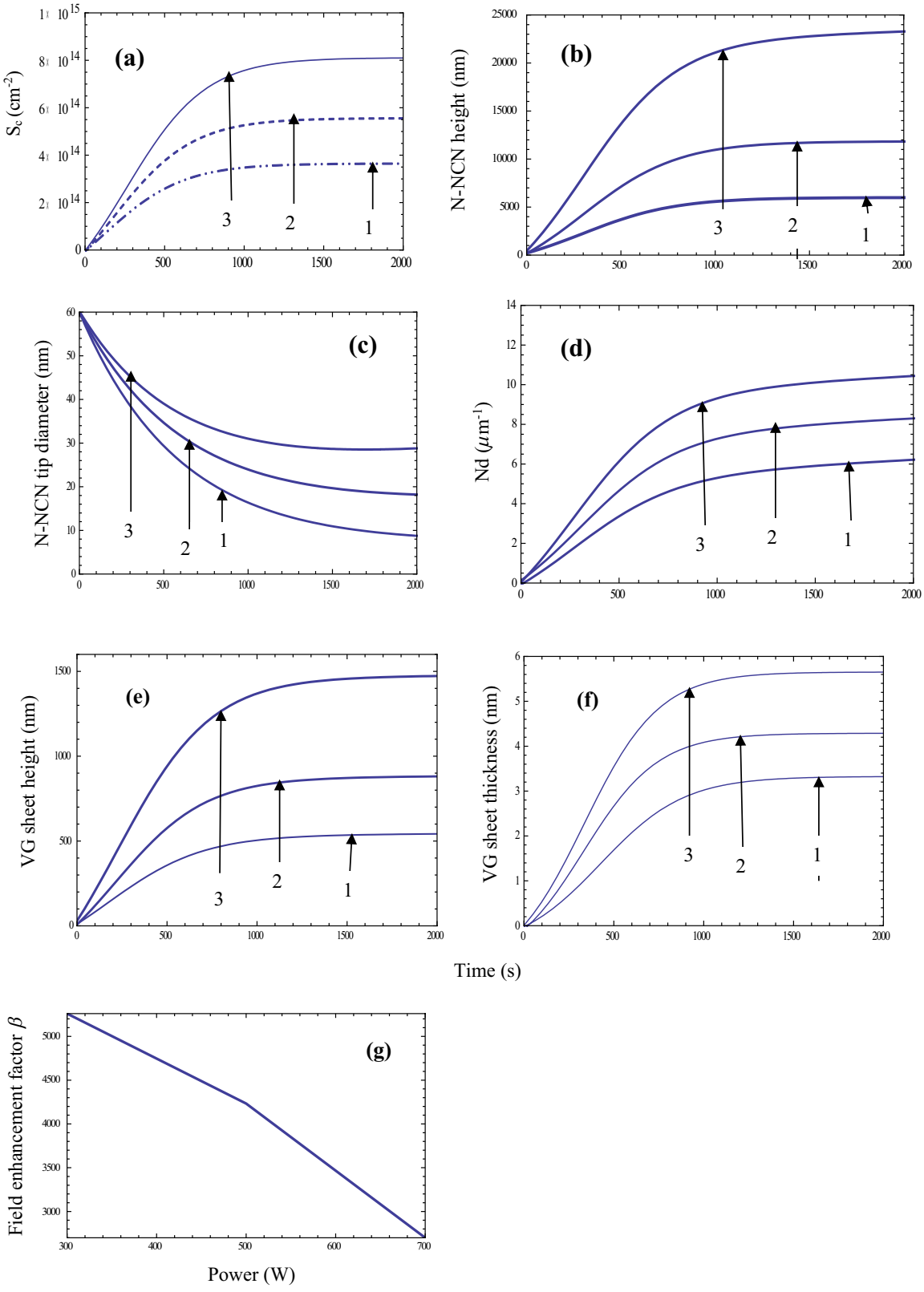


Fig. 6 Time variation of **a** carbon surface concentration (S_c), **b** N-NCN height **c** N-NCN tip diameter **d** linear density of defects (Nd) **e** VG sheet height **f** VG sheet thickness, for different input power (1 = 700 W), (2 = 500 W), (3 = 300 W) at 5 Torr and **g** Field Enhancement factor β of N-NCN Hybrid at different input power at 5 Torr

Consequently the growth rate (length) of nanocone, [cf. Figure 6b] drops with increasing plasma power. Decrease in Sc and decreased growth rate of nanocone with plasma power are following a study conducted by Deng et al. [28]. Moreover, increasing plasma power favors hydrogen etchant [28] production [cf. Fig. 5(c)] due to increased NH_3 and C_2H_2 dissociation, which also contributes to decline in growth rate of nanocone with plasma power. Further, as power increases the high energy etchants (H, H_2^+, NH_3^+, HCN^+) etches radius of nanotip so nanocone of smaller tip radius will be formed with increasing plasma power [cf. Fig. 6c] following the study of Merkulov et al. [80].

Figure 6d shows time varying, decreasing linear defect density over nanocone with increasing plasma power. This can be explained by availability of carbon (Sc) on the nanocone surface. As power increases from 300 to 700 W, Sc over nanocone surface decreases as discussed above, hence decreasing the defective structure along the nanocone surface [73]. Also, the number of defects is proportional to the length of the nanocone [74].

Relatively low Sc at nanocone surface at high plasma power leads to generation of small graphitic cluster at decreased defective sites. So we infer the formation of small size graphene island with decreased number density due to diffusion and agglomeration of these clusters. Hence, we outline that VG sheet number density decreases with increasing plasma power. In addition, (i) increase in active hydrogen concentration, with power, leads formation of volatile hydrocarbon species upon interaction with carbon on VG surface. (ii) Increase in HCN number density within plasma with plasma power, restrains carbon supply for VG sheet growth, owing to C–N triple bond strength 748 kJ mol^{-1} and H–CCH bond strength 556 kJ mol^{-1} [63]. This makes C_2H_2 to decompose preferably on nanocone surface than HCN, [cf. Fig. 5d] showing time-varying HCN number density in plasma with increasing plasma power [79]. Consequently, growth rate of VG sheet decreases with increasing plasma power [cf. Figure 6e].

Moreover, highly energetic etchants at high power etches the sidewalls of the growing VG sheets, so thickness of VG sheet decreases with increase of plasma power [cf. Figure 6f]. The decrease in growth rate and thickness of VG sheet with increase in plasma power complies with Deng et al. [28] and Nang et al [81].

As discussed above with increasing power, the height of the nanocone and VG sheet decreases along with tip radius of the nanocone and thickness of the VG sheet. The finding of the above work is employed to find the field enhancement factor of the N-NCN-VG hybrid with varying plasma power (300–700 W). The field enhancement factor of nanocone $\beta_{NCN} \approx \left(\frac{h}{r_1}\right)$ and of graphene $\beta_{VG} \approx \left(\frac{h_{gr}}{t_{gr}}\right)$, respectively. Therefore, field enhancement of the N-NCN-VG hybrid may be calculated as $\beta_{hybrid} = \beta_{NCN}\beta_{VG}$ [82]. But in that case, the order of field enhancement factor of hybrid is about two orders more than the experimental values of β_{hybrid}. So in the present work β_{hybrid} is estimated as a ratio of $(h + h_{gr})$ to t_{gr}. Figure 6g manifests that the field enhancement factor of N-NCN-VG hybrid decreases with increasing plasma power. Though thinner, nanocone and VG sheet are obtained at relatively high power, but nanocone and VG sheet (forming N-NCN-VG hybrid) obtained at relatively low power are much higher as compared to high power. Thus, obtaining N-NCN-VG hybrid of increased field enhancement factor at relative low plasma power [cf. Figure 6g]. Increase in β with decrease in plasma power comply with experimental works of Deng et al. [28].

Conclusion

Theoretical model, comprising plasma sheath kinetics and surface processes is manoeuvred, to study the growth characteristics (height, tip diameter of nanocone, thickness, and height of VG sheet) of the N-NCN-VG hybrid by varying the plasma control parameters i.e., total gas pressure and input plasma power. Effect of ion energy, etchant number density, carbon surface concentration has been investigated concerning plasma control parameters. It is observed that N-NCN and VG sheet growth rate increases with an increase of total gas pressure and decreases with an increase of plasma power. The findings of varying plasma power on growth rate of N-NCN-VG hybrid are employed to calculate the field enhancement factor. It turns out that the field enhancement factor decreases with an increase of plasma power. This theoretical result complies with the experimental observation of Deng et al. [28]. The present study result can be widened to be realized as efficient field emitter devices.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11090-022-10229-3.

Acknowledgements The author (Shruti Sharma) is thankful to the SERB-DST, Government of India for providing the necessary financial assistance

References

1. Cole MT, Parmee RJ, Milne WI (2016) Nanotechnology 27:8
2. Li X, Zhou J, Wu Q, Liu M, Zhou R, Chen Z (2019) J Vac Sci Technol B 37:051203
3. Liang X, Xia J, Dong G, Tian B, Peng I (2016) Top Curr Chem 374:80
4. Levchenko I, Xu S, Teel G, Mariotti D, Walker MLR, Keidar M (2018) Nat Commun 9:879
5. Zhang J, Wang C, Zhou C (2012) ACS Nano 6:8
6. Singh LA, Walker MLR, Sanborn GP, Turano SP, Ready WJ, "Operation of spindt-type, carbon nanotube cold cathodes in a Hall effect thruster environment," paper presented at the 33rd International Electric Propulsion Conference, Washington D.C., USA, 6- 10 October 2013, Paper IEPC-2013–348
7. Deng JH, Liu RN, Zhang Y, Zhu WX, Han AL, Cheng GA (2017) J Alloys Compd 723:75–83
8. Liu J, Zeng B, Wang X, Wang W, Shi H (2013) Appl Phys Lett 103:053105
9. Qi JL, Zhang F, Zhang LX, Cao J, Feng JC (2014) Mater Res Express 1:025601
10. Levchenko I, Xu S, Baranov O, Bazaka O, Ivanova EP, Bazaka K (2021) Molecules 26:4091
11. Bazaka K, Levchenko I, Lim JWM, Baranov O, Corbella C, Xu S, Keidar M (2019) J Phys D Appl Phys 52:18
12. Gong J, Yang H, Yang P (2015) Compos B Eng 75:250–255
13. Koh ATT, Chen T, Pan L, Sun Z, Chua DHC (2013) J Appl Phys 113:174909
14. Li C, Li Z, Zhu H, Wang K, Wei J, Li X, Sun P, Zhang H, Wu D (2010) J Phys Chem C 114:33
15. Park SK, Choi K, Lee SH, Oh IK, Park S, Park HS (2017) Carbon 116:500–509
16. Ye M, Hu C, Lv L, Qu L (2016) J Power Sources 305:106–114
17. Xu T, Yang D, Fan Z, Li X, Liu Y, Guo C, Zhang M, Yu ZZ (2019) Electrochim Acta 152:134–143
18. Park H, Ambade RB, Noh SH, Eom W, Koh KH, Ambade SB, Lee WJ, Kim SH, Han TH (2019) ACS Appl Mater Interfaces 11:9011–9022
19. Torrisi L, Rosinski M, Terwinska D, Tchorz P, Cutroneo M, Torrisi A (2020) Contrib Plasma Phys 60:1–15
20. Zhu Z (2017) Nanomicro Lett 9:25
21. Yap YK, Menda J, Vanga LK, Kayastha V, Wang J, King LB, Dimovski S, Gogotsi Y (2004) MRS Online Proc Libr 821:147–151
22. Kaushik V, Shukla AK, Vankar VD (2014) Appl Phys A Mater Sci Process 117:2197–2205
23. Xu Y, Deng P, Chen G, Chen J, Yan Y, Qi K, Liu H, Xia BY (2019) Adv Funct Mater 30:1906081
24. Jeong H, Nguyen DM, Lee MS, Kim HG, Ko SC, Kwac LK (2018) Mater Sci Eng C 90:38–45

25. Dasgupta K, Khosravifar M, Sawant S, Adusei PK, Kanakaraj SN, Kasik J, Shanov V (2020) C-J Carbon Res 6:40
26. Kotal M, Sharma A, Jakhar S, Mishra V, Roy S, Sahoo SC, Sharma HK, Mehta SK (2020) Cryst Growth Des 20:4627–4639
27. Gautier LA, Borgne VL, Khakani MAE (2016) Carbon 98:259–266
28. Deng JH, Cheng GA, Zheng RT, Yu B, Li GZ, Hou XG, Zhao ML, Li DJ (2014) Carbon 67:525–533
29. Gupta N, Sharma SC (2018) Phys Plasmas 25:073509
30. Mao M, Bogaerts A (2010) J Phys D Appl Phys 43:20
31. Wang BB, Xu XZ, Zhang B (2007) Appl Surf Sci 253:6951–6956
32. Levchenko I, Ostrikov K, Long JD, Xu S (2007) Appl Phys Lett 91:113115
33. Wang BB, Ostrikov K (2009) J Appl Phys 105:083303
34. Suzuki S (2011) In: Marulanda JM (ed) Low-Energy Irradiation Damage in Single-Walled Carbon Nanotubes, IntechOpen. https://www.intechopen.com/chapters/16851
35. Zeng L, Lei D, Wang W, Liang J, Wang Z, Yao N, Zhang B (2008) Appl Surf Sci 254:1700–1704
36. Bo Z, Mao S, Han ZJ, Cen K, Chen J, Ostrikov K (2015) Chem Soc Rev 44:2108–2121
37. Wang BB, Zhu K, Ostrikov K, Shao RW, Zheng K (2016) J Appl Phys 119:024302
38. Wang BB, Zheng K, Cheng QJ, Ostrikov K (2015) Appl Surf Sci 325:251–257
39. Mehdipour H, Ostrikov K (2012) ACS Nano 6:11
40. Economou DJ, Evans DR, Alkire RC (1988) J Electrochem Soc 135:756–763
41. Mehdipour H, Ostrikov K, Rider AE (2010) Nanotechnology 21:45
42. Mehdipour H, Ostrikov K, Rider AE, Han Z (2011) Plasma Processes Polym 8:386–400
43. Marvi Z, Xu S, Foroutan G, Ostrikov K (2015) Phys Plasmas 22:013504
44. Srivastava SK, Shukla AK, Vankar VD, Kumar V (2005) Thin Solid Films 492:124–130
45. Sodha MS, Misra S, Misra SK, Srivastava S (2010) J Appl Phys 107:103307
46. Ostrikov K, Xu S, Plasma-Aided Nanofabrication From Plasma Sources to Nanoassembly , Wiley Publishers Since 1807, https://download.e-bookshelf.de/download/0000/6030/70/L-G-0000603070-0002364804.pdf
47. Dang C, Wang BB (2006) Thin Solid Films 514:76–80
48. Bohdansky J, Roth J, Bay HL (1980) J Appl Phys 51:2861–2865
49. Hopf C, Jacob W (2005) J Nucl Mater 342:141–147
50. Schluter M, Hopf C, Jacob W (2008) New J Phys 10:053037
51. Chang JP, Arnold JC, Zau GCH, Shin HS, Sawin HH (1997) J Vac Sci Technol A 15:1853–1863
52. Sigmund P (1969) Phys Rev 184:383–416
53. Gupta N, Gupta N, Sharma SC (2018) Phys Plasmas 25:043504
54. Gupta R, Sharma SC (2018) Contrib Plasma Phys 59:72–85
55. Li J (2014) Thesis theoretical investigations of catalytic methane cracking and carbon nanotube growth. UWSpace. http://hdl.handle.net/10012/8629. https://uwspace.uwaterloo.ca/bitstream/handle/10012/8629/N_Li_Jingde.pdf?sequence=1&isAllowed=y
56. Sharma SC, Gupta N (2015) Phys Plasmas 22:123517
57. Tewari A, Sharma SC (2015) Phys Plasmas 22:023505
58. Smith HMM, Langmuir I (1926) Phys Rev 28:36–45
59. Sodha MS, Kinetics of Complex Plasmas, Springer Series on Atomic, Optical, and Plasma Physic**81** http://77.151.93.113/doc47/LIVRES%20Recopies/Physiques%20Atomic%2C%20Optical%2C%20and%20Plasma/Sodha%20M.S.%20-%20Kinetics%20of%20Complex%20Plasmas%20-%202014.pdf
60. Kennedy RV, Allen JE (2003) J Plasma Phys 69:485–506
61. Mantzaris NV, Gogolides E, Boudouvis AG, Rhallabi A, Turban G (1996) J Appl Phys 79:3718–3729
62. See http://www.udfa.net/ for database of neutral-ion reactions rate coefficients.
63. Hash DB, Bell MS, Teo KBK, Cruden BA, Milne WI, Meyyappan M (2005) Nanotechnology 16:925–930
64. Wang P, Lu J, Zhou O (2008) Nanotechnology 19:185605
65. Roth J, Bohdansky J, Poschenrieder W, Sinha MK (1976) J Nucl Mater 63:222–229
66. Luo Z, Lim S, You Y, Miao J, Gong H, Zhang J, Wang S, Lin J, Shen Z (2008) Nanotechnology 19:255607
67. Mao M, Bogaerts A (2011) J Phys Conference Series 275:012021
68. Bell MS, Teo KBK, Lacerda RG, Milne WI, Hash DB, Meyyappan M (2006) Pure Appl Chem 78:6
69. Parinov IA (2015) Piezoelectrics and Nanomaterials: Fundamentals Developments and Applications. Nova Science Publishers, New York, pp 1–283
70. Chhowalla M, Teo KBK, Ducati C, Rupesinghe NL, Amaratunga GAJ (2001) J Appl Phys 90:5308–5317

71. Ganjipour B, Mohrajerzadeh S, Hesamzadeh H, Khodadadi A (2005) A Fuller Nanotube Carbon Nanostruct 13:365–373
72. Kato T, Hatakeyama R (2010) J Nanotechnol 2010:256906
73. Ismail I, Mamat MS, Adnan NL, Yunusa Z, Hasan IH (2019) J Nanomaterials 2019:5717180
74. Ferrari AC, Basko DM (2013) Nat Nanotechnol 8:235–246
75. Garg R, Rastogi SK, Lamparski M, Barrera SC, Pace GT, Nuhfer NT, Hunt BM, Meunier V, Karni TC (2017) ACS Nano 11:6301–6311
76. Wei HW, Leou LC, Wei MT, Lin YY, Tsai CH (2005) J Appl Phys 98:044313
77. Hao J, Xu T, Liu W (2005) J Non-Crystalline Solids 351:3671–3676
78. Wang Y, Li L, Cheng Q, He C (2015) J Lumin 161:7–13
79. Teo KBK, Hash DB, Lacerda RG, Rupesinghe NL, Bell MS, Dalal SH, Bose D, Govindan TR, Cruden BA, Chhowalla M, Amaratunga GAJ, Meyyappan M, Milne WI (2004) Nano Lett 4:921–926
80. Merkulov VI, Hensely DK, Melechko AV, Guillorn MA, Lowndes DH, Simpson ML (2002) J Phys Chem B 106:10570–10577
81. Nang LV, Kim ET (2012) J Electrochem Soc 159:K93–K96
82. Lv S, Li Z, Liao J, Wang G, Li M, Miao W (2015) Sci Rep 5:15035

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

2021 IEEE 8th Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON) | 978-1-6654-0962-9/21/$31.00 ©2021 IEEE | DOI: 10.1109/UPCON52273.2021.9667649

Electric Vehicle Charging Standards in India and Safety Consideration

Sakshi Chalia
Delhi Technological University
Electrical Engineering Department
Delhi, India
sakshichalia_2k20pes18@dtu.ac.in

Aakash Kumar Seth
Delhi Technological University
Electrical Engineering Department
Delhi, India
akrseth@gmail.com

Mukhtiar Singh
Delhi Technological University
Electrical Engineering Department
Delhi, India
smukhtiar_79@yahoo.co.in

Abstract— **Electric vehicles (EVs) are very efficient, cheap, and cause less pollution to environment that internal-combustion (IC) engine base vehicles. As EVs are gaining lot of popularity globally so it becomes very important to standardize its infrastructure. This paper summarizes the various charging standards followed by different country regions globally and focuses the standards adopted by India. As of now, every standard identified with charging infrastructure depend on the European ambient conditions. As India being a tropical country, subsequently settling the EV charging standards, vehicle adoption pattern and India's encompassing temperature will become basic contemplations. Therefore, India has followed and modified various other standards to comply with Indian conditions. Moreover, the safety consideration for conductive EV charging utilized in India is also summarized.**

***Index Terms*— Electric Vehicles, Charging Levels, Charging Standards, Safety Consideration.**

I. INTRODUCTION

With increasing population, engine-based vehicle use may become 2.5 billion by 2050 which may lead to the shortage of fossil fuels and may covered all the cities with extreme air pollution [1]. To limit this, a paradigm shift in perspective is needed to maintainable vehicle it implies utilization of low or zero discharge vehicles, most extreme usage of public vehicle, utilization of sustainable power source.

Electric vehicles (EVs) are the best example of sustainable transportation. Its contribution in world's transportation sector to remove the harmful emissions is increasing day by day. Electric vehicles are the best replacement of internal-combustion (IC) engine-based vehicles for addressing enhancement in pollution, global warming, depleting natural resources etc. As gasoline, which is petrol, diesel, compressed natural gas (CNG) can be used as a fuel in vehicles. There are various types of pollutants and greenhouse gases are released as emission from vehicles which becomes a reason for smog, and air pollution. The particulates released in the combustion cycle are in range of micrometers. Due to small size these cannot be filtered out by human nose/throat and effects heart or lungs after entering the body [2].

However, in terms of efficiency, an EV has higher energy efficiency. If fuel tank is compared to vehicle efficiency, an IC engine-based vehicle has 16% efficiency. Whereas an EV and fuel cell electric vehicle (FCEV) provides 72% and 40% efficiency respectively at the wheels if the starting source is taken as 100% [3].

Despite of so many benefits of EV, there are some issues that need to be considered are limited driving range, high costs, battery life issues and charging. It is found that driving range go down by 41% when the temperature dropped to 20 degrees Fahrenheit [4]. However, the driving range can be extended by increasing the energy levels of battery pack, but it can be increased only up to a certain size and mass to achieve good results. Charging infrastructure is also a problem for EVs, a smaller number of charging station are built as compared to gas stations.

Generally, EV has battery pack with Li-ion cells and to charger this pack, three types of charging Levels are utilized i.e., level 1, level 2 & level 3. Level 1 is slow charging corresponds to supply voltage of 110V–120 V while level 2 (semi-fast) is of 220V–240 V and level 3 (fast) or DC fast charging (DCFC) is of 200V–800 V [1]. On-board battery chargers (OBC) are restricted by size, weight, and volume for this reason they are typically viable with level 1 (supports 110-120V) and level 2 (220-240V) chargers. The level 3 (200-800V) charger requires high charging voltage and current, therefore it cannot be installed onto the vehicle as it is higher in weight and requires larger space [5].

As EVs are evolving globally, it becomes essential to standardize the related aspects for the uniform operation. ISO i.e., International Organization for Standardization is among the EV charging component standardization organizations that deals with EVs standardization and the rest works on the specification of component level [2]. Every country follows different standards like society of automotive engineers (SAE), institute of electrical and electronics engineers (IEEE), international electrotechnical commission (IEC), Guobiao (GB/T) and CHAdeMO. The USA follows IEEE standards where IEC standards is utilized by Europe. The CHAdeMO standards is followed by Japan. Whereas the IEC and SAE are mostly used standards in worldwide [6]-[8].

The EV charging standards used in India are AC-001 specified by Indian standard (IS): 17017 and DC-001 specified by IS: 17017-1 for low voltage EV applications. India is in the stage of developing the EV standards for medium as well as high voltage applications. Whereas right now India is using IEC

978-1-6654-0962-9/21/$31.00 ©2021 IEEE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:46:42 UTC from IEEE Xplore. Restrictions apply.

standards for AC high voltage charging and SAE/CHAdeMO for DC high voltage charging applications [9]-[11].

With the global regulatory landscape dynamic evolvement, the standardization of charging infrastructure is very crucial and with this electrical safety is a key concern. Safety, compatibility, and performance are three basic regulation aspects. In India, some safety regulations of international and national standard are taken up by Indian agencies like automotive industry standards committee (AISC) for featuring the need of harmonization [12].

This paper discusses about the international standards for EV charging and further focus on charging standards used by Indian organizations and safety requirements.

II. Charging Power Level and International Standards

Gas and diesel-based vehicles have been in existence for more than 100 years, and the business has built up a vigorous refueling foundation for the help of billions of vehicles. Nonetheless, with the new ascent in emerging trends of electric vehicles, new "refueling" methods are developing practices. The three EV charging levels i.e., level 1, level 2 & level 3 (DCFC) are summarized in table I.

Table I

Summary of charging levels standards for electric vehicles.

Charging Specification	**Level 1**	**Level 2**	**Level 3 (DCFC)**
Voltage	120 V AC @60Hz	240 V AC @60Hz	200-800V D.C
Current (A)	15	40	60
Charge Time	Up to 20 hrs.	Up to 7 hrs.	Up to 30 min
Vehicle Charge Port	Nema 515, Nema 520	SAE J1772, Nema 1450, Nema 6-50	CHAdeMO, SAE
Power Equivalent	Toaster	Clothes dryer	15 Central A.C.
Installation	Self	Professional	Professional

Level 1: The basic EV charging level is Level 1 charger. It is charged from a standard 120 V household outlet, which gives around 4 to 5 miles of reach in one hour of charging. Most of EV producers incorporate Level 1 EVSE cord set, therefore no extra charging hardware is required. A totally drained EV battery requires around 20 hours to fully re-energize [3]. Level 1 charging is for the most part kept to North, South and Central America, Europe and a large part of the world uses a 220 V inventory framework. Around 500-880 USD installation cost is required for this type of charging method.

Level 2: Level 2 charging equipment utilizes 220 V supply system private or 208 V for an outside charging station [3]. A 3.3 KW on-board charger gives around 15 miles of travelling in an hour of charging. However, a 6.6 KW on-board charger gives around 30 miles of travel in same time span of charging. Level 2 EVSE uses gear explicitly intended to give speed up re-charging and requires proficient electrical establishment utilizing a devoted electrical circuit. Around 1000-3000 USD establishment cost is required for Level 2 charging system.

Level 3 OR DCFC: The Level 3 charging system requires 480 V AC supply system. This type of charger gives around 80-100 miles of travelling in 20-30 minutes of charging. The DCFC converts AC supply into DC and then DC supply is sent directly to EV battery pack. This level of charging is suitable only for commercial usage not for residential locations. The executing cost is between 30000-160000 USD. The summary of all charging levels is listed in Table I.

Electric Vehicle Charging Standards: - There are a few guidelines accessible globally which manage EV charging foundation. Fig. 1 shows the country region and their EV charging standard organization. The SAE and IEEE standards are utilized in USA. based makers while IEC is limitlessly utilized in Europe. Japan and China have their own EV charging norms named CHAdeMO and Guobiao (GB/T) standard respectively. Guobiao (GB/T) standard given by the Standardization Administration of China and Chinese National Committee of ISO and IEC for AC and DC charging. However, GB/T AC charging principles are like IEC norms. All these major standards are explained below and summarized in Tables.

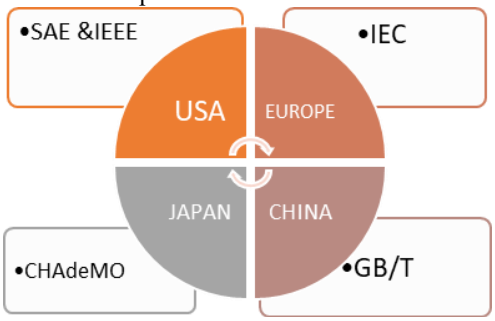


Fig. 1. EV charging standards followed by various countries.

SAE Standards: The SAE international have 128,000 members from over 100 nations. This standard for Level 1 and level 2 EVSE is accommodated onto the vehicle. Whereas for fast DC charging it should be outside the vehicle. The SAE worldwide coordinates the advancement of technical standards based on best practices distinguished and portrayed. They developed the EV charging standards for both AC and DC supply system. SAE: J1772 discuss all the equipment ratings for EV charging includes circuit breaker, current rating of charging voltage and so on. The voltage and current ratings of AC and DC charging system according to SAE: J1772 are listed in Table II and III respectively [13].

Table II

SAE AC charging levels

Charging System	**Supply System**	**Maximum Current (A)**	**Branch Circuit Breaker Rating (A)**	**Power Rating (KW)**
AC Level 1	1-φ, 120 V	12	15	1.44
		16	20	1.92
AC Level 2	1-φ, 208-240 V	≤80	AS PER NEC625	UP TO 19.2

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:46:42 UTC from IEEE Xplore. Restrictions apply.

Table III
SAE DC charging levels

DC Charging System	EVSE DC Output Voltage (V)	Maximum Current (A)	Power Rating (KW)
Level 1	200-450	80	36
Level 2	200-450	200	90
Level 3	200-600	400	240

IEC Standards: The IEC is a non-benefit, non-administrative worldwide norms association that prepares International Standards for electrical, electronic, and related technologies. The IEC: 62196 is an international standard for plugs, sockets outlet, electrical connectors, and vehicle inlet for EVs. [3]. Moreover, a Combined Charging system (CCS) is developed which have combo 1 and 2 connectors. These connectors are the extension of IEC: 62196 type 1 and 2 connectors with additional DC contacts for high power up to 350 KW DC fast charging. IEC defines the power levels in Modes as summarized in Table IV [14].

Table IV
IEC-62196 SUMMARY

Source	Mode	Voltage (V)	Supply System	Maximum Current (A)
AC	Mode 1	240	1-φ	16
	Mode 2	240	1-φ	32
	Mode 3	415	3-φ	250
DC	Mode 4	600	DC	400

IEEE Standards: The IEEE is world's biggest specialized proficient affiliation devoted to propelling innovation to support humankind. The IEEE Standard 2030.1.1TM-2015 explains fundamental interoperability and characterizes prerequisites for the plans of EVs and DC fast chargers. The IEEE: 2030.1.1 is for fast DC charging of EVs [4]. IEEE: 1547 summarized "Standards for interconnecting distributed resources with electric power systems." It is pertinent for all Distributed Energy Resources (DER) technologies with an aggregate limit of 10 MVA, covers necessities applicable to the performance, operation, testing, safety contemplations and maintenance for DERs interconnection, and emphasizes on the establishment of DERs on primary and secondary networks distribution systems. [15].

CHAdeMO Standards: The CHAdeMO Research and Development began in 2005 determined to build up a public foundation of quick chargers that empowers individuals to drive EVs. The principal business CHAdeMO charging foundation was appointed in 2009 and CHAdeMO was distributed as IEC standard in 2014 (IEC: 62196-3 arrangement AA for connector and IEC: 61851-23 for charging framework). Around the same time, CHAdeMO was published as European Committee for Electrotechnical Standardization, followed by the distribution as IEEE Standard 2030.1.1TM-2015. CHAdeMO is for DC fast charging delivering up to 400 KW by 400 A, 1000 V DC [16].

GB/T: The Standardization Administration of China (SAC) issued Chinese national standards i.e., GB standards. The GB/T: 20234 specifies about the general requirementsfor conductive charging of EVs i.e., the specifications of plugs, socket-outlets, vehicle couplers and vehicle inlets [17].

III. Application of EV Charging Standards

Various nations adheres to various charging standards. The major contrast among these charging principles is of ports/connectors plan. To avoid the charging standards conflicts manufactures are attempting to concoct a typical charging connector contention. In the USA, SAE: J1772 connectors can be used for AC as well as DC charging ability. Tesla has manufactured its own connector that additionally upholds both AC as well as DC quick charging. Tesla additionally has planned a connector for other vehicle models, which changes over SAE: J1772 connectors to empower them by Tesla supercharging stations.

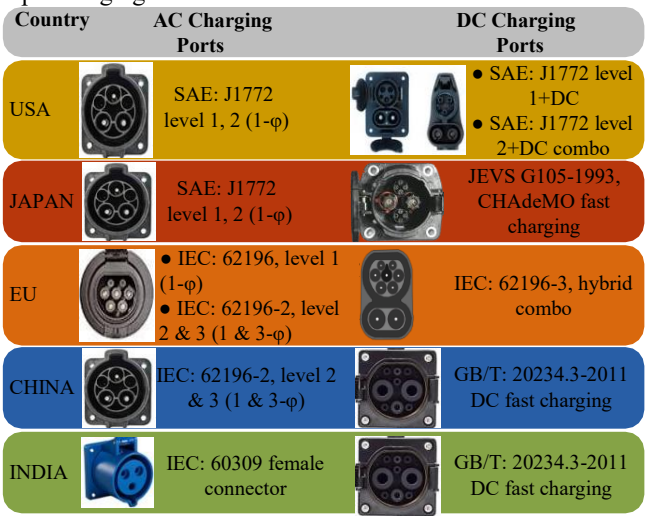


Fig. 2. AC and DC charging ports followed by different countries.

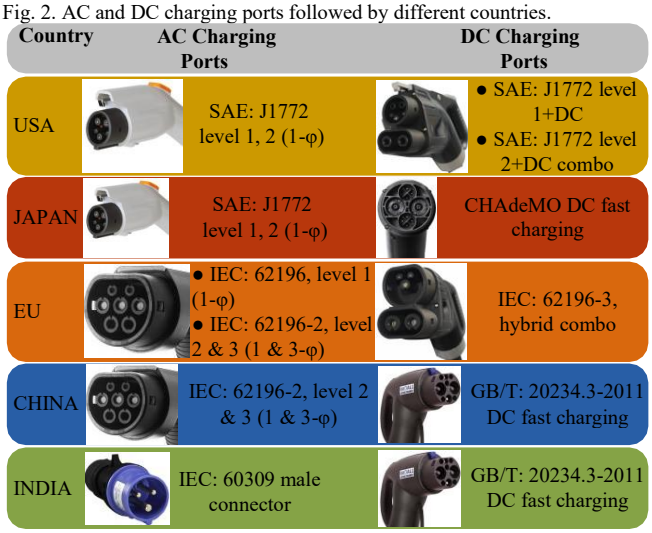


Fig. 3. AC and DC charging connectors followed by different countries.

A "combo," charging connector is used in Europe. This combo has separate DC charging pins added with the existing AC chargers' connectors. EV & charging hardware producers are attempting to homogenize the charging standards and accompany a widespread arrangement of the EV charging gadget. Fig. 2 and 3 shows charging ports and connectors respectively of various standards individually for both AC and DC charging.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:46:42 UTC from IEEE Xplore. Restrictions apply.

IV. Overview of Charging Standards in India

In India, light vehicle category of EV is widely adopted, with more modest battery pack and lower power prerequisites. As of 2017, 2-wheeler (2W) and 3-wheeler (3W) represented 99% of the aggregate EVs in India, with 3-wheeler assessing 78% (17.5 lakhs) and 2-wheeler assessing 21% (4.75 lakh). So, it is necessary to characterize the standards for category of EVs.

The connectors between EV and EVSE, particular of communication conventions and characterizing the level of voltage/current are prerequisites for standardization. India is expecting to deploying a modular charging infrastructure giving all the accessible alternatives like Bharat chargers, CHAdeMO & CCS to EV users. Primary requirement for standardization includes defining the level of voltage/current and finding connectors between the EV and EVSE.

A committee on standardization of protocol for EV formed by the ministry of heavy industries in late 2017. On the basis of voltage and charging time, they framed draft standards for charging stations named as Bharat EV charger-AC001 (BEVC-AC001) and Bharat EV charger-DC001 (BEVC-DC001). In India, bureau of Indian standards (BIS) drafted IS:15886 for standardization of EVs and hybrid electric vehicles (HEVs) and their components.

The automotive research association of India (ARAI) also drafted some of the standards which includes automotive industry standard (AIS)-138. It is for DC charging system by existing international standards which includes IEC: 61851-1 specifies general requirements, IEC: 61851-23 is for charging station of EVs and IEC: 61851-24 is for digital communication [5]. The AIS: 102 (Part 1 and 2) and AIS: 123 include central motor vehicles rules (CMVR) type approval for HEVs. Further, AIS: 131 specifies approval procedure for EVs and HEVs presented in market for demonstration projects for government schemes.

With the type of EV (2W, 3W, 4W and buses), the design of charger depends on battery selection, system architecture and electronics for vehicle. India has adopted BEVC-AC001 (15 A, 3.3 KW, IEC-60309 connector) and BEVC-DC001 (200 A, 15 KW, GB/T-20234 connector) for EVs low voltage applications. Right now, India is in the stage to characterize the standards for various, i.e., low or high, voltage applications as shown in Fig. 4. The EV chargers demand for DC-DC converters and inverters is likely to increase in near future, going forward.

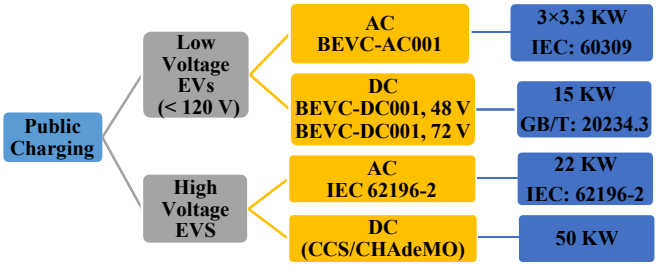


Fig.4. Public charging characterization of EVs in India.

BEVC-AC001: - The detailed specifications of low voltage AC EV charger are shown below. These specifications apply to charge three EVs simultaneously from single phase AC supply at 230 V with maximum 15A output current and maximum 3.3 KW output power with IEC: 60309 connectors. The global reference to this standard is IS: 12360.

General Requirements

- **EVSE:** AC.
- **Energy Transfer Mode**: Conductive mode.

Input Requirements

- **Input Supply System:** AC ,3-φ, 5 wire system (3P+N+PE).
- **Input Voltage**: 415 V as per IS: 12360.
- **Input Frequency:** 50 Hz (± 1.5 Hz).
- **Failure Backup for Input Supply:** For the control system and billing unit minimum 1 hour battery backup for. During back up time in case battery drains out, data logs should be synchronized with CMS (Central management system).

Output Requirements

- **Number of Ports:** 3
- **Each Output type:** 230 V single phase, 15 A as per IS-1-2360 AC.
- **Output Details:** Three independent charging sockets.
- **Output Current:** 15 A current each for three vehicles charging simultaneously, with IEC: 60309 connectors.

BEVC-DC001: - This standard gives all the specifications for low voltage DC EV charging station [6]. This DC charger works with 3-φ AC input voltage, 415 V. In this, the requirements of digital communication between charging station and EV for controlling the charging has also been specified. This standard is referred to the IEC: 61851 Part 1, 23 and GB/T: 20234.

General Requirements

- **EVSE**: Dual- connector DC EVSE.
- **Energy Transfer Mode**: Conductive modes.
- **Charging Mode**: 4 charging modes.
- **Reliability and Serviceability:** Modularity, fault codes, self- diagnostic features, and easy field serviceability

Input Requirements

- **Input Supply System:** AC, 3-φ, 5 wire system (3p+N+E).
- **Input Voltage:** 415 V as per IS:12360.
- **Input Frequency**: 50 Hz (± 1.5 Hz).
- **AC Connector for Supply Side**: IEC: 62196 Type 2.
- **Input Supply Failure Backup:** For control system and billing unit minimum 1 hour battery backup.

Output Requirements

- **Output Details:** Apt for 48 V and 72 V vehicle battery configuration.
- **Charger Configuration Type**:
  - Type 1: 48 V or 72 V for single vehicle charging with maximum power of 10 KW, or charging at 48 V with maximum 3.3 KW of power for 2W EV.
  - Type 2: 48 V for single vehicle charging with maximum power of 10 KW and 72 V with 15 KW power.
- **Output Current:** 200 A Max.
- **Output Connector**: 2 connectors.
- **Compatibility of Output Connector**: One connector with GB/T: 20234.3.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:46:42 UTC from IEEE Xplore. Restrictions apply.

- **Communication System**: Controlled area network (CAN) between EV-EVSE and open charge point protocol (OCPP) between EV-CMS.
- **Converter Efficiency**: At nominal output power (> 92 %)
- **Power Factor**: At full load (≥ 0.90).

V. Safety Consideration of EV Charging Standards in India

With regards to electrical safety, the standardization of charging infrastructure stays a vital challenge in a rigorously developing global regulatory landscape. In India, some safety regulations of international and national standard are picked up by Indian agencies like AISC for featuring the need of harmonization. The AISC is set up by MoRT&H (Ministry of Road Transport & Highways) under CMVR-Technical Standing Committee (TSC). This committee review about the design, operation safety, construction, and motor vehicle's maintenance.

Under National Electric Mobility Mission Plan (NEMMP), adoption and manufacturing of EV and HEV in India scheme is introduced by Department of Heavy Industry, Govt. of India. This will need infrastructure support for AC and DC charging stations [12].

IS: 1271-1985	• Electrical insulation classification and thermal evaluation
IS/IEC: 60309 (Part 1 and Part 2)-2002	• Plugs, socket-outlets and couplers for industrial purposes
IS: 694-1990	• PVC insulated cables for rated voltages up to and including 1100 V
IS: 11000 (Part 2/Sec 2)-2008	• Testing for fire hazard
IS: 2824-2007	• Method for determnation of proof and comparitive tracking idices of insulating materials
IS: 15382 (Part 1)-2003	• Insulation coordination for equipment within low voltage
IS/IEC: 60990-1999	• Touch current and protective conductor current measurement methods.

Fig. 5. Indian electrical safety regulations based on IS: 13252-2010.

Automotive EV application involves both consumer and automotive domains for high voltage electrical safety. For drafting the new standards, some of the existing standards are considered for the electrical safety around on-board as well as off board chargers' infrastructure. Chargers, power connectors the administrative structure concerning electrical and functional safety is grounded by the IS: 13252-2010 standard alongside the two later alterations published by the BIS as shown in Fig. 5 [7].

AIS: 138 (PART2)	• EV conductive AC charging system
AIS: 138 (PART2)/DF/MAY 2017	• EV conductive DC charging system
AIS: 038 (REV.1)-2015	• Electric power train vehicles construction and functional safety requirements
AIS: 049 (REV.1)	• Electric power train vehicles - CMVR type

Fig. 6. Indian standard for EV safety.

Safety standards are submitted to CMVR-TSC in the form of recommendation by AISC for approval. Then AISC recommendations are considered by CMVR-TSC. It may either approve or sends the recommendations for improvements to the AISC. The MoRT&H takes final decision after CMVR-TSC submission of proposal with the recommendations of CMVR. After getting final approval, the test agency will allotment a certificate to developer. Some of the presently appropriate safety regulations of India are shown in Fig. 6.

VI. Conclusion

The most preferrable EV charging standards globally like SAE, IEEE, IEC, CHAdeMO and GB/T are summarized. Further, the EV charging standards followed by India has been discussed in detail. In 2015, faster adoption and manufacturing of HEVs & EVs in India was initiated. The program aimed at promoting research, development, and demonstration of projects on EVs. The Govt. of India, ministry of heavy industries & public enterprises and department of heavy industry constituted a committee and submitted its report on 11th October 2017. The committee has recommendations in the form of specifications for AC and DC chargers namely BEVC-AC001 and BEVC-DC001 for low voltage applications. The specifications of both the chargers are also discussed. The BEVC-AC001 is specified for charging up to 3 KW using standard 15 A industrial plug. However, BEVC-AC002 will sooner be specified for on-board EV charger. In BEVC-DC001, DC chargers with output less than 100 V/15 KW is specified but BEVC-DC002 will sooner be defined for higher voltages ranges (300 V-1000 V) and high-power ratings (30 KW and more). Furthermore, for high voltage applications, currently India is following the international standards like IEC 62196-2 for AC and CCS/CHAdeMO for DC charging. Moreover, the safety standards for EV charging in India also summarized. In India, safety regulations are taken by AISC and safety standards for conductive EV charging are based on IS: 13252-2010.

References

[1] L. Wang, Z. Qin, T. Slangen, P. Bauer and T. van Wijk, "Grid Impact of Electric Vehicle Fast Charging Stations: Trends, Standards, Issues and Mitigation Measures - An Overview," in IEEE Open Journal of Power Electronics, vol. 2, pp. 56-74, 2021.

[2] S. S. Williamson, A. K. Rathore and F. Musavi, "Industrial Electronics for Electric Transportation: Current State-of-the-Art and Future

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:46:42 UTC from IEEE Xplore. Restrictions apply.

Challenges," in IEEE Transactions on Industrial Electronics, vol. 62, no. 5, pp. 3021-3032, May 2015.
[3] D. Sahu and P. S. R. Nayak, "A Substantial Modelling and Analysis of Solar Powered e-Rickshaw Drive System," *2021 International Conference on Sustainable Energy and Future Electric Transportation (SEFET)*, 2021.
[4] Aakash Kumar Seth, Mukhtiar Singh, "Unified adaptive neuro-fuzzy inference system control for OFF board electric vehicle charger," International Journal of Electrical Power & Energy Systems, Vol. 130, 2021.
[5] A. Narula and V. Verma, "Bi – directional trans – Z source boost converter for G2V/V2G applications," *2017 IEEE Transportation Electrification Conference (ITEC-India)*, pp. 1-6, 2017.
[6] M. C. Falvo, D. Sbordone, I. S. Bayram and M. Devetsikiotis, "EV charging stations and modes: International standards," *2014 International Symposium on Power Electronics, Electrical Drives, Automation and Motion*, pp. 1134-1139, 2014.
[7] S. Habib, M. M. Khan, F. Abbas, L. Sang, M. U. Shahid and H. Tang, "A Comprehensive Study of Implemented International Standards, Technical Challenges, Impacts and Prospects for Electric Vehicles," in IEEE Access, vol. 6, pp. 13866-13890, 2018.
[8] Das, H.S. & Rahman, Mohammad & Li, S. & Tan, C.W., "Electric vehicles standards, charging infrastructure, and impact on grid integration: A technological review," Renewable and Sustainable Energy Reviews, Vol. 120, 2020.
[9] A. Sharma and R. Gupta, "Bharat DC001 Charging standard Based EV Fast Charger," *IECON 2020 The 46th Annual Conference of the IEEE Industrial Electronics Society*, pp. 3588-3593, 2020.
[10] Standardisation of protocol for EV charging infrastructure 2017, [online] Available: https://dhi.nic.in/writereaddata/UploadFile/REPORT%20OF%20COMMITTEE636469551875975520.pdf
[11] Propelling Electric vehicles in India 2017, [online] Available: https://beeindia.gov.in/sites/default/files/2019%20-%20EY%20to%20BEE%20-%20Technical%20study%20on%20EVs%20%26%20Charging%20Infrastructure.pdf
[12] A. Gambhir, "Safety considerations for EV charging in India: Overview of global and Indian regulatory landscape with respect to electrical safety," 2017 IEEE Transportation Electrification Conference (ITEC-India), pp. 1-5, 2017.
[13] SAE Electric Vehicle and Plug-in Hybrid Electric Vehicle Conductive Charge Coupler, SAE J1772, Jan. 2010.
[14] International standard IEC 62196 2003, [online] Available: https://webstore.iec.ch/preview/info_iec62196-1%7Bed1.0%7Den.pdf.
[15] "Standard Technical Specifications of a DC Quick and Bi-directional Charger for Use with Electric Vehicles", *IEEE 2030.1.1*.
[16] CHAdeMO protocol development 2021, [online] Available: https://www.chademo.com/activities/protocol-development.
[17] Connection set for conductive charging of electric vehicles 2015, [online] Available: https://www.chinesestandard.net/PDF.aspx/GBT20234.1-2015.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:46:42 UTC from IEEE Xplore. Restrictions apply.

Electrostatic Analysis of Ferroelectric and High Dielectric Layer Assisted MOSFET

Rashi Mann
Microelectronics Research Lab, Department of Engineering Physics
Delhi Technological University
Bawana Road, Delhi, India
rashimann08@gmail.com

Rishu Chaujar
Microelectronics Research Lab, Department of Engineering Physics
Delhi Technological University
Bawana Road, Delhi, India
chaujar.rishu@dtu.ac.in

2021 7th International Conference on Signal Processing and Communication (ICSC) | 978-1-6654-2739-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICSC53193.2021.9673258

***Abstract*— Report presents a hybrid structure of MOSFET with insulator and the ferroelectric assisted substrate is proposed. The structure has two layers of insulator and one layer of ferroelectric material. The applied gate voltage controls the channel conductivity. The comparison investigation is done for conventional MOSFET and FE-MOSFET structures. In comparison, it is observed that the leakage current for the FE-MOSFET structure is reduced significantly, which leads to a better switching ratio (I_{on}/I_{off}) and reduced subthreshold swing (SS) value, which occurs due to short channel. In addition, the drain induced barrier lowering (DIBL) becomes better in the proposed device structure. The minimum voltage required to turn on the device is also reduced in the case of new proposed architecture of MOSFET. As on state current drive current (I_{on}) of MOSFET, it depends upon the channel electron mobility. Thus, the factors like potential, electric field, electron mobility, and energy bandgap are discussed for both structure. These parameters are carried out by using the Cogenda Visual TCAD tool. It has been observed in simulation that the leakage current is reduced by 97.05% with an increase of 3.60% in on-state drive current for FE-MOSFET structure.**

Keywords*—FE-MOSFET, short channel effects, negative capacitance.*

I. INTRODUCTION

With the establishment of semiconductor generation, the advantages of silicon like low cost, small mass, maximum wafer diameter and good carrier ability proves its superiority. Silicon on insulator(SOI) based transistor is advisable for reducing the short channel effects(SCEs) and also the delicate SOI structures encourage the CMOS technology to work on the nano-scale regime [1-6]. After going through the properties of materials that can be used in device structures to enhance their capability. One of them is using ferroelectric (FE) material because of the negative capacitance phenomenon. When FE material is used in the gate stack with a combination of the gate oxide of conventional MOSFET, it shows the phenomenon of negative capacitance which means when the applied voltage increases, the charge goes down and vice versa. This structure is known as NCFET [7].

Another approach to reducing the leakage current of a device structure is using the high-k dielectrics as a gate oxide in place of silicon oxide. Which gives immunity to the device. In the last two to three decades, SOI structures are explored with new design architectures and technologies. Some of which are double-gate, multi-gates, FINFET, NCFET, tunnel FETs etc [8-11]. Using these two approaches, we tried a new structure in which we used both ferroelectric material and high-k dielectric material to enhance the performance of the device structure. Ferroelectric material has the advantage of negative capacitance phenomenon due to which the subthreshold swing value can be reduced beyond the limit value 60 mV/dec. Basically, the SS value defines the voltage that is needed to change the current by one order of magnitude. Which directly reduces the power consumption. Also due to use of high-k dielectric as gate oxide minimize the leakage current in the off state. Another SCEs also shows the improvement with this factor [11-16].

II. DEVICE STRUCTURE AND SIMULATION METHOD

The conventional MOSFET structures are fabricated using a planer CMOS mechanism on bonded SOI wafers. For less complexity, we divide the body region into different sections like insulator region, FE region and substrate region. Source and drain regions have been doped with n-type Gaussian doping and substrate region are uniformly doped with p-type. The doping concentration of the source/drain and substrate region is 1×10^{20} and 1×10^{17}, respectively. Figure 1 shows the proposed device structure. The 25 nm gate length is used for the device fabrication and a 1 nm combination of gate oxide (0.5 nm SiO_2 +0.5 nm HfO_2) is used in the gate stack. The source and drain are extended with 10 nm length on both sides of the channel region. The gate electrode of Npoly silicon is used with the work function of 4.5 eV. The ferroelectric layer of hafnium oxide based FE (HfO_2FE) is sandwiched in-between the insulator layer of HfO_2. Landau's parameters for the FE material are α= - 1e+07 and β=8.9e+08 [17]. For simulation work, we used the meshed device with the mesh size given in Table 1. The visual TCAD simulator is used for the analysis. The simulation work is done at room temperature (300 K) and the constant drain-source voltage of 0.5 V. The simulated results have been calibrated with the published work of G. Roy et al. The calibrated results of 35 nm MOSFET are presented in Figure 2 with constant V_{ds} of 50 mV and 850 mV [18].

III. RESULTS AND DISCUSSION

Figure 3 shows the variation of drain current (I_d) for both structure MOSFET with supplied gate voltage at constant drain-source voltage. And, from the Figure 3, it is clear that the leakage current is reduced almost by one decimal point which gives strength to the device performance [19-21]. Also, the on-current for the proposed device structure increases slightly which gives a better switching ratio to the device. Figure 4 shows the comparison of switching ratio for conventional and proposed device structures. The maximum value of switching ratio 0.581×10^6 is observed for the proposed device structure. In the same Figure 4, the other factor subthreshold swing that describes the immunity of structure is shown. Smaller the SS value, the better is the device performance because less voltage is needed to change the current by one decimal point [22-25]. So, a better SS is

978-1-6654-2739-5/21/$31.00 ©2021 IEEE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:38:44 UTC from IEEE Xplore. Restrictions apply.

observed for the FE-MOSFET structure when compared to the conventional one.

Figure 5 shows the comparison of DIBL and threshold voltage for both device structures. Drain induced barrier lowering is another name for channel length modulation and is the result of short channels. For better device performance, this should be minimized. From the results, we observed that the DIBL is reduced significantly in the case of the proposed device structure. DIBL is calculated by taking the horizontal shift in the sub-threshold characteristics (in millivolts) divided by change in the V_{ds}.

Another factor is the threshold voltage, the minimum voltage that is required to turn on the device. Smaller the V_{th} value, the lesser the power consumption. When compared both devices, it is observed that FE-MOSFET shows the less threshold voltage value. All these parameters with the width of device (W_g) are tabulated in table 2.

Further, Figure 6 shows the surface potential variation along the channel for both the device structure. The formulation of surface potential is given as

$$\Phi_m = \frac{kT}{q} \ln \left[\frac{n}{n_i} \right.$$

Where q signifies the elementary charge, n_i represents the intrinsic carrier concentration and k and T shows the Boltzmann's constant and temperature respectively. So, the concentration increases in the case of FE-MOSFET structure due to the introduction of ferroelectric material, hence the potential increases in the channel region of FE-MOSFET than the conventional MOSFET structure. The electric field variation along the channel of both the device is given in the Figure 7. From the Figure 7, we can see that the electric field at the drain end of the FE-MOSFET structure is less than at the source end, which shows the reduction of corner effects. The energy band diagram of FE-MOSFET and conventional MOSFET gives the details about the working of device structure that is shown in the Figure 8. This energy band diagram for both structures is obtained across the gate length at V_{ds}= 0.5 V and V_{gs} is varied from 0 to 1 V. From the representation, we can see that the conduction band and valence band of FE-MOSFET shift towards lower values of energy when compared to the conventional MOSFET one. Furthermore, the electron mobility of both device structures is shown in Figure 9. This variation is taken along the channel of both devices. The electron mobility is observed higher in FE-MOSFET in comparison to conventional MOSFET which directly gives the higher on current for the device. Electron mobility is an important factor for device performance. The net doping of FE-MOSFET is shown in Figure 10, which shows the maximum doping around the source and drain region and is of n-type. The transient analysis for FE-MOSFET based inverter shows the immunity to noise margin and transition region.

IV. Conclusion

The steady-state analysis of insulator and ferroelectric assisted MOSFET at 25 nm gate length has been done by using Cogenda visual TCAD tool. It is concluded from the simulation that the parameters shows better immunity for the FE-MOSFET when compared to the conventional MOSFET device. The on-current is improved by 3.60% and the leakage current is reduced by 97.05%. The SS slope is reduced by 16.75% and DIBL is improved by a 56.51 % reduction. Also, the threshold voltage is decreased by 20.37%. The potential is increased by 5.10 % in the channel region for the FE-MOSFET device structure. The electron mobility and electric field also show improved results with a 16.21 % increment in channel region and 19.66 % decrement at the drain end respectively. In the channel region. The energy band shifts towards the lower energy value. From the above result and discussion, we conclude that the modified configuration of MOSFET with ferroelectric and insulator layers can be an alternate choice for designing analogue circuit devices.

TABLE I. Default device dimension and parameters are taken for the simulation.

Parameter	Conventional MOSFET	Modified FE-MOSFET
L_g (nm)	25	25
$L_{S/D}$ (nm)	10	10
T_{ox} (nm)	1	1(0.5 SiO_2+ 0.5 HfO_2)
T_{FE} (nm)	0	8
T_{ins} (nm)	0	9
$N_{S/D}$ (cm^{-3})	1×10^{20}	1×10^{20}
N_{ch} (cm^{-3})	1×10^{17}	1×10^{17}
Φ_m (eV)	4.5 (Npoly Si)	4.5 (Npoly Si)
Mesh size (nm)	0.005	0.005

TABLE II. SS and DIBL comparison for conventional MOSFET and FE-MOSFET.

Device	W_g [um]	L_g [um]	V_{th} [V]	DIBL [mV/V]	SS [mV/dec]
MOSFET	0.05	0.025	0.341	0.086	87.843
FE-MOSFET	0.05	0.025	0.272	0.037	73.131

Fig. 1. Schematic representation of proposed FE-MOSFET architecture.

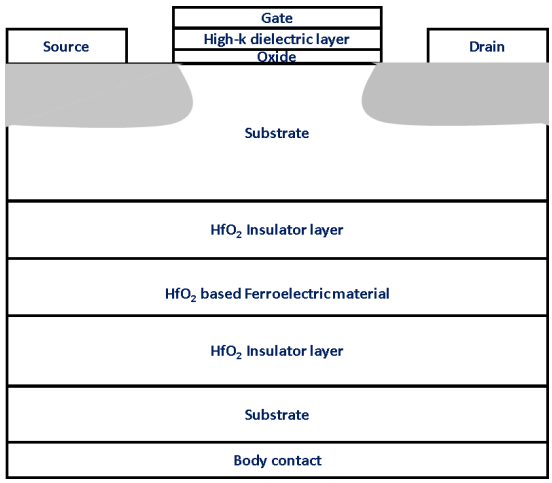


Fig. 2. . Calibration of simulated results with experimental work of 35 nm MOSFET device structure.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:38:44 UTC from IEEE Xplore. Restrictions apply.

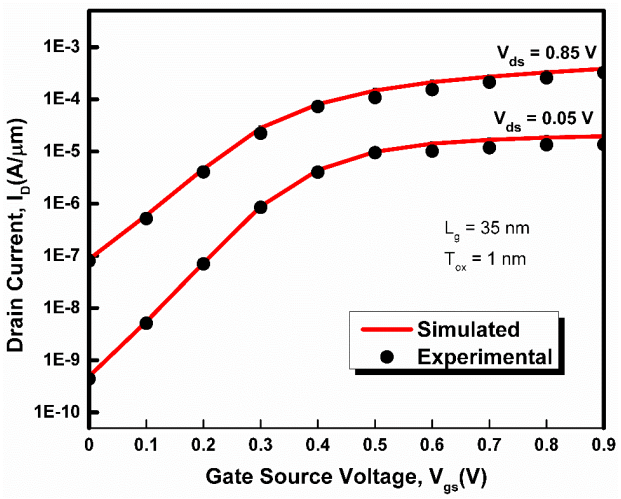


Fig. 3. The transfer (I_D-V_G) characteristic of conventional MOSFET and FE-MOSFET devices at V_{ds}= 0.5 V.

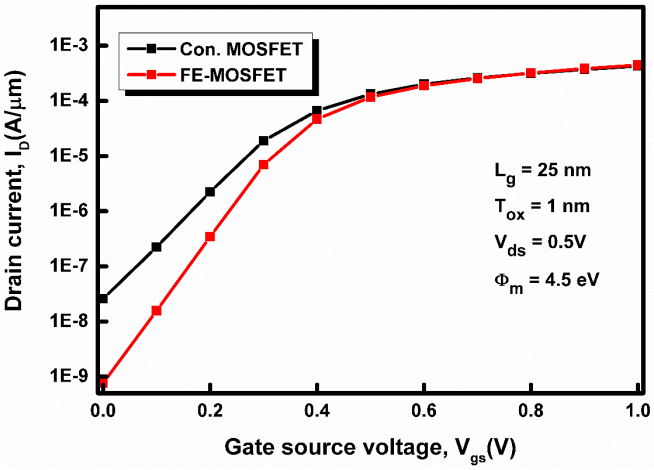


Fig. 4. The switching ratio and subthreshold-swing value for conventional MOSFET and FE-MOSFET.

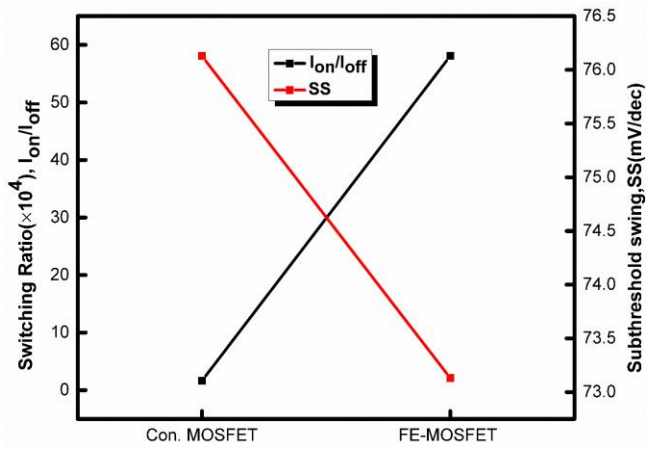


Fig. 5. The threshold voltage and DIBL comparison for both conventional MOSFET and FE-MOSFET.

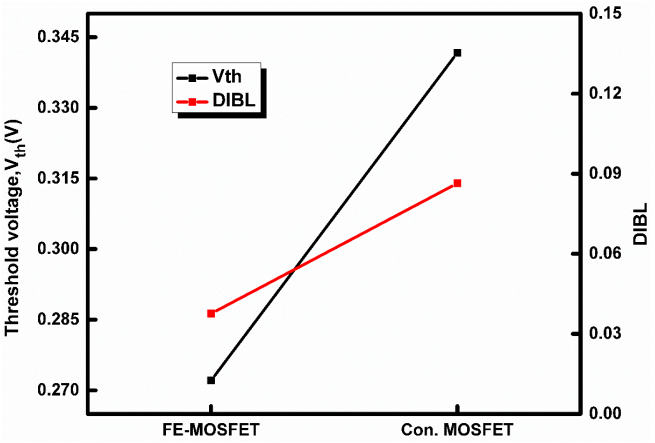


Fig. 6. The potential variation across the channel at Vds=0.5V for conventional MOSFET and FE-MOSFET devices.

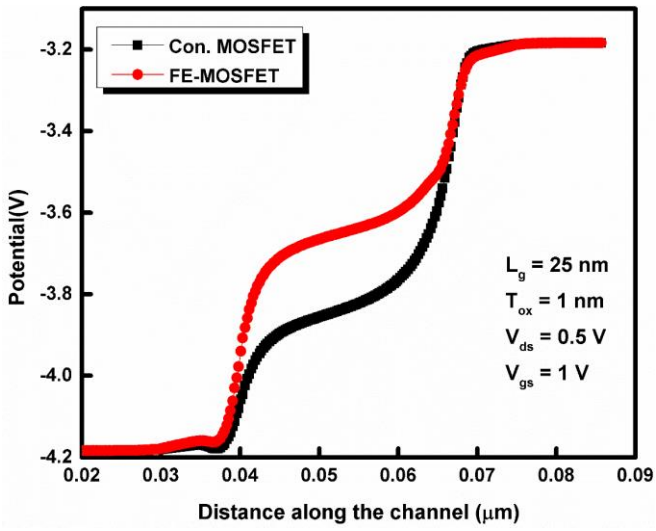


Fig. 7. The electric field across the channel at V_{ds}= 0.5 V for conventional MOSFET and FE-MOSFET devices.

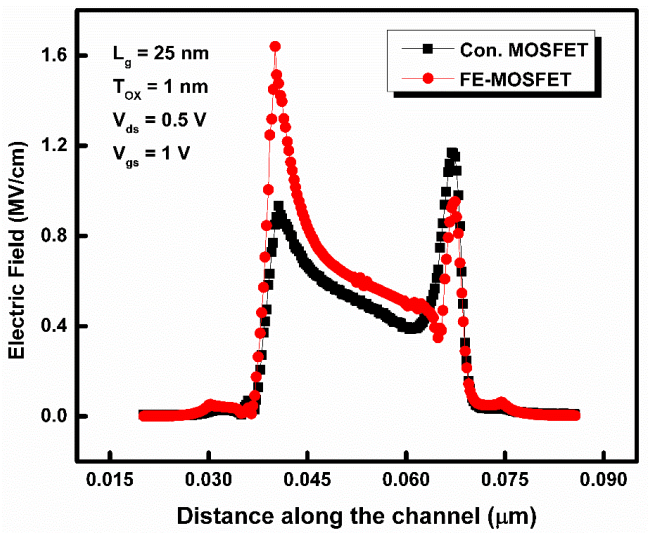


Fig. 8. The energy band diagram across the length of both devices at V_{ds}= 0.5 V.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:38:44 UTC from IEEE Xplore. Restrictions apply.

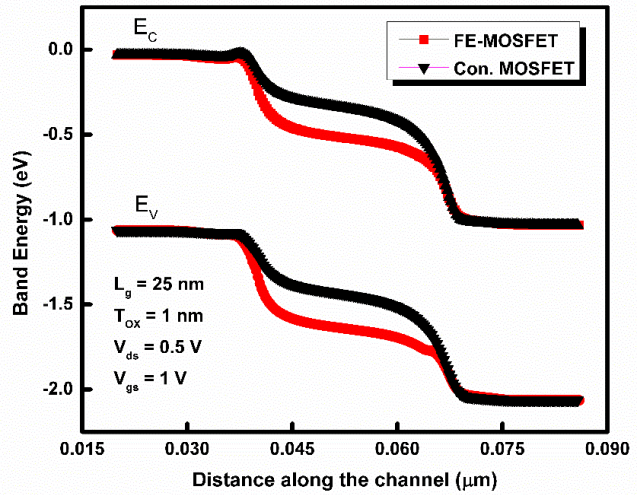


Fig. 9. The electron mobility across the channel at V_{ds} = 0.5 V for conventional MOSFET and FE-MOSFET devices.

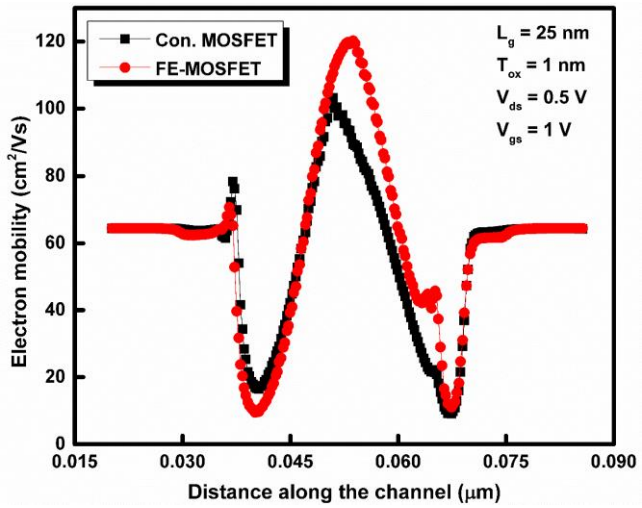
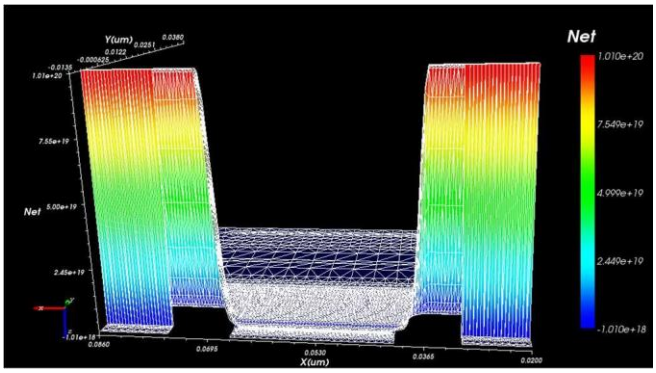


Fig. 10. Net doping of FE-MOSFET across the length of the device.



REFERENCES

[1] G. E. Moore, Cramming more components onto integrated circuits, Proc.IEEE 2002, 86, 82.

[2] P. K. Bondyopadhyay, Moore's law governs the silicon revolution, Proc.IEEE 1998, 86, 78.

[3] C. A. Mack, Fifty years of Moore's law, IEEE Trans. Semicond. Manuf. 2011, 24, 202.

[4] G. E. Moore, Progress in digital integrated electronics, Tech. Dig. - Int.Electron Devices Meet. 1975, 11.

[5] M. M. Waldrop, The chips are down for Moore's law, Nature, 2016, 530,144.

[6] E. Mollick, Establishing Moore's Law, IEEE Ann. Hist. Comput. 2006, 28, 62.

[7] R. Aitken, V. Chandra, J. Myers, B. Sandhu, L. Shifren, G. Yeric, On device design for steep-slope negative-capacitance field-effect-transistor operating at sub-0.2V supply voltage with ferroelectric HfO2 thin film, Symp.VLSI Technol., Dig. Tech. Pap., 2014, 1.

[8] L. B. Kish, End of Moore's law: thermal (noise) death of integration in micro and nano electronics, Phys. Lett. A, 2002, 305, 144.

[9] G. E. Moore, Why optical lithography will live forever, Proc. SPIE, 1995, 2439, 2.

[10] J. R. Powell, The economic limit to Moore's Law, Proc. IEEE, 2008, 96, 1247.

[11] R. R. Schaller, Moore's law: past, present and future, IEEE Spectrum, 1997, 34, 52.

[12] S. E. Thompson, S. Parthasarathy, Moore's law: the future of Si microelectronics, Mater. Today, 2006, 9, 20.

[13] K. J. Kuhn, In 13th int. workshop computational electronics, IEEE, Piscataway, NJ, 2009, p. 1.

[14] K. Boucart, A. M. Ionescu, Double-gate tunnel FET with high-gate dielectric, IEEE Trans. Electron Devices, 2007, 54, 1725.

[15] Y. Khatami, K. Banerjee, Steep subthreshold slope n-and p-type tunnel-FET devices for low-power and energy-efficient digital circuits, IEEE Trans.Electron Devices, 2009, 56, 2752.

[16] G. Fiori, G. Iannaccone, Ultralow-voltage bilayer graphene tunnel FET, IEEE, Electron Device Lett., 2009, 30, 1096.

[17] S. Salahuddin, S. Datta, Use of Negative Capacitance to Provide Voltage Amplification for Low Power Nanoscale Devices, Nano Lett., 2008, 8, 405-410.

[18] G. Roy, F. A. Lema, A. R. Brown, S. Roy, A. Asenov, Simulation of combined sources of intrinsic parameter fluctuations in a 'real' 35 nm MOSFET, Proc. ESSDERC, 2005, 337–340.

[19] J. D. Bude, Drain disturb during CHISEL programming of NOR flash EEPROMs-physical mechanisms and impact of technological parameters, Symp. VLSI Technol., Dig. Tech. Pap., 1995, 101.

[20] A. Padilla, C. W. Yeung, C. Shin, C. Hu, T. J. K. Liu, Feedback FET: A novel transistor exhibiting steep switching behavior at low bias voltages, Tech. Dig. - Int. Electron. Devices Meet., 2008, 1.

[21] D. E. Fulkerson, Feedback FET logic: a robust, high-speed low-power GaAs logic family, IEEE J. Solid-State Circuits, 1991, 26, 70.

[22] A. Kumar, M.M.Tripathi, R. Chaujar, Comprehensive analysis of sub-20 nm black phosphorus based junctionless-recessed channel MOSFET for analog/RF applications, SUPERLATTICE MICROST, 2018,116, 171-180.

[23] N. Gupta, R. Chaujar, Investigation of temperature variations on analog/RF and linearity performance of stacked gate GEWE-SiNW MOSFET for improved device reliability, Microelectronics Reliability, 2016, 64, 235–241.

[24] R. Chaujar, R. Kaur, M. Saxena, M. Gupta, R.S. Gupta, Laterally amalgamated DUal Material GAte Concave (L-DUMGAC) MOSFET for ULSI, Microelectronic Engineering, 2008, 85, 566–576.

[25] J. Madan, R. S. Gupta, R. Chaujar, Mathematical modeling insight of hetero gate dielectric-dual material gate-GAA-tunnel FET for VLSI/analog applications, Microsystem Technologies, 2017, 23, 409-4098.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:38:44 UTC from IEEE Xplore. Restrictions apply.

Emotion Recognition Using Rational Dilation Wavelet Transform For Speech Signal

Ravi
Delhi Technological University (DTU),
Shahbad Daulatpur,
New Delhi,110042
ravi_2k20phdec12@dtu.ac.in

Sachin Taran
Delhi Technological University (DTU),
Shahbad Daulatpur,
New Delhi,110042
sachintaran@dtu.ac.in

2021 7th International Conference on Signal Processing and Communication (ICSC) | 978-1-6654-2739-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICSC53193.2021.9673412

Abstract*—Speech based emotion recognition is an emerging area in human computer interaction. It can improve the performance of smart assistant systems. In this work, a wavelet based approach is proposed for the classification of emotions. In proposed method, rational dilation wavelet transform, which has adjustable quality factor and high frequency resolution, is employed. Wavelet based features are extracted and statistically examined by using Kruskal-Wallis test. Further, these features are utilized for training and testing of K-nearest neighbors, ensemble, and decision tree classifiers variants. Among, all the classifiers optimizable ensemble classifier achieves the highest accuracy.

Keywords- Emotion recognition, Speech signal, Rational dilation wavelet transform, machine learning.

I. INTRODUCTION

Speech is one of the most common method for information exchange among the human beings. So, it attracts the attention of researchers in the domain of human computer interaction (HCI) [1]. There are number of speech signal based HCI applications like human emotion recognition, audio mining, vehicle card system, call center application and automatic translation system [2]. It is found that emotion recognition can improve the HCI system performance. Emotion reflects the human brain activity so that it contains a set of behavioral and psychological information of the speaker [3]. As human speech has biomedical features, so the same utterances spoken by different speakers have noticeable dissimilarities. Further, the voice recorded at a different time from the same speaker also has dissimilarity [4]. A reliable emotion recognition HCI system can be design by using the speech signal.

The major difficulty lies in the field of emotion recognition using speech signals is that to select the optimum features.

For speech emotion recognition (SER), there are limited features identified in comparison to speech recognition. The other challenges in the designing of SER system are listed as follows: the number of emotion classes is different in different datasets, different languages have different optimal feature sets, optimal speech feature sets also vary with the variation of speech length, gender, vocal tract shape and age of the speaker [5].

In recent research, there are multiple machine learning and deep learning-based techniques are suggested for different SER classes [6-13]. In [6], happy and sad emotions of Ryerson Audio-Visual (RAVDESS) dataset are used for the training of convolution neural network (CNN) and 66.41% classification accuracy is achieved [6]. In another method bi-directional long short-term memory is used for the classification of same emotion classes happy and sad of RAVADESS dataset and achieved 70.4% recognition accuracy [7]. In [8], author used simple task (ST), multitask feature selection/learning (MTFS/MTFL) and group MTFS/MTFL model for the classification of four classes neutral, angry, happy and sad of RAVDESS and university of Michigan databases. The highest accuracy for four classes classification is 57.14%. While, In [9], for the same four classes neutral, angry, happy and sad of interactive emotional dyadic motion capture database, down sampling and deep stride CNN model used and 81.7% accuracy was achieved. In [10], SVM classifier along with bag of words (BOW) and speeded-up robust features are utilized for five common emotion classes (sad, anger, happy, neutral and fear) classification for different datasets. The maximum and minimum recognition accuracy 83% and 43% are achieved, respectively [10]. Six different emotions neutral, disgust, anger, fear, happiness and sadness of Berline database were classified in [11] and [12]. In [11], mel frequency cepstral coefficient (MFCC), pitch and energy features along with *K*-nearest neighbors (KNN) and Gaussion mixture model

978-1-6654-2739-5/21/$31.00 ©2021 IEEE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:38:15 UTC from IEEE Xplore. Restrictions apply.

classifier are used for the classification. While in [12], MFCC features with SVM, naive bayes (NB) and KNN classifier are used. In both cases maximum accuracy 87.7% was achieved. Deep belief network was introduced for the classification of six emotion classes surprise, neutral, anger, happiness, fear and sadness of CASIA Chinese speech emotion dataset and its accuracy is analyzed and compared with the back propagation (BP) and SVM classifier [13]. In this comparison, BPN and BP achieve average 92.5% and 90% recognition rate, respectively [13].

This work proposes an emotion recognition framework using speech signal. In the proposed work speech signal is preprocessed and decomposed into sub-bands (SBs) by rational dilation wavelet transform (RDWT). Features are extracted from RDWT provided SBs for the classification of happy and sad emotions. These features are used for the training and testing of multiple classifiers and best classifier is explored for the emotion recognition.

II. METHODOLOGY

The steps involved in the proposed method are shown in the Fig. 1. The workflow includes dataset, preprocessing, decomposition, feature extraction and classification.

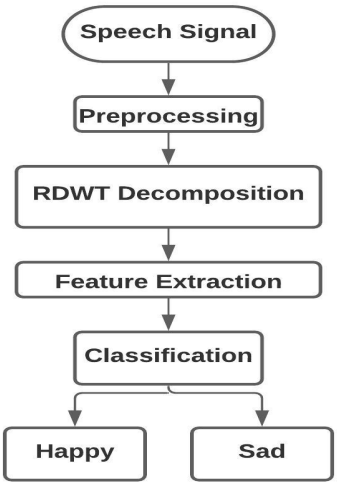


Fig. 1 Classification of emotions using RDWT based features.

2.1 Dataset

The proposed model uses a publicly available RAVDESS dataset [14]. This dataset is available with eight emotion classes surprise, neutral, anger, happiness, fear, calm and sadness. This dataset was recorded by 24 professional actors; 12 male and 12 female and the recording sampling frequency is 48KHz. In this database, 192 speech signals are available for each class except the neutral class. This work uses happy and sad emotions speech signals to form a binary classification problem. The example of raw speech signal is shown in Fig. 2.

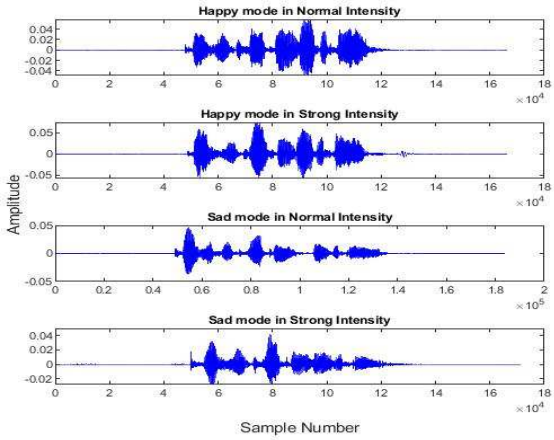


Fig. 2. Example of raw speech signal.

2.2 RATIONAL DILATION WAVELET TRANSFORMS (RDWT)

In literature, many variants of wavelet transform (WT) like dual tree complex WT and double density WT are employed in signal processing. The performance of these wavelet transforms is limited by frequency resolution and low quality (Q) factor [15]. While, for the optimum representation of the oscillatory signals, a high resolution and adjustable Q factor based wavelet transform are required.

Due to the oscillatory behavior of speech signal RDWT are selected for the decomposition [15]. For designing low-pass and high-pass filters of RDWT, Fast Fourier Transform based circular convolution is used. In RDWT, the perfect reconstruction is obtained for discrete signal, when the least common multiple of p and q is multiplied with the length of signal at each level. The RDWT is approximately shift invariant and discrete transform. It provides an adjustable Q factor and high frequency resolution. For performing the decomposition RDWT uses the repeated structure of two filter banks. Further, desired Q factor is obtained by adjusting positive integer variables p, q and s. These variables must satisfy $p/q + 1/s \geq 1$ and $1 \leq p < q$ equations, where p and q are co-prime and assuring $q>p$. In most of the WT redundancy factor is fixed while in RDWT redundancy factor is controllable. The redundancy in RDWT with iterated filter-bank defined as [15],

$$\text{red(p,q,s)} = \lim_{j\to\infty}(p,q,s) = \frac{1}{s}\left(\frac{1}{1-p/q}\right)$$

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:38:15 UTC from IEEE Xplore. Restrictions apply.

In RDWT, the expression for scaling φ(t) and wavelet function Ψ(t) is defined as [15],

$$\varphi(t) = \sqrt{\frac{q}{p}} \sum_{n} h(n) \varphi\left(\frac{q}{p} t - n\right)$$

$$\Psi(t) = \sqrt{\frac{q}{p}} \sum_{n} g(n) \Psi\left(\frac{q}{p} t - n\right)$$

where, *h(n)* and *g(n)* represent high pass and low pass filters respectively. The mathematical value of p and q are optimized to achieve higher frequency resolution.

In this work, the parameters chosen for the tuning of RDWT are *p=1, q =2, s=1* and *j=7*. With these parameters, the preprocessed speech signal is decomposed into 8 SBs, which are further used for feature extraction.

2.3 FEATURE EXTRACTION

In this work four basic features complexity, average amplitude change (AAC), mobility and zero crossing rate (ZCR) [16-19] are extracted from eight SBs of each frame. AAC is calculated by taking the mean of the difference between two successive samples. Mobility is measured by taking the ratio of the first derivative of variance of signal to the variance of the signal. ZCR is defined as the rate of the crossing of the zero axis by the speech signal. SBs are obtained from the RDWT decomposition of a frame of speech signal. As four features are extracted from each SB, hence there is a total of 32 effective features for a frame of speech signal. Further, these extracted features are utilized for the classification of emotions.

2.4 CLASSIFICATION MODELS

In this work, for the classification of emotions K-nearest neighbors (KNN), ensemble and decision tree (DT) classifier variants are tested [16-18]. The selection of these classifiers is based on their optimum behavior with the selected features. The ensemble method is also used to overcome the shortcoming of used classifiers. The KNN variants are Weighted KNN (WKNN), Cosine KNN (CosKNN), coarse KNN (CKNN), fine KNN (FKNN) and Medium KNN (MKNN) tested for selected feature set. The DT variants are Fine tree (FT), Medium tree (MT) and Coarse tree (CoT) tested for selected feature set. The ensemble variants are boosted tree (BoT), Bagged tree (BaT), optimizable ensemble (OE), RUSboosted tree (RUSBoT) and Subspace discriminant (SubD) tested for selected feature set.

III. RESULTS AND DISCUSSION

In the proposed model, two emotion classes happy and sad of RAVDESS dataset are used for the training and testing of classifiers. In this model, each speech signal is preprocessed and fragmented into 25ms duration frames. The preprocessing includes removing unvoiced part, apply hamming window (HW) and pre–emphasis [20]. The unvoiced part is removed by using optimize threshold value and HW is used for smoothing the edges of each frame. The example of a preprocessed signal is shown in Fig. 3. Further, these frames are decomposed into 8 SBs by using RDWT. After decomposition four features complexity, AAC, mobility and ZCR are extracted from each SB, consequently, a total of 32 features are extracted from each frame. These features are statistically examined using the probabilistic (p) - value of Kruskal-Wallis (KW) test. The p–value of KW test for all features is shown in Table I. Most of the features are showing a significantly low p-value for each SB. Further, these features are tested on different classifiers. For the training of classifiers, 10 fold cross validation is used.

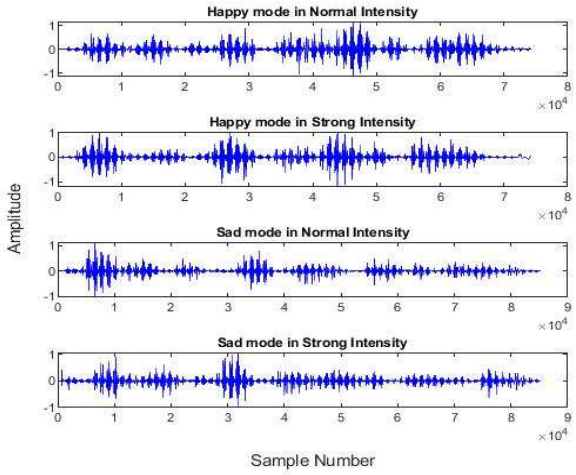


Fig. 3 Example of preprocessed speech signal.

TABLE I. KW test p-values of SB-wise extracted features.

Feature /SB	SB-1	SB-2	SB-3	SB-4	SB-5	SB-6	SB-7	SB-8
Complexity	1.0429 6e-120	8.1529 3e-22	0.0067	0.9571	3.7768 1e-220	5.6021 5e-08	3.8378 4e-29	2.4260 2e-98
AAC	1.2388 1e-18	2.9125 3e-31	2.2767 7e-35	4.3915 7e-18	0.683	0.1769	3.0563 6e-06	1.1663 2e-13
Mobility	3.2437 7e-12	1.1847 9e-36	0.0148	8.0950 6e-68	1.1677 e-117	0.3596	4.7471 6e-13	1.7908 9e-199
ZCR	1.6545 4e-62	3.3511 1e-43	0.1999	1.1069 9e-82	4.4060 2e-117	0.371	9.1758 1e-19	1.6499 6e-191

By observing Table II, for speech emotion classification KNN and ensemble classifier variants exhibit better performance in comparison to DT classifier variants. In KNN variants, FKNN achieves the highest 79.9% and

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:38:15 UTC from IEEE Xplore. Restrictions apply.

CKNN lowest 72.2% classification accuracy. Similarly, in ensemble variants, BoT achieves the lowest 70.2% and OE highest 83.3% classification accuracy. The classification accuracy of other variants such as FT, MT, CoT, MKNN, CosKNN, WKNN, BaT, SubD and RUSBoT is also shown in Table 2. Among all the classifiers OE classifier achieves the highest accuracy 83.3% for the selected classes (happy and sad) and dataset (RAVDESS). Fig. 4 represents the confusion matrix of the OE classifier. The truly classified signal in one class is 15710 and in the other class is 13967.

The performance of the OE classifier is further evaluated through the ROC curve in Fig. 5. In this figure, it can be observed that the area under curve (AUC=0.92) is very high. The high value of AUC represents the stable classification of happy and sad classes signals.

TABLE II. Proposed methods classification accuracies with DT, KNN and Ensemble classifiers variants.

DT		KNN		Ensemble	
DT variant	**Accuracy (%)**	**KNN variant**	**Accuracy (%)**	**Ensemble variant**	**Accuracy (%)**
FT	67.4	FKNN	79.9	BoT	70.2
MT	63.3	MKNN	76.9	BaT	78.9
CoT	60	CKNN	72.2	SubD	52.6
		CosKNN	76.9	RUSBoT	65
		WKNN	79.1	OE	**83.3**

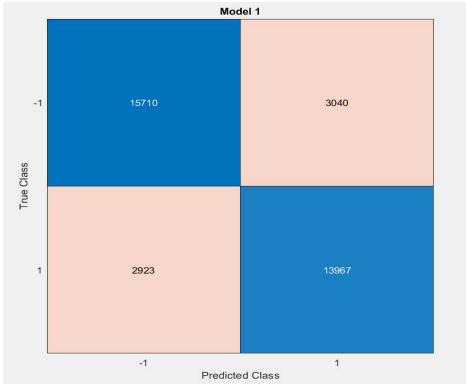


Fig. 4 Confusion Matrix of OE classifier.

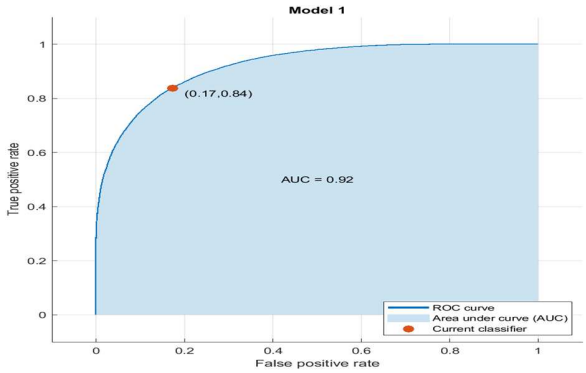


Fig. 5 ROC curve of OE classifier.

RAVDESS dataset recorded in high and low intensity and also includes both male and female speakers. Hence there are very few methods available, which include all variations. In Table III, the performance of the pre-established method on the same RAVDESS dataset is compared with the proposed method [6-7]. The classification accuracies achieved in previous methods are 66.41% by R. Jannat et al [6] and 70.14% by M. A. Jalal et al [7], while the proposed method achieves 83.3% classification accuracy. So, this work has proposed more accurate method for emotion recognition using speech signals.

TABLE III. Performance report of same dataset methods.

Authors	Classified emotions	Classifier	Accuracy (%)
R. Jannat et al [6]	Happy, Sad	CNN	66.41
M. A. Jalal et al [7]	Happy, Sad	BLSTM	70.14
Proposed Method	Happy, Sad	OE	83.3

IV. Conclusion

This work is introduced RDWT-based features for speech signals of different emotions. The extracted features are tested on multiple machine learning algorithms for emotion recognition. First, RDWT is used for the decomposition of preprocessed speech signals. It decomposes each frame into 8-SBs. Further, four features complexity, AAC, mobility and ZCR are extracted from each SB and their statistical significance is examined by using the *p*-value of KW test. These features are utilized for the training and testing of multiple classifiers variants. It is observed that KNN and ensemble variants exhibit better performance in comparison to DT variants. Among all the classifiers OE classifier exhibits the highest accuracy 83.3%, which is also highest as compared to existing works.

References

[1] Y. Sönmez and A. Varol, ``In-depth analysis of speech production, auditory system, emotion theories and emotion recognition," in Proc. 8th ISDFS, Beirut, Lebanon, 2020, pp. 18.

[2] A. Mencattini, E. Martinelli, G. Costantini, M. Todisco, B. Basile, M. Bozzali, and C. Di Natale, ``Speech emotion recognition using amplitude modulation parameters and a combined feature selection procedure," Knowl.-Based Syst., vol. 63, pp. 6881, Jun. 2014.

[3] M. Egger, M. Ley, and S. Hanke, ``Emotion recognition from physiological signal analysis: A review," *Electron. Notes Theor. Comput. Sci.*, vol. 343, pp. 35_55, May 2019.

[4] E. S. Paula, S. Sherb, M. Tamiettoc, P. Winkielmane, and M. T. Mendla, "Towards a comparative science of emotion: Affect and consciousness in humans and animals," ScienceDirect, Neuroscience and Biobehavioral Reviews, Vol.108, pp.749-770, 2020.

[5] J. Zhang, Z. Yin, P. Chen, and S. Nichele, ``Emotion recognition using multi-modal data and machine learning techniques: A tutorial and review," Inf. Fusion, vol. 59, pp. 103126, Jul. 2020.

[6] R. Jannat, I. Tynes, L. L. Lime, J. Adorno, and S. Canavan, ``Ubiquitous emotion recognition using audio and video data," in Proc. ACM Int. Joint Conf. Int. Symp. Pervas. Ubiquitous Comput. Wearable Comput.,Oct. 2018, pp. 956-959.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:38:15 UTC from IEEE Xplore. Restrictions apply.

[7] M. A. Jalal, E. Loweimi, R. K. Moore, and T. Hain, ``Learning temporal clusters using capsule routing for speech emotion recognition," in Proc. Interspeech, Graz, Austria, Sep. 2019, pp. 17011705, doi: 10.21437/Interspeech.2019-3068.
[8] S. R. Livingstone and F. 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, vol. 13, no. 5, May 2018, Art. no. e0196391, doi: 10.1371/journal. pone.0196391.
[9] Mustaqeem and S. Kwon, ``A CNN-assisted enhanced audio signal processing for speech emotion recognition," Sensors, vol. 20, no. 1, p. 183, Dec. 2019, doi: 10.3390/s20010183.
[10] E. Spyrou, R. Nikopoulou, I. Vernikos, and P. Mylonas, ``Emotion recognition from speech using the bag-of-visual words on audio segment spectrograms," Technologies, vol. 7, no. 1, p. 20, Feb. 2019, doi: 10.3390/technologies7010020.
[11] C. Prakash,V. Gaikwad, R. R. Singh, and O. Prakash, ``Analysis of emotion recognition system through speech signal using KNN & GMMclassier," IOSR J. Electron. Commun. Eng., vol. 10, no. 2, pp. 5561, 2015.
[12] S. Emerich, E. Lupu, and A. Apatean, ``Emotions recognition by speech and facial expressions analysis," in Proc. 17th Eur. Signal Process. Conf., Glasgow, Scotland, Aug. 2009, pp. 16171621.
[13] B. Chen, Q. Yin, and P. Guo, ``A study of deep belief network based Chinese speech emotion recognition," in Proc. 10th Int. Conf. Comput. Intell. Secur., Kunming, China, Nov. 2014, pp. 180184, doi:10.1109/CIS.2014.148.
[14] S. R. Livingstone and F. 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, vol. 13, no. 5, May 2018, Art. no. e0196391, doi: 10.1371/journal. pone.0196391.
[15] I. Bayram and I. W. Selesnick, "Frequency-Domain Design of Overcomplete Rational-Dilation Wavelet Transforms," in IEEE Transactions on Signal Processing, vol. 57, no. 8, pp. 2957-2972, Aug. 2009, doi: 10.1109/TSP.2009.2020756.
[16] Sachin Taran, Smith K. Khare, Varun Bajaj, and G. R. Sinha. " Classification of Motor-Imagery Tasks from EEG Signals using Rational Dilation Wavelet Transform." Modelling and Analysis of Active Biopotential Signals in Healthcare-Volume 2, 2020.
[17] Krishna A H, Sri A B, Priyanka K Y V S, Taran S and Bajaj V 2018 Emotion classification
using EEG signals based on tunable-Q wavelet transform IET Sci. Meas. Technol. 13 375–80.
[18] Chada S, Taran S and Bajaj V 2020 An efficient approach for physical actions classification using surface EMG signals Health Inf. Sci. Syst. 8 3.
[19] Sravani C, Bajaj V, Taran S and Sengur A 2020 Flexible analytic wavelet transform based features for physical action identification using sEMG signals IRBM 41 18–22.
[20] Mehmet Berkehan Akçay, Kaya Oğuz, Speech emotion recognition: Emotional models, databases, features, preprocessing methods, supporting modalities, and classifiers,Speech Communication, Volume 116, 2020, Pages 56-76, ISSN 0167-6393.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:38:15 UTC from IEEE Xplore. Restrictions apply.

Flow Measurement and Instrumentation 83 (2022) 102109

Contents lists available at ScienceDirect

Flow Measurement and Instrumentation

journal homepage: www.elsevier.com/locate/flowmeasinst

Energy dissipation of flow over the type-B Piano Key Weir

Deepak Singh [a], Munendra Kumar [b,*]

[a] Dept. of Civil Engineering, Delhi Technological University, Delhi, India
[b] Dept. of Civil Engineering, Delhi Technological University, Delhi, India

ARTICLE INFO

Keywords:
Energy dissipation at base of PKW
Residual energy
Low head efficient structure
Steps at outlet key

ABSTRACT

Piano Key Weir (PKW) is an evolution with an efficient implementation over conventional labyrinth weir, especially at low heads. It is used as the crest of spillways and specific flow regulation applications for overflow control structures. The effects of a PKW on the downstream flood regime have not yet been thoroughly studied; only a few studies have focused so far on the downstream characteristics of the PKWs. The consideration of energy-dissipation is the main component of downstream elements for all hydraulic structures. Therefore, there is a need to log and assess details on the energy dissipation of the PKWs. This paper presents the results of the experimental study to investigate the energy dissipation at the base of type-B PKW under free-flow conditions in a rectangular channel. According to the current study's findings, the type-B PKWs show a higher energy dissipation efficiency at low heads than other nonlinear or ordinary weirs, but dissipation efficiency decreases as the discharge or head over the weir increases. The energy dissipation was found to decrease with the increasing magnification ratio and relative width ratio; however, the energy dissipation was found to increase with increasing cycle number (for a constant-width channel). In addition, the sensitivity of the weir, i.e., using the steps at the outlet key instead of the slopping floor, increases the efficiency of energy dissipation for low head discharges about 6.43% while decreasing by 2.05% for higher head discharges.

1. Introduction

Dams and reservoirs are among the most effective ways for our community to have long-term water reserves and flood protection. In order to the betterment of any country, dams and reservoirs are played a vital role [4]. Older dams require more maintenance or rehabilitation expenses than new ones because their storage capacity decreases due to sedimentation or siltation of the reservoir. Some of them face the problem of submergence [27]. The control of floods is generally assured that the passage or release of floods without creating any hazardous incident on dam structures. One of the essential aspects of dam building and exploration is protection since an accident can often have significant effects, to varying degrees, depending on the amount of water contained [40]. Higher peak flow values can result from the availability of more extensive hydrologic data series and climate change, causing dam overtopping due to undersized spillways. As a result, increasing the flow release capability of these structures is essential to improve their protection. The introduction of the labyrinth-style weir is one of the most common solutions. Indeed, this form of the weir will increase discharge while maintaining the same length as a traditional linear weir [2,23].

The PKW is continued to improve the form labyrinth concept, first depicted in 2003 by Ref. [24]. The first developers aimed to create a new labyrinth weir type with a smaller footprint while maintaining a structurally economical and straightforward structure that could be easily built [24,29]. At present, more than 35 PKWs have been built worldwide since 2006, Goulours dam first PKW installation in France [8], and research continues around all over the world, with an average of 15 significant contributions in scientific journals every year since 2010. In addition to the above developments, the idea of using a PKW as a frontal weir for flood release, energy dissipation, and aeration efficiency has recently shown interest in inside weirs applications [19].

Many studies have been conducted on the effect of hydraulics and geometric parameters on discharge efficiency, but very few studies have focused on energy dissipation over the PKW. The first systematic experimental studies have been ongoing with stepped spillway [17](or stilling basin [39,45]. Moreover, the aforementioned general research studies [3,10,16,21,22] have conducted project-specific and special studies on the energy dissipation of labyrinth and PKWs. Silvestri et al. [42] investigated the dissipation of energy overstepped chutes with a PKW crest. They noticed that the low residual energy at the spillway toe

* Corresponding author.
E-mail addresses: deepak_phd2k18@dtu.ac.in (D. Singh), munendrakumar@dtu.ac.in (M. Kumar).

https://doi.org/10.1016/j.flowmeasinst.2021.102109
Received 2 September 2021; Received in revised form 17 November 2021; Accepted 13 December 2021
Available online 30 December 2021
0955-5986/© 2022 Elsevier Ltd. All rights reserved.

and increases with the discharge and spillway length. According to Ho Ta Khanh et al. [16], PKWs combined with short stilling basins on small gravity dams can effectively dissipate energy. However, PKWs are better integrated with a short stilling basin and a stepped spillway on medium-to-large barriers/barricades. Erpicum et al. [10] found that the PKWs placed over stepped spillways provide faster energy dissipation than ogee weirs under similar conditions. They also stated that more work should be done to create a shorter stepped spillway. Al-Shukur and Al-Khafaji [1] investigated the effect of PKW slopes on energy dissipation. They observed that as the slope decreases, so does energy dissipation. The essential aspect for measuring the energy dissipation is the hydraulic jump distance measured from the toe of the PKW, which was zero for low slopes. They believed that as the distance to the hydraulic jump increased, the dissipation of energy decreased. Furthermore, Eslinger and Crookston [12] performed an experimental study to clarify the energy dissipation analysis at the base of type-A PKWs and proposed two empirical prediction equations (both are the functions of head-water ratio (H/P) and W_i/W_o, where H is the head over the weir, P is the weir height, W_i and W_o, represent the width of the inlet and outlet key, respectively) to predict the relative residual energy downstream of the PKW in order to encourage design efforts.

In addition, Silvestri et al. [42] described the crucial work of [9] by measuring four stepped spillway lengths and concluding that shorter-stepped spillway lengths downstream of the PKWs achieve uniform flow conditions and flow capacity, not those downstream of ogee-crested weirs. Studies of local scour at the toe of PKWs placed in canals and rivers are also related to the energy dissipation of PKWs. According to Noui and Ouamane [30], a ski-jump at the downstream end of the outlet key allows the water jet to avoid impacting the dam body; however, the performance of the PKW decreases while energy dissipation over the weir increases. Further, Jüstrich et al. [18] developed empirical equations to assess proper scouring and ridge generation on the downstream toe of PKW due to the loosening riverbed, which helps define the appropriate depth for the PKW foundation, installed on a movable riverbed. Pfister et al. [39] investigated the creation of toe scour at the PKW and noticed that if the foundation on rock is not possible, then toe-scour occurring during flood discharges is relevant to weir stability. Under certain conditions, the downstream face of the foundations can influence the flow emerging from the PKW to reduce its scouring potential. Palermo et al. [34] investigated the effects of various tested variables at equilibrium conditions on one uniform non-cohesive granular bed content under different discharge and tailwater depth combinations. They provided a preliminary analysis of scouring morphology downstream of a PKW.

Despite numerous valuable studies on energy dissipation over the labyrinth and PKWs, designers lack adequate knowledge from comprehensive research on energy dissipation and residual energy for the various types (e.g., Type-A, Type-B, Type-C, and Type-D) PKWs. To better understand this, the current study employed 12 laboratory-scale type-B, PKW models in various parametric configurations (see Table 1) to estimate energy dissipation downstream of the weir and corresponding flow characteristics. Furthermore, the sensitivity of the weir, i.e., using the steps at the outlet key instead of the slopping floor were also discussed. In this study, the experiments were carried out considering the channel flow approach (i.e., experiments were conducted over PKW models in laboratory flume); thus, this study covered mainly PKWs installed at river barrages or as a control structure in a canal.

2. Materials and method

2.1. Experimental setup and model configuration

In order to perform the experimental study, an experimental setup is built up in the "Fluid Mechanics and Hydraulics Research Laboratory" at Delhi Technological University, Delhi. The tests were performed in a

Table: 1
Ranges of data collected in the present study Type-B.

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 $\left(\frac{E_1}{E_1}\right)$	Range of $\left(E_r = \frac{E_2}{E_1}\right)$	N (No. of cycles)	No. of runs
PKW-1	0.005–050	0.0133–0.165	1.28	0.088	0.069	0.15	5	0.343	0	0.2286	0.8493–0.1671	0.1724–0.8329	3	25
PKW-2	0.005–050	0.0132–0.164	1.28	0.064	0.05	0.15	5	0.259	0	0.165	0.860–0.1396	0.1512–0.8604	4	25
PKW-3	0.005–050	0.0135–0.174	1.28	0.05	0.039	0.15	5	0.208	0	0.139	0.870–0.1667	0.1379–0.8333	5	25
PKW-4	0.005–050	0.0132–0.174	1.28	0.088	0.069	0.15	6	0.427	0	0.2842	0.833–0.1711	0.1924–0.8289	3	25
PKW-5	0.005–050	0.0135–0.172	1.28	0.064	0.05	0.15	6	0.322	0	0.214	0.8456–0.1406	0.1757–0.8595	4	25
PKW-6	0.005–050	0.0132–0.169	1.28	0.05	0.039	0.15	6	0.259	0	0.173	0.8597–0.1368	0.1555–0.8632	5	25
PKW-7	0.005–050	0.0136–0.155	1.0	0.078	0.078	0.185	6	0.427	0	0.2842	0.831–0.1531	0.1581–0.8465	3	20
PKW-8	0.005–050	0.0138–0.157	1.1	0.082	0.075	0.185	6	0.427	0	0.2842	0.822–0.146	0.1687–0.846	3	20
PKW-9	0.005–050	0.0139–0.154	1.2	0.085	0.071	0.185	6	0.427	0	0.2842	0.818–0.1535	0.182–0.84654	3	20
PKW-10	0.005–050	0.0138–0.154	1.3	0.089	0.068	0.185	6	0.427	0	0.2842	0.8012–0.1563	0.1987–0.8432	3	20
PKW-11	0.005–050	0.0141–0.151	1.4	0.091	0.065	0.185	6	0.427	0	0.2842	0.7934–0.1478	0.2065–0.8521	3	20
PKW-12	0.005–050	0.0138–0.158	1.5	0.094	0.063	0.185	6	0.427	0	0.2842	0.7843–0.1301	0.2156–0.8698	3	20
PKW-1 (Stepped)	0.005–050	0.0134–0.169	1.28	0.088	0.069	0.15	5	0.343	0	0.2286	0.893–0.1574	0.1077–0.8426	3	25
PKW-4 (Stepped)	0.005–050	0.0132–0.168	1.28	0.088	0.069	0.15	6	0.427	0	0.2842	0.865–0.1616	0.1356–0.8384	3	25

straight rectangular tilting flume of 10 m long, 0.506 m wide, and a height of 0.6 m. The width of the test section was reduced to 0.30 m (however, models were fabricated based on total width, 0.506 m (see in Fig. 1 a)). The channel is fed by a 20 HP (15 kW) pump connected to a series of 4-inch (0.102 m) supply pipes, delivering discharges of up to 50 L/s. The flume supply line is calibrated using an orifice meter (0.25% uncertainty), and the discharge can be adjusted with the assistance of a supply valve. Water enters the flume via the head tank, which contains a metal screen gate or baffle wall as well as a synthetic membrane or a manufactured film to ensure uniform flow conditions (to improve approach stream consistency). A tailgate fixed to the trunk channel controls the water depth at the downstream site. In order to see the visual assessment of the flow pattern, the Plexiglas sheet has been fixed up to 6.5 m in length on the sidewalls of the channel. Flume is furnished with a 4–20 mA ultrasonic level sensor (accuracy ± 0.2% ± 1 mm) instrumentation carriage as well as a ±0.1 mm pointer gauge for measuring the height of the water surface and crest elevations in various sections. The discharge will be measured using 4–20 mA electromagnetic flowmeters (accuracy ± 0.2% ± 1 L/s), as shown in Fig. 1 (a). The ADV (Acoustic Doppler Velocimeter) was used to measure the mean flow velocities. ADV is highly reliable for capturing turbulence characteristics in an open channel stream in the laboratory's research facility.

This study aims to enhance the understanding of energy dissipation downstream of the PKW (type-B) and corresponding flow characteristics by conducting experimental research on 12 PKW models with a constant parapet wall (R) of 0.02 m for each model, as shown in Fig. 2. Under the present study, the models are fabricated using a 0.008 m thick transparent acrylic sheet and affixed with the help of chloroform. The models were assembled so that the inlet-outlet width ratio varies between $1.0 \leq W_i/W_o \leq 1.5$. The magnification ratio (L/W) ratio varies between 5 and 6, and the height of the weir (P) varies between 0.15 and 0.185 m (the first six models are 0.15 m height, and the remaining six are 0.185 m height, see in Table 1). The inlet and outlet key slopes are $0.9 \leq S_i \leq 1.3$, and $0.33 \leq S_o \leq 0.43$, and the two overhang portions are such that $B_i = 0$, $B_o = 2/3B_b$, where B_b is the length of PKW along the channel in (m), B_i and B_o are the inlet and outlet overhang portions, respectively.

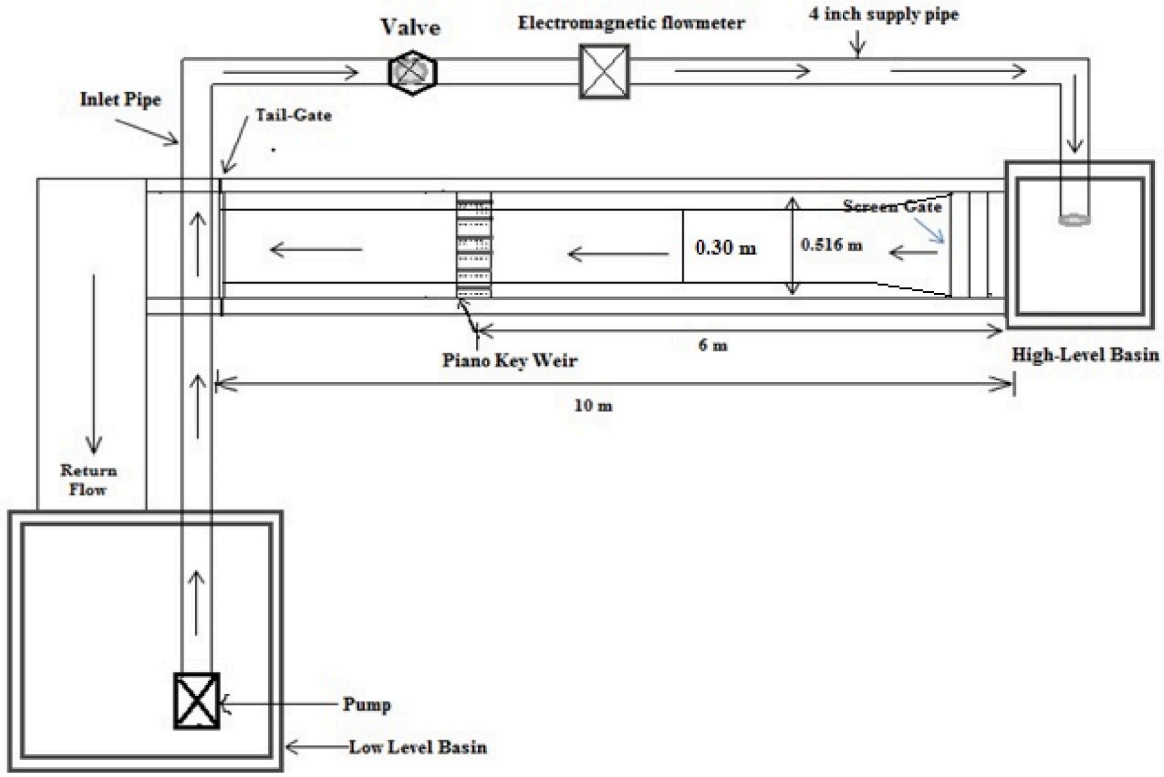


Figure: 1 (a). Schematic plan view of the experimental setup.

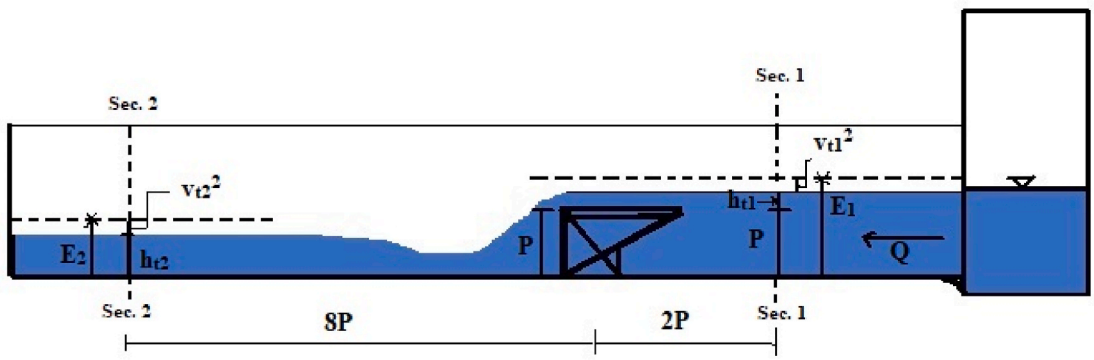


Figure: 1. (a). Schematic plan view of the experimental setup. (b). Schematic plot for specific energy measurement.

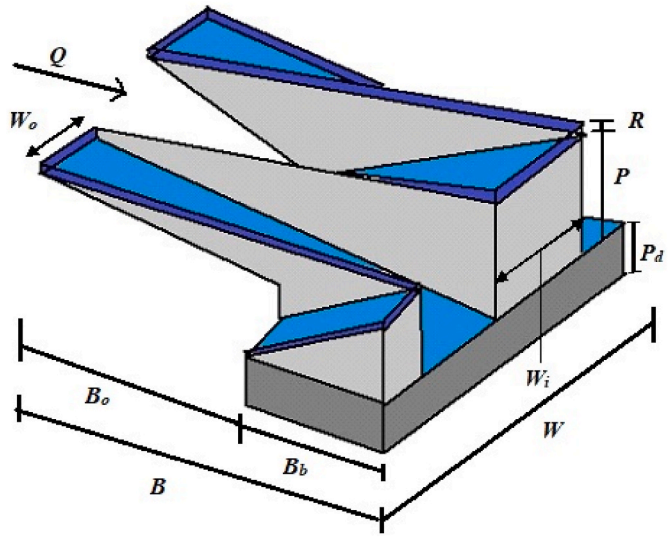


Figure: 2. Clarification sketch of PKW model 3D-view type-B.

2.2. Dimensional analysis

The current study was performed under free-flow conditions, and therefore the downstream or tailwater was not controlled artificially. The specific energy upstream and downstream of the PKW was calculated as follows:

$$E_i = P + h_{ti} + \frac{V_{ti}^2}{2g} \tag{1}$$

where E represents the specific energy at section i, P represents the height of the weir in (m) (for downstream of the weir the $P = 0$ m), h_t represents the head at section i, V_t is the velocity at section i, and i represents the section (i.e., $i = 1, 2 \ldots$, Section 1 was considered upstream, and section 2 was considered downstream as shown in Fig. 1 (b)). Upstream flow depth ($P + h_{t1}$) and downstream flow depth h_{t2} were measured carefully with a 4–20 mA ultrasonic level sensor as well as a pointer gauge of ±0.1 mm (accuracy ± 0.2% ± 1 mm) after the water surface had been allowed to steady-state for at least 3–5 min or it can increase for the higher flow rate) were taken in such a way that the flow or water level has been stabilized. In general, a more stabilization time is required for a higher discharge rate through the channel. With the help of ADV, the mean approach flow velocity V_{ti} was measured for 1-min records at the same section where the head h_{ti} values were measured. The heads over the PKWs were measured at a distance upstream $X = 2$ x P, where X is the distance from the lateral centreline of the PKWs [32].

The downstream site head was measured at a distance of $8P$ from the lateral centreline of the PKWs, after the downstream water level approximately stabilized, as shown in Fig. 1 (b). The ranges of the data collected in the present study are shown in Table 1.

The values of E_1 and E_2 were used to compute the relative energy dissipation based on Eqn. (2) and relative residual energy ($E_r = E_2/E_1$) is calculated considering the total energy downstream is equal to 100%.

$$E_L = \frac{(E_1 - E_2)}{E_1} \times 100 = \left(1 - \frac{E_2}{E_1}\right) \times 100 \tag{2}$$

and,

$$E_r = 1 - E_L = \frac{E_2}{E_1} \tag{3}$$

where E_L denotes total relative energy dissipation or energy dissipation ratio, and E_r (=E_2/E_1) is the relative residual energy downstream of the PKW.

Hydraulic and geometrical parameters involved in energy dissipation are arranged in Eqn. (4) to determine those that affect energy dissipation over the PKW.

$$E_L = f\left(\rho,\ g,\ V_t,\ H_t, L, L_{cy}, P, W, W_i, W_o, W_{cy}, S_i,\ S_o, R,\ h_s,\ l_s\right) \tag{4}$$

In which ρ is the density of flowing fluid, g is the acceleration of gravity, and H_t is the total upstream hydraulic head ($h_t + V_t^2/2g$) over the PKWs. The other geometrical parameters are related to the PKW, where L = total developed crest length; L_{cy} = crest length of one cycle; P = height of PKW, W = total width of PKW; W_i and W_o = widths of inlet and outlet keys sections, respectively W_{cy} = width of one cycle and S_i and S_o represents the inlet and outlet key slopes variation, respectively, R = height of parapet wall, l_s and h_s = length and height of the steps, and N = number of cycles. Using the Buckingham theorem as a dimensional analysis technique and by considering the ρ, g, and P as repetitive parameters (Eqns. (5)–(7), the dimensionless parameters are derived as Eqn. (8). Usually, the flow regime upstream of weirs constructed on the perpendicular to the river course is always subcritical; hence, the Froude number is not considered in evaluating their hydraulic properties [15]. Therefore, Eqn. (8) can be rewritten as Eqn. (9).

$$\prod_1(H_t) = \frac{H_t}{P},\ \prod_2(L) = \frac{L}{P},\ \prod_3(L_{cy}) = \frac{L_{cy}}{P},\ \prod_4(W) = \frac{W}{P},\ \prod_5(W_i) = \frac{W_i}{P},\ \prod_6(W_o) = \frac{W_o}{P},\ \prod_7(W_{cy}) = \frac{W_{cy}}{P},$$
$$\prod_8(S_i) = S_i\ ,\ \prod_9(S_o) = S_o,\ \prod_{10}(R) = \frac{R}{P},\ \text{and}\ \prod_{11}(V_t) = \frac{V_t}{\sqrt{gP}} = F_r,\ \prod_{12}(l_s) = \frac{l_s}{P},\ \prod_{13}(h_s) = \frac{h_s}{P} \tag{5}$$

$$\prod(L) \times \frac{1}{\prod(W)} = \frac{L}{P} \times \frac{P}{W} = \frac{L}{W},\ \prod(W_i) \times \frac{1}{\prod(W_o)} = \frac{W_i}{P} \times \frac{P}{W_o} = \frac{W_i}{W_o},\ \prod(S_i) \times \frac{1}{\prod(S_o)} = \frac{S_i}{S_o},\ \prod(h_s) \times \frac{1}{\prod(l_s)} = \frac{h_s}{l_s} \tag{6}$$

$$\prod(L) \times \frac{1}{\prod(L_{cy})} = \frac{L}{L_{cy}} = N = \prod(W) \times \frac{1}{\prod(W_{cy})} = \frac{W}{W_{cy}} = N \tag{7}$$

By combining or rearranging these parameters, the relationship is obtained:

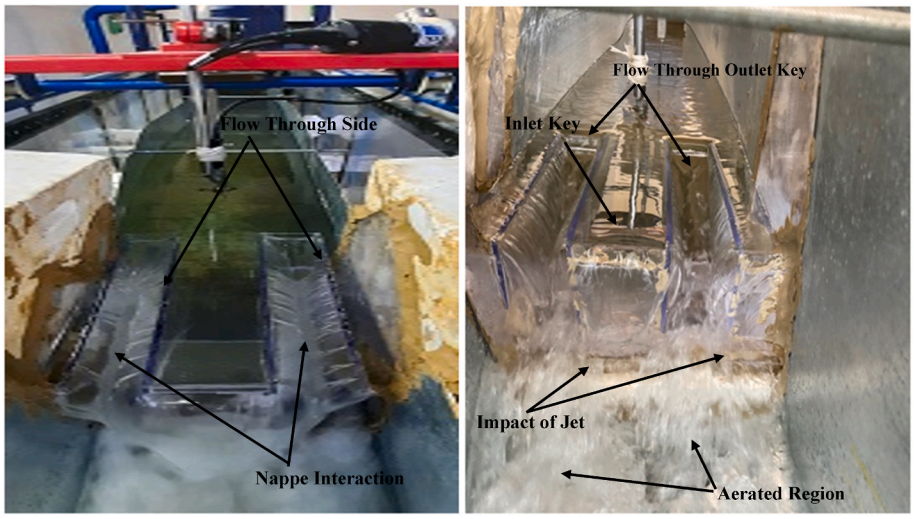


Figure: 3. Flow pattern over PK Weir.

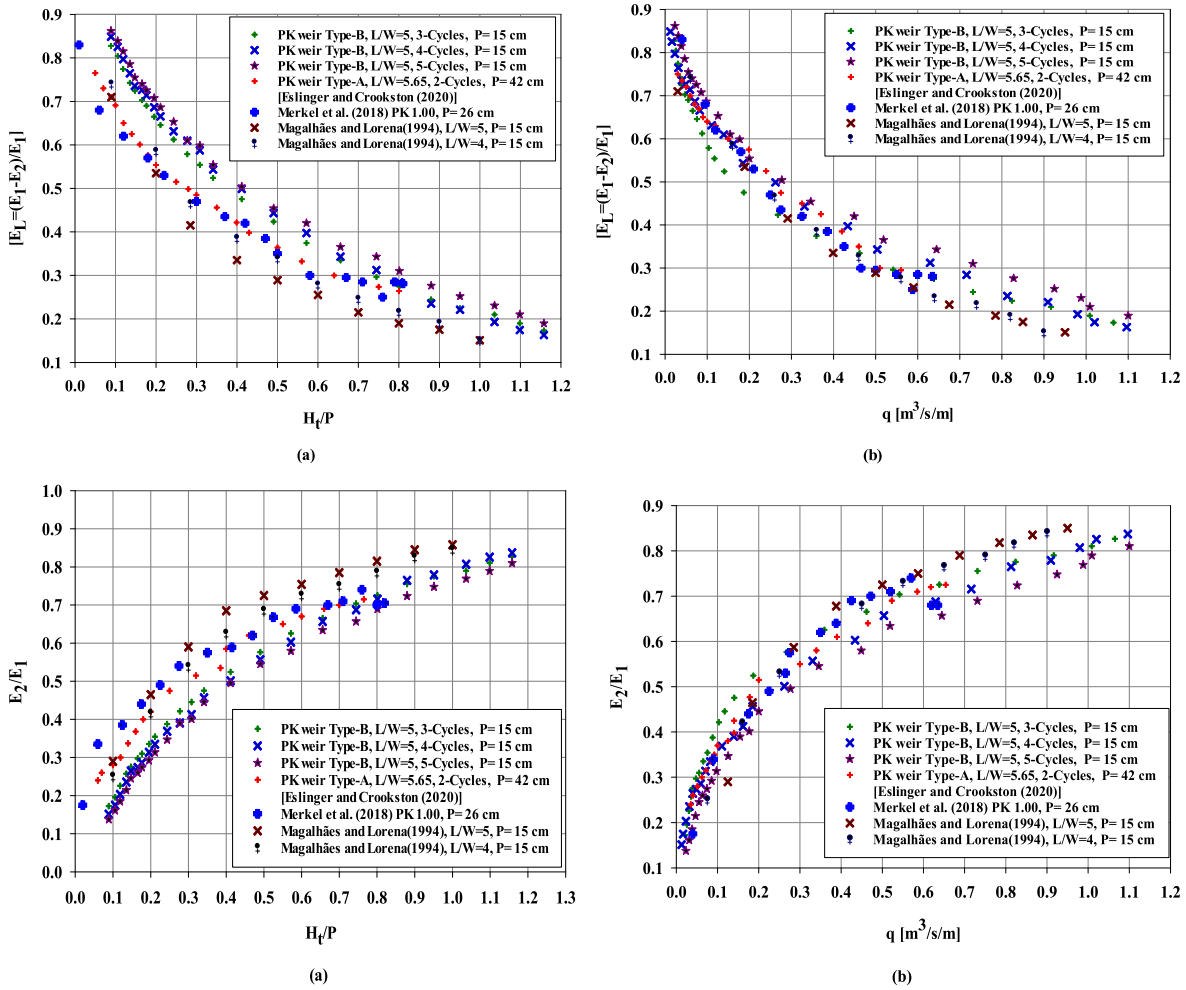


Figure: 4. (i) (Relative energy dissipation [$E_L = (E_1\text{-}E_2)/E_1$] with respect to (a) the head water ratio [H_t/P] and (b) the unit discharge [q]) for [$L/W = 5$]. (ii) (Relative residual energy [E_2/E_1] with respect to (a) the head water ratio [H_t/P] and (b) the unit discharge [q] for [$L/W = 5$].

$$\frac{E_L}{E_1}=f\left(\frac{H_t}{P},\ \frac{L}{W},\frac{W_i}{W_o},\ \frac{S_i}{S_o},\frac{h_s}{l_s},\ N,F_r,\frac{R}{P}\right) \tag{8}$$

For subcritical flow

$$\frac{E_L}{E_1}=f\left(\frac{H_t}{P},\ \frac{L}{W},\frac{W_i}{W_o},\ \frac{S_i}{S_o},\frac{h_s}{l_s},\ N,\frac{R}{P}\right) \tag{9}$$

3. Experimental results and discussions

The basic concept of the flowing discharge over the PKW is that flow converges into the inlet key and diverges away from the outlet key [25]. The total flowing discharge over the weir summarizes three discharges, discharge through the inlet keys, discharge through the outlet keys and discharge through the side crest walls. All the three modes of the flow of discharges meet consequently in an intricate flow [20]. The flow through the PKW is highly aerated and three-dimensional, with splash and spray regions within the outlet keys and at the base of the structure (see Fig. 3). The area of spray and sprinkling only increased marginally proportional to the trajectory of H_t and the planar jet that started downstream on the crest. But the aeration region increased significantly with H_t, partly because the local speed increased, resulting in greater advection levels and turbulent mixing. The nappes passing over the PKW crest are considered as three planar jets interacting over a slanting surface (see Fig. 3), exhibiting jet expansion, jet interaction subjected by opposing nappes with a water surface conflict zone, crossing region, and small wakes near the upstream apexes.

According to Pagliara et al. [33], the nappes acting in opposite directions contribute to energy dissipation, as previously observed in crossing jet studies. As upstream heads H_t and discharge Q increase the volumetric space of the outlet keys, the flow field evolves clearly, limiting jet interaction and energy dissipation caused by crossing jets. During the investigation, the local submergence is noticed to increase and occupy a more significant percentage for the region of $H_t/P \geq 0.25$ until the majority of the $B_o + B_b$ streamwise distance is locally submerged. As discussed in the following section, local submergence reduces discharge capacity and effectively considers energy dissipation reduction for higher H_t/P ratios.

The primary goal of this study was to collect data on energy dissipation over type-B PKWs and calculate specific energy using Eqn. (1) based on upstream and downstream head measurements. The relative energy loss over the PKW and relative residual energy was determined using Eqn. (2) & Eqn. (3), respectively. The analysis was carried out using various parametric ratios, a total of 320 tests comprising these new datasets, and detailed observations of the flow field. The experimental data results were documented and plotted with various hydraulic characteristics such as the relative energy dissipation [$E_L = (E_1\text{-}E_2)/E_1$], relative residual energy (E_2/E_1), and the H_t/P ratio, unit discharges (q) are shown in Fig. 4 to 10.

As presented in the material and methods section, the flow energy dissipation over the PKWs are proportional to the various non-dimensionless parameters. In the present study, the values of the flow energy dissipation over the PKWs have been analyzed in three ways: the first one was to determine the effect of cyclic variation of the weir on energy dissipation over the PKW while keeping the L/W and W_i/W_o ratios were constant. The second aspect is comprehending the impact of varying the L/W ratio on energy dissipation by keeping the number of cycles and the W_i/W_o ratio constant. And the third aspect is to see how changing the inlet-outlet width (W_i/W_o) ratio affects the energy dissipation at the base of the PKW while keeping the L/W ratio and number of cyclic counts constant. The detailed discussion is described in the following subsections later on.

The testing discharges were varied over the model between 0.005 $m^3/s \leq Q \leq 0.05\ m^3/s$. The head-discharge data associated with the testing were collected $0.088 \leq H_t/P \leq 1.16$, for 0.15 m height models; $0.075 \leq H_t/P \leq 0.85$, for 0.185 m height models, and other parametric variations are as followed; 0.013 m $\leq H_t \leq$ 0.174 m, $B_i/P = 0$, $1.1 \leq \{B_o/P\} \leq 1.89$. In the present study, the total head over the PKW varies 0.013 m $\leq H_t \leq$ 0.174 m, at which few H_t values are less than 0.03 m so that the scale effect would come to the role for these critical values [11, 31]. Scale effects consideration on PKWs were discussed in detail by Erpicum et al. [11]. The projects that come under the physical modelling are to be designed by considering the large-scale variables (i.e., small dimensional physical models). The flow characteristics of such models may be affected by not-scaled parameters such as atmospheric pressure, water viscosity, and surface tension. The viscosity and surface tension are fluid properties that cannot be scaled, resulting in scale effects on weir crests for low overflow depths. According to Machiels [25], the surface tension and viscous forces at the crest-water-air interface come to the role of existence, and for a few seconds, no discharge over the weir exists, for low head $H_t/P \leq 0.06$. When the head is high enough to overcome surface tension everywhere, the entire crest becomes involved in discharge conveyance. As a consequence, scaling does not fully achieve model and prototype similarity. Such discrepancies are referred to as size-scale effects. They can alter the head-discharge relationship, nappe trajectory, and air entrainment [46]. According to Tullis et al., [46], the total head over the PKW and scaled-to-prototype head thresholds above which negligible scale effects were observed are $H_t/P \geq 0.09$. The more energy dissipates for low (H_t/P) values in this study, and the authors did not consider the scale effect in the present study.

Table: 2
Q comparison for corresponding H_t between the present study and published data by [Young (2018) and (Khassaf et al., 2015)].

	Models	Young (2018)			(Khassaf et al., 2015)		
S. No.	Type-B	MAPE	RMSE	R^2	MAPE	RMSE	R^2
1	PKW-1	2.50%	0.001	0.998	2.80%	0.001	0.967
2	PKW-2	3.70%	0.001	0.997	2.02%	0.003	0.985
3	PKW-3	2.50%	0.003	0.995	3.23%	0.004	0.985
4	PKW-4	3.70%	0.006	0.99	2.90%	0.003	0.978
5	PKW-5	4.04%	0.007	0.977	2.44%	0.002	0.976
6	PKW-6	3.59%	0.003	0.973	3.81%	0.032	0.971
7	PKW-7	2.50%	0.018	0.990	2.30%	0.007	0.987
8	PKW-8	2.70%	0.003	0.981	3.91%	0.003	0.975
9	PKW-9	1.50%	0.011	0.977	2.15%	0.019	0.968
10	PKW-10	4.70%	0.008	0.965	2.71%	0.004	0.991
11	PKW-17	3.32%	0.006	0.994	2.32%	0.002	0.956
12	PKW-12	3.14%	0.004	0.957	4.14%	0.005	0.964
7	PKW-1-stepped	2.7%	0.006	0.992	2.85%	0.028	0.989
8	PKW-4-stepped	2.15%	0.005	0.994	3.12%	0.031	0.976

[1]MAPE = Mean Absolute Percentage Error.
[2]RMSE = Root Mean Square Error.
[3]R^2 = Coefficient of determination.

Furthermore, the mean absolute percentage error (MAPE), coefficient of determination (R^2), and root mean square error (RMSE) between the current study and that of [20,48] were computed first and compared Q for a measured H_t value as shown in Table 2. Then [$E_L = (E_1\text{-}E_2)/E_1$] and ($E_r = E_2/E_1$) were compared to datasets by Eslinger and Crookston [12], of PKW Type-A, L/W = 5.65, 2-Cycles, P = 42 cm; trapezoidal labyrinth weirs of $L/W = 4, 5$; P = 15 cm, by Magalhães and Lorena [26]; rectangular labyrinth weirs of P = 26 cm, by Merkel et al., [28]. After comparing the data of current and published studies, it was noticed that the type-B PKW has the highest energy dissipation at low heads compared to the other types of weirs. This may partly be due to the geometric configuration like slopping inlet-outlet key bottoms, overhangs length, and partly due to the complex flow behavior of the PKW.

3.1. Effects of the cyclic variation (N) on energy dissipation

The primary consequence of the study is to find out the impacts of cyclic variation on energy dissipation. After investigating the

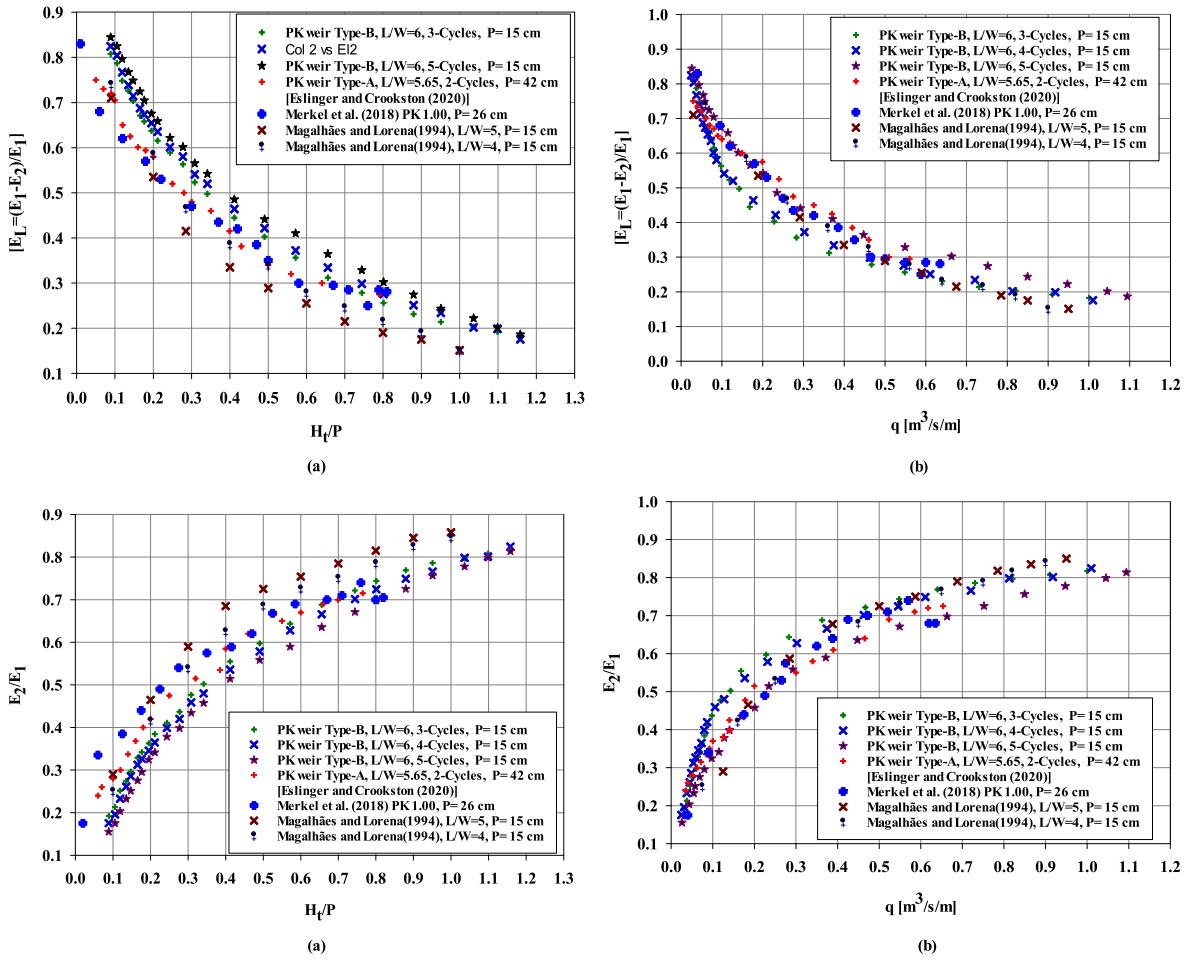


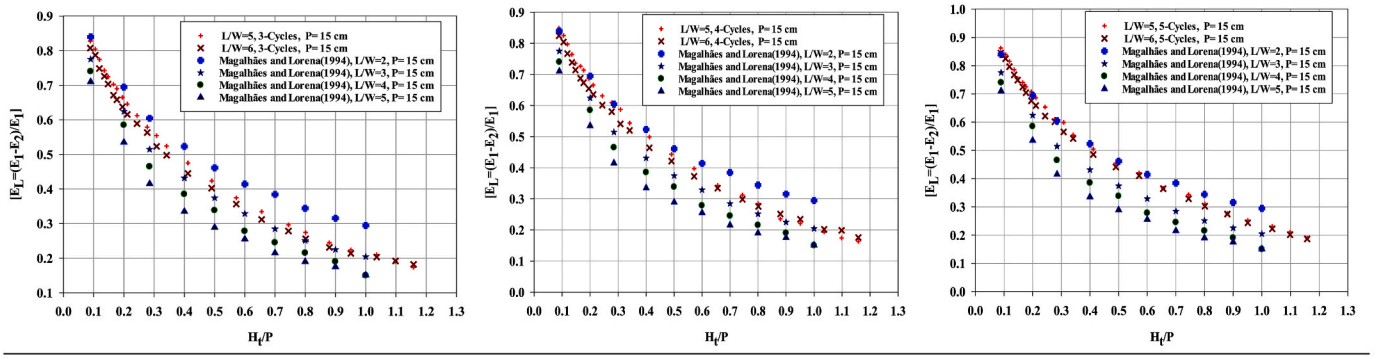
Figure: 5. (i) (Relative energy dissipation [E_L=(E_1-E_2)/E_1] with respect to (a) the head water ratio [H_t/P] and (b) the unit discharge [q] for [$L/W = 6$]. (ii) (Relative residual energy [E_2/E_1] with respect to (a) the head water ratio [H_t/P] and (b) the unit discharge [q] for [$L/W = 6$].

experimental data, it was evident that the PKW shows significant energy dissipation variation with varying cycles or units. According to the current study's findings, the data for PKW type-B show a pattern of convergence in relative energy dissipation at both low and high values of H_t/P. The relative energy dissipation for the nonlinear weirs [$E_L = (E_1$-$E_2)/E_1$] has shown to be inversely proportional to H_t/P with a nonlinear pattern, as are shown in Figs. 4 and 5.

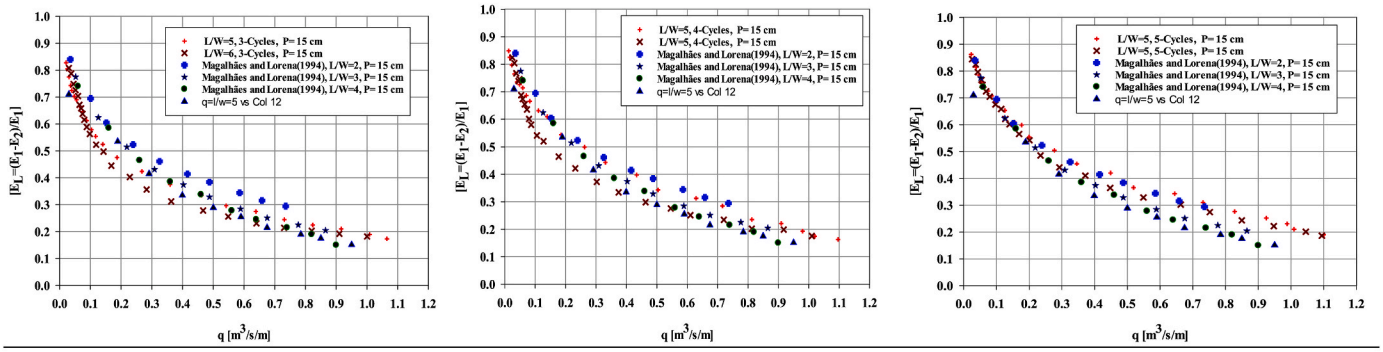
During the investigation, it was observed that the lowest flow depth has the greatest energy dissipation, and the PKW is less effective at dissipating energy at higher heads. A marginal change in H_t/P results in a relatively significant difference in relative energy dissipation when the headwater ratio H_t/P is less than 0.25. The relative energy dissipation was found to be constant $H_t/P > 1.1$ and between $0.25 \leq H_t/P \leq 1.1$, the rate of [$E_L = (E_1$-$E_2)/E_1$] relative energy dissipation steadily decreases as H_t/P increases. This is because the energy is dissipated in the outlet key section for small discharges, and the nappes are aligned partially laterally and partially in the flow direction. The water level in the keys rises for higher discharges due to local submergence effects [6], and the nappes are only aligned in the flow direction. As a result, energy dissipation decreases or remains nearly constant. This point concludes that the PKW represents the better relative energy dissipation efficiency at low discharges or low head level; an increasing the discharges or heads, its relative energy dissipation decreases gradually or sometimes rapidly.

The number of cycles plays an essential role in influencing the discharge capacity when the width is limited or constrained. According to Tulli et al. [47], as increasing the number of units or cycles for fixed width, the discharge carrying capacity of the folded weir decreases. In the current study, the models were fabricated and tested while keeping the width of the model constant. So, this section of the study is mainly focused on the influence of cyclic variation on the dissipating energy at the base of the type-B PKW and found that for the same Magnification ratio (L/W), the relative energy dissipation [$E_L = (E_1$-$E_2)/E_1$] at the base of PKWs increases with the number of cycles (units). In addition, a closer look at Figs 4 and 5 reveal that the relative energy dissipation increases more with the number of cycles at low values of H_t/P (both the L/W ratio), and in contrast, for higher H_t/P values intermixing patterns of energy dissipation were observed. This is partly due to the progressive change in the nappe portion affected by local submergence [6]. The energy dissipation provided by jet collision and downstream impact is altered and thus reduced as the labyrinth or PKW nappe conditions transition from aerated to partially aerated and then drown [7]. The abovementioned trends hold true when comparing relative energy dissipation to unit discharge (q) of dissimilar structures such as vertical drops and overflow weirs.

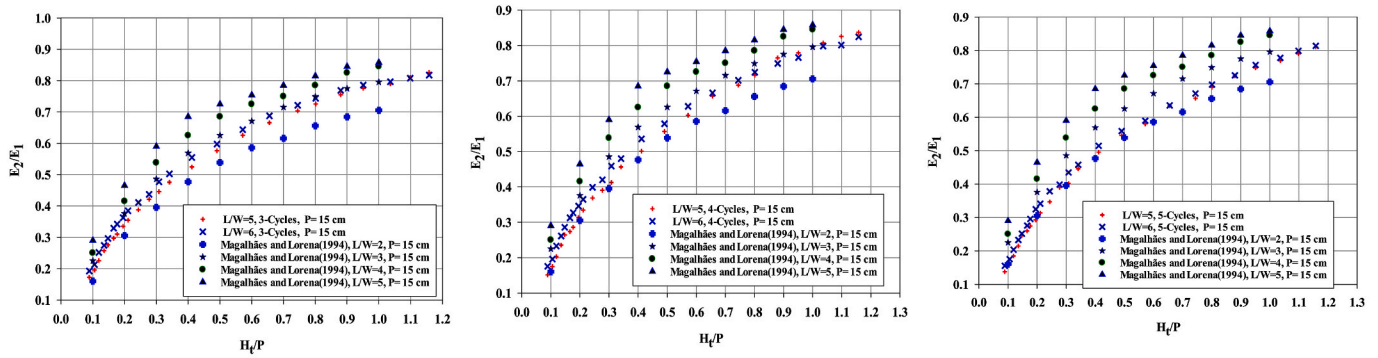
The maximum relative energy dissipation was observed in the present study corresponding to the highest cyclic count (i.e. $E_L = 0.8621$ or 86.21% the corresponding $N = 5$ and $L/W = 5$) and the less energy dissipation for the lowest cyclic counts (i.e. $E_L = 0.8077$ or 80.77% the corresponding $N = 3$, and $L/W = 6$) (see Fig. 4). As discussed above, the relative energy dissipation of PKWs exhibits a dependence on the number of cycles (units) at $0.25 \leq H_t/P \leq 1.1$. Thus the number of cycles is a convenient parameter. Moreover, for comparison and reference, published data for laboratory-scale trapezoidal labyrinth weirs [26], rectangular labyrinth weirs [28], and PKW [12] were included. While the results were compared to previous studies within the accuracy of the measurements, all of these models had shown a very similar trend to



(a)



(b)



(a)

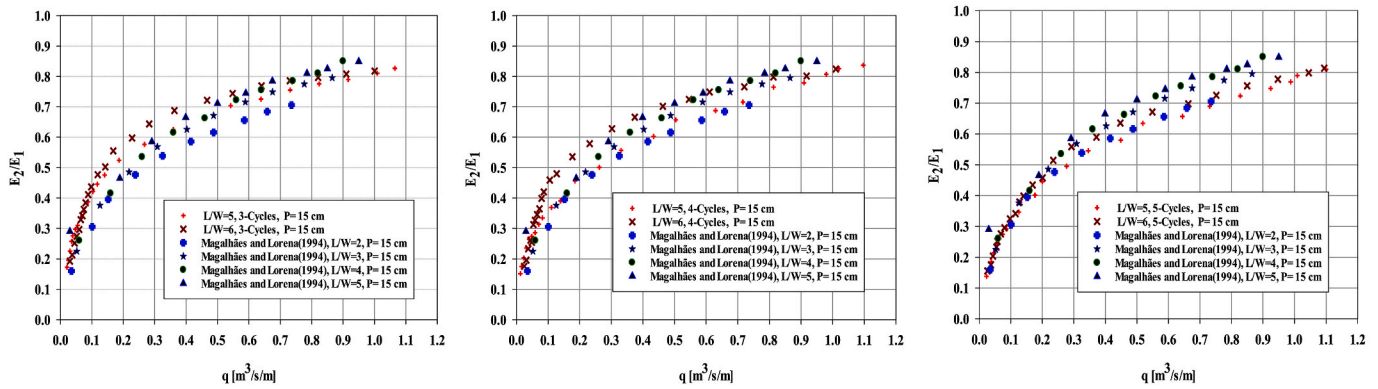


Figure: 6. (i) (Relative energy dissipation $[E_L=(E_1\text{-}E_2)/E_1]$ with respect to (a) the head water ratio $[H_t/P]$ and (b) the unit discharge $[q]$ for $[L/W = 5, 6]$. (ii) (Relative residual energy $[E_2/E_1]$ with respect to (a) the head water ratio $[H_t/P]$ and (b) the unit discharge $[q]$ for $[L/W = 5, 6]$.

other tested models. However, the minor differences in energy dissipation are due, in part, to experimental uncertainties and different experimental geometries.

3.2. Effects of the relative length (L/W) on energy dissipation

The second consequence of this study is to highlight the effects of the magnification ratio (L/W) on the relative energy dissipation of the type-B PKW. The magnification ratio (L/W) is the most crucial parameter influencing the discharge carrying capacity of PKWs. According to Leite Ribeiro et al. [23], the discharge carrying capacity of the PKW enhances with the L/W ratio. However, the energy dissipation capacity of folded shape weir decreases with increasing the magnification ratio (L/W) by Magalhães and Lorena [26].

Magalhães and Lorena [26] investigated one-cycle trapezoidal labyrinth weirs with WES-type crest profiles for a different magnification ratio, and they noticed that the relative residual energy (E_2/E_1) at the base of labyrinth weirs increases with the magnification ratio (L/W). This trend may be highlighted because more significant discharges are required for the same relative total upstream head (H_t/P) and magnification ratio (L/W). They also reported more relative residual energy for $L/W = 5$, which may be due to the difference B and in crest shape (overhanging profile) and cycle geometry. Finally, they concluded that the energy loss of labyrinth weirs is highly dependent on weir geometry.

In the present study, the authors put the effort to understand better the cyclic influence and the effects of the magnification ratio on the energy dissipation capacity of the PKW and compared it with the previously published data. The test results indicated that the energy dissipation at the base of the PKW shows a decreasing trend with increasing the L/W ratio. In this approach, L varies with the head as the effective crest length decreases with increasing heads because of local submergence on the upstream apex. Flow drowning and lateral jet overcrossing occur as the head increases, reducing the dissipation efficiency of the PKW.

The current study shows that the energy dissipation over the PKW (Type-B) is higher than its alternatives, except $L/W = 2$ of the trapezoidal labyrinth weirs [26], but it follows the same trend. Fig. 6 (i) & (ii) depicts that regardless of the magnification ratio (L/W), the relative energy dissipation [$E_L = (E_1\text{-}E_2)/E_1$]at the base of PKWs decreases with the relative total upstream head (H_t/P), especially for smaller values of the (H_t/P). The relative residual energy at the base of PKWs (E_2/E_1) increases with the magnification ratio (L/W), which was similar to Magalhães and Lorena [26], as shown in Fig. 6 (ii). The relative residual energy at the base of PKWs ranges from 0.14 to 0.84, corresponding to the energy losses of 0.86 and 0.16. As a result, PKWs (type-B) are highly effective at dissipating energy, especially at low relative total upstream

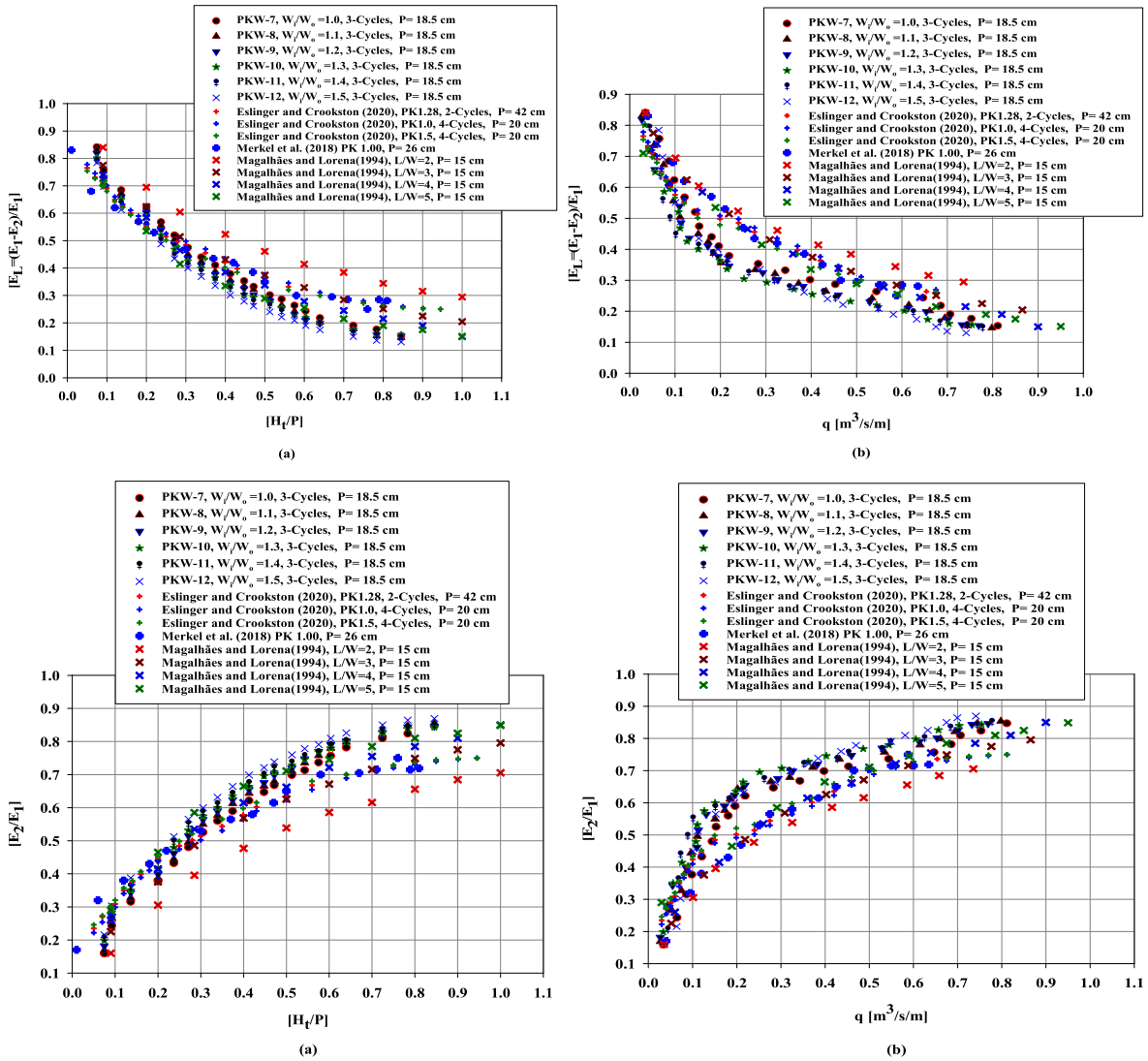


Figure: 7. (i) (Relative energy dissipation [E_L=(E_1-E_2)/E_1] with respect to (a) the head water ratio [H_t/P] and (b) the unit discharge [q] for different W_i/W_o ratios. (ii) (Relative residual energy [E_2/E_1] with respect to (a) the headwater ratio [H_t/P] and (b) the unit discharge [q] for different W_i/W_o ratios.

head (H_t/P) values.

Indeed, the L/W ratio demonstrates how effectively a design uses the available width W to maximize the crest length L. The flow control section must then be located along the crest to maximize the discharge on the developed crest length L. It means that the inlet should not restrict flow over the developed crest, and flow conditions in the outlet should not cause it to submerge. The sidewalls of the inlet section account for the majority of the developed crest length. In order to maximize the crest discharge, the inlet cross-section must have a discharge capacity at each point that is at least equal to one of the downstream developed crest length sections. The outlet key is responsible for evacuating the upstream streamwise flow and the lateral overflow from the inlet key. The operation of the outlet determines the efficiency of a given PKW.

3.3. Effects of the relative width ratio(W_i/W_o) on energy dissipation

To provide additional insight and demonstrate the effects of different inlet-outlet key width ratios on energy dissipation, six laboratory-scaled type-B PKW models (see in Table 1, PKW-7 to PKW-12) were tested. Following testing, the rating curves for each PKW were established, analyzed, and relative energy dissipations were compared. For comparison and reference, published data for laboratory-scale trapezoidal labyrinth weirs [26]; rectangular labyrinth weirs [28]; and PKW [12] are included (see Fig. 7 (i) & (ii)). All the models showed very similar trends for relative energy dissipations, while their results were compared with previous studies within the accuracy of the measurement.

In the present study, the energy dissipation rate was found to be more when H_t/P is less than 0.27 (approximate for all the models) than in previous studies. The rate of relative energy dissipation was found to be less for $H_t/P > 0.48$ and between $0.27 \leq H_t/P \leq 0.48$, the rate of $[E_L = (E_1\text{-}E_2)/E_1]$ relative energy dissipation has been observed intermixing in nature.

Fig. 7 (i) & (ii) show the variation of the relative energy dissipation $[E_L = (E_1\text{-}E_2)/E_1]$, and residual energy [Er = (E2/E1)] at the PKWs toe as a function of the upstream head ratio (H_t/P) and the specific discharge q. The relative residual energy (E_2/E_1) at the base of PKWs increases with H_t/P, particularly for smaller values, regardless of the relative width ratio (W_i/W_o). From Fig. 7 (i) & (ii), it is clear that the maximum relative energy dissipation was observed in the present study corresponding to the lowest width ratio (i.g. $E_L = 0.8513$ or 85.13% the corresponding $W_i/W_o = 1$) and the less energy dissipation for the highest width ratio (i.g. $E_L = 0.772$ or 77.20% the related $W_i/W_o = 1.5$). Increasing inlet key width reduces overall head losses due to the flow entering the inlet key, and increasing the inlet key flow area increases the flow carrying capacity of the inlet key. A high W_i/W_o value improves flow approach and distribution within PKWs inlet keys, whereas a high W_i/W_o value increases submergence effects in outlet keys. Submergence effects in the outlet cycles (regions where the flow depth in the outlet cycle exceeds the weir crest elevation) can reduce the discharge efficiency of the weir.

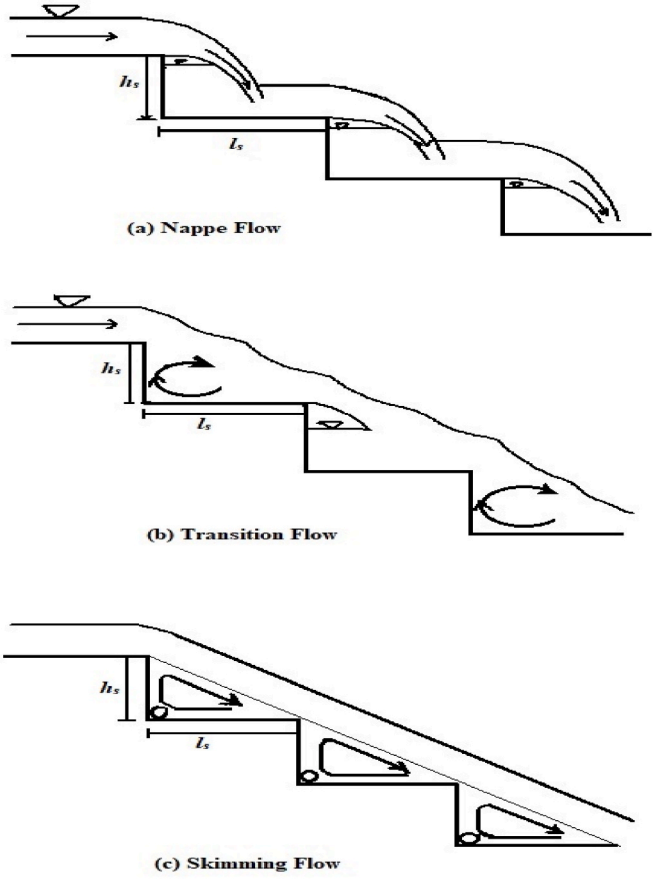


Figure: 9. Flow regimes above stepped weir: (a) Nappe Flow, (b) Transition Flow & (c) Skimminf Flow.

4. Sensitivity analysis

In this section, the sensitivity of the PKW has been assessed by constructing the steps at the outlet sections instead of the slanted floors (See in Fig. 8). The stepped spillway significantly dissipates the energy of passing flow and markedly reduces the occurrences of cavitation [4, 35]. Hence, the current study has focused on the energy dissipation mechanism using the steps at the outlet key section of the PKWs.

In order to determine the effect of the steps at the outlet key rather

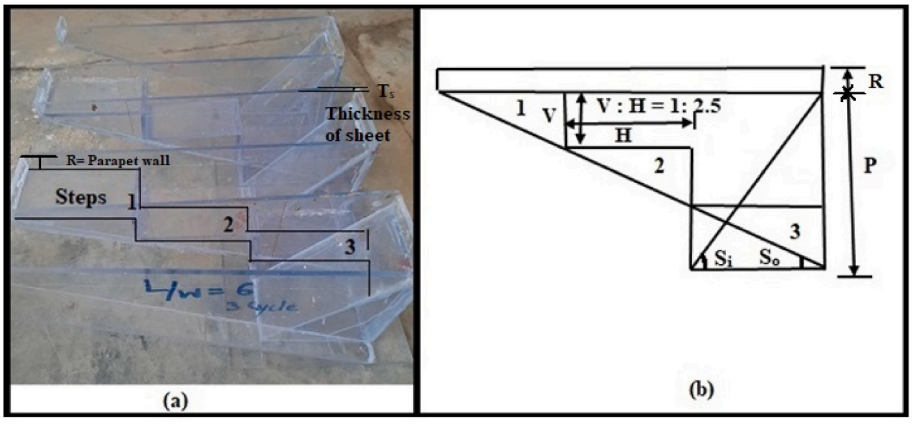


Figure: 8. Type-B PK weir after using the steps at outlet key (a) laboratory model (b) Sectional view.

than the sloping floor, two above tested laboratory-scaled PKW models (with the specification $L/W = 5$, 3 cycles, $P = 15$ cm, and $L/W = 6$, 3 cycles, $P = 15$ cm) were modified using the three uniform steps (all with a slope of 1*V*: 2.5*H*) at outlet key, whereas all other geometric parameters were kept constant. The testing discharges were varied over the model between 0.05 $m^3/s \leq Q \leq 0.05$ m^3/s. Other relative parametric variations associated with the testing are $0.088 \leq H_t/P \leq 1.13$, 1.32 cm $\leq H_t \leq 16.90$ cm, $B_i/P = 0$, $1.52 \leq \{B_o/P\} \leq 1.89$.

To better understand the effect of steps at the outlet key, knowing the flow behavior over the PKW with or without steps is necessary. The steps significantly increase the energy dissipation rate along the face of the spillway and eliminate or significantly reduce the need for a large energy dissipator at the weir's toe [4]. According to Parsaie et al. [35], drop number, critical depth to the height of steps ratio, and Froude number are the most influential parameters on energy dissipation of flow over stepped spillways. The step's roughness also plays a vital role in energy dissipation over the stepped spillway by Torabi et al. [44].

The flow configuration over stepped spillways can be classified into three types (napped flow, transition flow, and skimming flow) based on observation of the flow pattern (see Fig. 9). Napped flow condition occurs at low head/discharges values. In this condition, flow leaves the upper step and falls onto the lower step [13,43]. During this condition, energy dissipation is caused by the collision of a flow jet with steps and hydraulic jumps that may occur either entirely or partially. On the other hand, for large discharge values, a skimming flow condition occurs, and in this situation, a pseudo-bottom is created between steps and passing flow. In this case, a large amount of energy is dissipated in maintaining stable vortices beneath the pseudo-bottom formed by the external edges of the steps. The vortices are strengthened through the turbulent shear stresses between the skimming stream and the recirculating fluid [4]. Notably, the turbulent kinetic energy dissipation above the steps must be carefully estimated, especially for large discharges per unit width associated with the skimming flow regime. The high turbulence level and free-surface aeration characterize skimming flows [5,14,38,41]. A transition regime is a condition between napped and skimming flow.

According to the present study, energy dissipation rating curves followed the same pattern as the without using the steps at the outlet key but with steps showing higher energy dissipation. In the case of the stepped spillway, the chute slope and the size of the steps greatly influence the energy dissipation [36,37]. A comparison of rating curves using steps at the outlet key and without using steps are shown in Figs 10 (i) and (ii). These results indicate a large proportion of the energy dissipated by water flowing over the steps at the outlet key rather than the sloping floor for both models. At a low head (H_t/P) ratio, the gain is approximately 5.67%, 6.43%, respectively, while at a high head (H_t/P) ratio, the decrease is about 2.05%, 1.67%, respectively (see Fig. 10). It partly maybe the water flows down the steps as a coherent free-stream at the low heads, skimming over the step edge's formed pseudo-bottom. The turbulent recirculation in the step cavities is maintained by the transmission of shear stress from the free stream. At the free surface, the air is constantly trapped and released. The water level in the outlet keys rises over the steps for higher heads or discharges due to local submergence effects. Consequently, a two-phase mixture interacts with flow turbulence, resulting in intricate air-water structures associated with

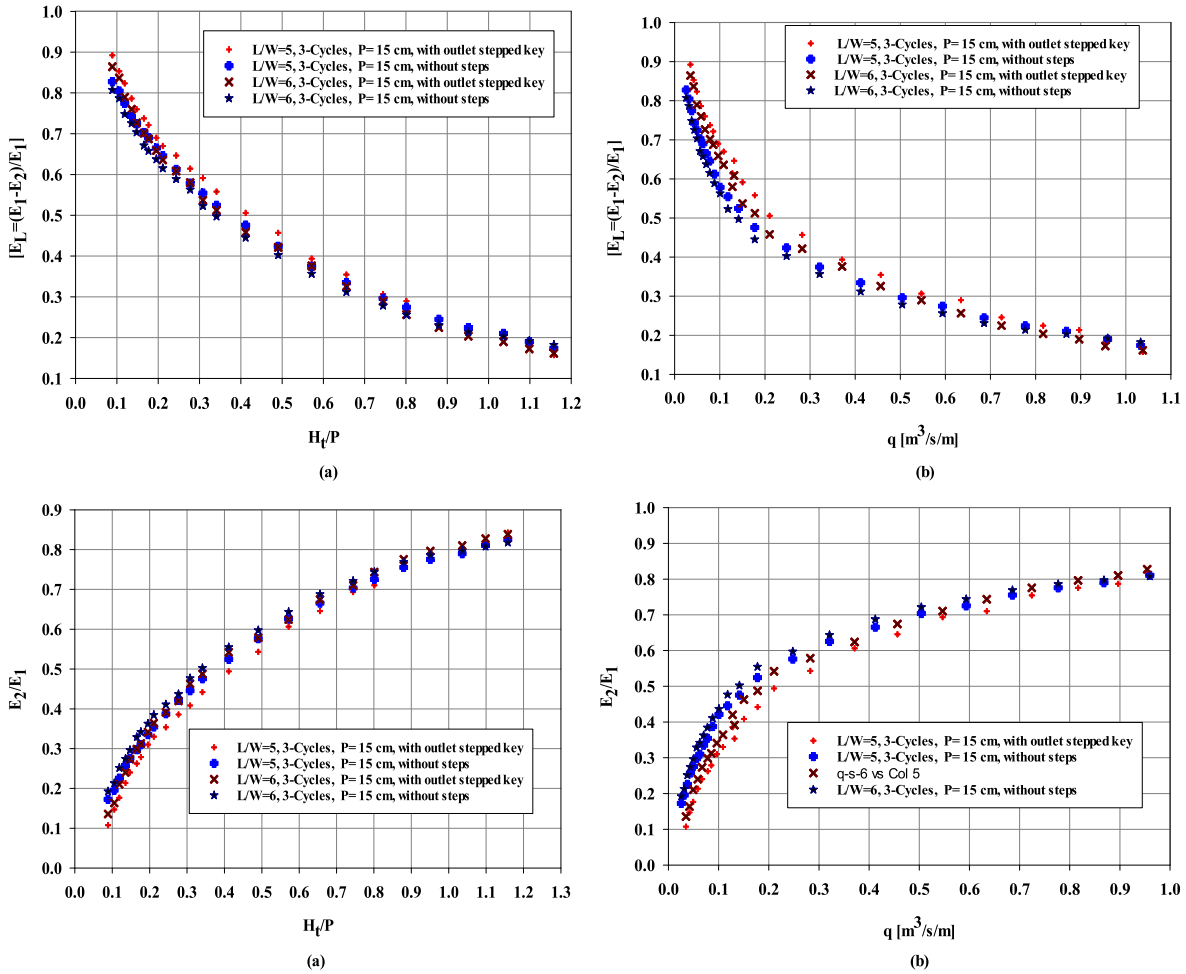


Figure: 10. (i) (Relative energy dissipation $[E_L=(E_1-E_2)/E_1]$ with respect to (a) the head water ratio $[H_t/P]$ and (b) the unit discharge $[q]$. (ii) (Relative residual energy $[E_2/E_1]$ with respect to (a) the headwater ratio $[H_t/P]$ and (b) the unit discharge $[q]$.

complicated energy dissipation mechanisms that reduce the steeped weir's energy dissipation capacity.

As a result, using the steps at the outlet key instead of the slanted floor increases the relative energy dissipation over the PKW for low head discharges while decreasing for higher head discharges. It will help design energy-dissipative structures and reduce the downstream scour, the length of hydraulic jump and apron length for the weir, etc.

5. Conclusions

PKWs are high-performance labyrinth weirs, especially at lower heads. In order to develop a better understanding of energy dissipation at the base of the type-B PKW, 12 laboratory-scale models were tested and analyzed. The results were presented with different hydraulic characteristics, such as the relative energy dissipation and residual energy compared to the energy dissipation downstream of rectangular and trapezoidal labyrinth weir datasets. Furthermore, a sensitivity analysis was also presented by constructing the outlet key steps to understand better how the outlet key's steps influence the energy dissipation of the PKW. Finally, the experimental data sets were used to develop the best relative residual model based on gene expression programming, which practitioners can use as design approximations to help researchers and design engineers in PKW's design. The following important conclusions can be drawn from this research:

1. The relative energy dissipation increases with the number of cycles at low values of H_t/P, and in contrast, for higher H_t/P values intermixing patterns of energy dissipation were observed.
2. The maximum relative energy dissipation was observed corresponding to the highest cyclic count (i.e. $E_L = 0.8621$ or 86.21% the corresponding $N = 5$ and $L/W = 5$) and the less energy dissipation for the lowest cyclic counts (i.e. $E_L = 0.8077$ or 80.77% the corresponding $N = 3$ and $L/W = 6$).
3. Energy dissipation at the base of the PKW shows a decreasing trend with increasing the L/W ratio. Energy dissipation is not linear, and it is most noticeable at low heads.
4. The relative residual energy at the base of PKWs (E_2/E_1) increases with the relative total upstream head (H_t/P), particularly for smaller values of H_t/P, as well as with the magnification ratio (L/W).
5. The maximum relative energy dissipation was observed corresponding to the lowest width ratio (i.e. $E_L = 0.8513$ or 85.13% the corresponding $W_i/W_o = 1$) and the less energy dissipation for the highest width ratio (i.e. $E_L = 0.7720$ or 77.20% the related $W_i/W_o = 1.5$).
6. The energy dissipation at the base of the PKW using the steps at the outlet key instead of the slopping floor increases for low head discharges (the gain is approximately 5.6–6.43%) while decreasing (−2.05 to −1.67%) for higher head discharges.

Overall, the PKWs are very efficient in dissipating energy, mainly when the relatively upstream head is low. Furthermore, the geometry of the PKW has a significant impact on the energy dissipation or relative residual energy at the base of the weirs. As a result, it is critical to design and built efficient hydraulic structures that will increase energy dissipation at the base of the weir and reduce scouring near the structure's downstream. The contents of this paper, which include flow field observations and a dataset, are intended to assist researchers and design engineers in designing PKWs and determining downstream features such as chutes and energy dissipators.

Data availability statement

No data, models, or code were generated or used during the study.

Funding information

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors (No funds were taken from any Organization).

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

Authors express their gratitude and sincere thanks to the Vice-Chancellor of Delhi Technological University, DRC, Chairman of Department, Head of Department of Civil Engineering, and Technical staff of the Fluid Mechanics & Hydraulic Laboratory for their valuable support during the research.

List of Symbols

B	Length of side weir ($B_i + B_b + B_o$)
B_b	Base length
B_i	Length of overhang portions at the inlet side
B_o	Length of overhang portions at the outlet side
E	Specific energy at section 'i'
F_r	Froude Number
g	Acceleration of gravity
H_t	Total energy head
h_t	Piezometric head
h_s	Height of steps
i	Represents the section
L	Total developed crest length
L_{cy}	Crest length of one cycle
l_s	Length of steps
N	Number of cycles
P	PKW height
Q	Discharge over the PKW
S_i	Inlet key slope
S_o	Outlet key slope

T_s	Sidewall thickness of the PKW
V_t	Mean flow velocity
$V_t^2/2g$	Approach velocity head
W	Channel width/Width of PKW
W_{cy}	Width of one cycle
W_i	Inlet key width
W_o	Outlet key width
X	Distance of measurement section from the lateral centreline of the PKWs
ρ	Density of flowing liquid

References

[1] A.-H.K. Al-Shukur, G.H. Al-Khafaji, Experimental study of the hydraulic performance of piano key weir, Int. J. Energy Environ. 9 (2018) 63–70.
[2] R.M. Anderson, B.P. Tullis, Piano key weir: reservoir versus channel application, J. Irrigat. Drain. Eng. 138 (8) (2012) 773–776.
[3] M. Bieri, M. Federspiel, J.-L. Boillat, B. Houdant, L. Faramond, F. Delorme, Energy Dissipation Downstream of piano key weirs—a case study of Gloriettes Dam (France), in: In Labyrinth and Piano Key Weirs—PKW 2011, CRC Press, Leiden, The Netherlands, 2011, pp. 120–130.
[4] H. Chanson, Hydraulics of nappe flow regime above stepped chutes and spillways, Trans. Inst. Eng. Aust. Civ. Eng. 36 (1) (1994) 69–76.
[5] H. Chanson, L. Toombes, Experimental investigations of air entrainment in transition and skimming flows down a stepped chute, Can. J. Civ. Eng. 29 (1) (2002) 145–156.
[6] B.M. Crookston, B.P. Tullis, Labyrinth weirs: nappe interference and local submergence, J. Irrigat. Drain. Eng. 138 (2012) 757–765.
[7] B.M. Crookston, B.P. Tullis, Hydraulic design and analysis of labyrinth weirs. II: nappe aeration, instability, and vibration, J. Irrigat. Drain. Eng. 139 (2013) 371–377.
[8] B.M. Crookston, D. Ph, M. Asce, Hydraulics of labyrinth and piano key weirs : 100 Years of prototype structures, advancements, and future research need 145 (12) (2019) 1–7.
[9] S. Erpicum, V. Nagel, F. Laugier, Piano key weir design of raviege dam, in: Proc. Int. Conf. Labyrinth and Piano Key Weirs Liège B, vols. 43–50, CRC Press, Boca Raton, FL, 2011.
[10] S. Erpicum, F. Laugier, M. Pfister, M. Pirotton, M. Cic, J. Schleiss, Labyrinth and Piano Key Weirs II PKW 2013, 2013.
[11] S. Erpicum, B.P. Tullis, M. Lodomez, P. Archambeau, B.J. Dewals, M. Pirotton, Scale effects in physical piano key weir models, J. Hydraul. Res. 54 (6) (2016) 692–698.
[12] K.R. Eslinger, B.M. Crookston, Energy Dissipation of Type a Piano Key Weirs, Water, MDPI, Basel, Switzerland, 2020.
[13] N. Fen, D.B. Kozlov, I.S. Rumyantsev, Hydraulic studies of stepped spillways of various design, Power Technol. Eng. 49 (2016) 337–344, https://doi.org/10.1007/s10749-016-0625-7.
[14] C.A. Gonzalez, H. Chanson, Turbulence and cavity recirculation in air-water skimming flows on a stepped spillway, J. Hydraul. Res. IAHR 46 (1) (2008) 65–72.
[15] A.H. Haghiabi, M.R. Ghaleh Nou, A. Parsaie, The energy dissipation of flow over the labyrinth weirs, Alex. Eng. J. (2021), https://doi.org/10.1016/j.aej.2021.08.075, 0–4.
[16] M. Ho Ta Khanh, D. Sy Quat, D. Xuan Thuy, PKWs under design and construction in Vietnam (2010), in: Labyrinth and Piano Key Weirs-PKW 2011, CRC Press, London, 2011, pp. 225–232.
[17] M. Ho Ta Khanh, T.C. Hien, N.T. Hai, "Main Results of the PKW Models Test in Vietnam. (2004 to 2010)". Labyrinth and Piano Key Weirs-PKW 2011: 191-198, CRC Press, London, 2011.
[18] S. Jüstrich, M. Pfister, A.J. Schleiss, Mobile riverbed scour downstream of a piano key weir, J. Hydraul. Eng. 142 (11) (2016) 1–12.
[19] M. Karimi, J. Attari, M. Saneie, M.R.J. Ghazizadeh, Side weir flow characteristics: comparison of piano key, labyrinth, and linear types, J. Hydraul. Eng. 144 (12) (2018) 1–13.
[20] S.I. Khassaf, L.J. Aziz, Z.A. Elkatib, Hydraulic behavior of piano key weir type B under free flow conditions, Int. J. Sci. Technol. Res. 4 (8) (2015) 158–163.
[21] M. Leite Ribeiro, J.L. Boillat, A. Schleiss, F. Laugier, C. Albalat, Rehabilitation of St-Marc dam – experimental optimization of a piano key weir, in: Proc. Of 32nd Congress of IAHR., Vince, Italy, 2007.
[22] M. Leite Ribeiro, J.-L. Boillat, A.J. Schleiss, Experimental parametric study for hydraulic design of PKWs, in: B. Liege (Ed.), Proc. Int. Conf. Labyrinth and Piano Key Weirs, CRC Press, Boca Raton, FL, 2011, pp. 183–190.
[23] M. Leite Ribeiro, M. Bieri, J.L. Boillat, A.J. Schleiss, G. Singhal, N. Sharma, Discharge capacity of piano key weirs, J. Hydraul. Eng. 138 (2) (2012) 199–203.
[24] F. Lempérière, A. Ouamane, The Piano Keys weir: a new cost-effective solution for spillways, Int. J. Hydropower Dams 10 (5) (2003) 144–149.
[25] O. Machiels, "Experimental Study of the Hydraulic Behaviour of Piano Key Weirs, PhD Thesis ULgetd-09252012- 224610, University of Liege, 2012 (Belgium).
[26] A.P. Magalhães, M. Lorena, Perdas de Energiado do Escoamento Sobre Soleiras em Labirinto, in: Energy Losses in Flow over Labyrinth Weirs), SILUSBA, Lisboa, Portugal, 1994, pp. 203–211 (In Portuguese).
[27] K.G. Mehta, M.M. Khan, A review of parametric studies of piano key weir, Int. J. Res. Eng. Appl. Sci. 6 (7) (2016) 24–28.
[28] J. Merkel, F. Belzner, M. Gebhardt, C. Thorenz, Energy dissipation downstream of labyrinth weirs, in: Proceedings of the 7th IAHR International Symposium on Hydraulic Structures, Aachen, Germany, 15–18 May 2018, 2018.
[29] M. Noseda, I. Stojnic, M. Pfister, A.J. Schleiss, Upstream erosion and sediment passage at piano key weirs, J. Hydraul. Eng. 145 (8) (2019), 04019029.
[30] A. Noui, A. Ouamane, Study of Optimization of the Piano Key weir." Labyrinth and Piano Key Weirs II, CRC/Balkema, Leiden, Netherlands, 2013, pp. 175–182.
[31] P. Novak, V. Guinot, A. Jeffrey, D.E. Reeve, Hydraulic Modelling: an Introduction: Principles, Methods, and Applications, 2010.
[32] M. Oertel, Sensitivity analysis for discharge coefficients of piano key weirs, 6th International Symposium on Hydraulic Structures: Hydraulic Structures and Water System Management, ISHS 2016 (2016) 540–548, 3330628160.
[33] S. Pagliara, M. Palermo, I. Carnacina, Scour process due to symmetric dam spillways crossing jets, Int. J. River Basin Manag. 9 (2011) 31–42.
[34] M. Palermo, B.M. Crookston, S. Pagliara, Analysis of equilibrium morphologies downstream of a PKW structure, in: Proceedings of the EWRI World Environmental and Water Resource Congress, Las Vegas, NV, USA, 2020, 17–21 May 2020.
[35] A. Parsaie, A.H. Haghiabi, M. Saneie, H. Torabi, Prediction of energy dissipation of flow over stepped spillways using data-driven models, Iran. J. Sci. Technol. Trans. Civil Eng. 42 (1) (2018) 39–53, https://doi.org/10.1007/s40996-017-0060-5.
[36] A. Parsaie, A.H. Haghiabi, Inception point of flow aeration on quarter-circular crested stepped spillway, Flow Meas. Instrum. 69 (August) (2019), 101618, https://doi.org/10.1016/j.flowmeasinst.2019.101618.
[37] A. Parsaie, A.H. Haghiabi, The hydraulic investigation of circular crested stepped spillway, Flow Meas. Instrum. 70 (August) (2019), 101624, https://doi.org/10.1016/j.flowmeasinst.2019.101624.
[38] L. Peyras, P. Royet, G. Degoutte, Ecoulement et Dissipation sur les Déversoirs en Gradins de Gabions, J. La Houille Blanche 1 (1991) 37–47 (in French).
[39] M. Pfister, S. Jüstrich, A.J. Schleiss, Toe-scour formation at piano key weirs, in: Labyrinth and Piano Key Weirs III—PKW 2017, CRC Press, Leiden, The Netherlands, 2017, pp. 147–156.
[40] M. Pinto, Energy dissipation on stepped spillways with a piano key weir : an experimental study (2017) 1–11.
[41] N. Rajaratnam, Skimming flow in stepped spillways, J. Hydraul. Eng. ASCE 116 (4) (1990) 587–591.
[42] A. Silvestri, P. Archambeau, M. Pirotton, B. Dewals, S. Erpicum, Comparative analysis of the energy dissipation on a stepped spillway downstream of a piano key weir, in: S. Erpicum, F. Laugier, M. Pfister, M. Pirotton, G.M. Cicero, A.J. Schleiss (Eds.), in Labyrinth and Piano Key Weirs II, CRC Press/Balkema, Leiden, The Netherlands, 2013, pp. 111–120.
[43] M. Tabbara, J. Chatila, R. Awwad, Computational simulation of flow over stepped spillways, Comput. Struct. 83 (2005) 2215–2224, https://doi.org/10.1016/j.compstruc.2005.04.005.
[44] H. Torabi, A. Parsaie, H. Yonesi, E. Mozafari, Energy dissipation on rough stepped spillways, Iran. J. Sci. Technol. Trans. Civil Eng. 42 (3) (2018) 325–330, https://doi.org/10.1007/s40996-018-0092-5.
[45] H. Truong Chí, S. Huynh Thanh, M. Ho Ta Khanh, Results of some piano keys weir hydraulic model tests in Vietnam, in: Proceedings of the 22nd ICOLD Congress, 2006. Barcelona, Spain, 18–23 June 2006.
[46] B.P. Tullis, D. Ph, M. Asce, B.M. Crookston, D. Ph, M. Asce, N. Young, M. Asce, Scale effects in free-flow nonlinear weir head-discharge relationships (2020).
[47] B.P. Tullis, M. Asce, B.M. Crookston, M. Asce, J. Brislin, S.M. Asce, T. Seamons, Geometric effects on discharge relationships for labyrinth weirs 146 (10) (2020) 1–10.
[48] N.L. Young, DigitalCommons @ USU Size-Scale Effects of Nonlinear Weir Hydraulics, 2018.



Deepak Singh is currently a research scholar of the Civil Engineering Department at Delhi technological University, India. He obtained his B.Tech. (Civil) from UTU and M.Tech. in Hydraulic Engineering GBPUAT, Pantnagar. He has published, 4 papers in refereed international and national journals, and 2 chapters in conference proceedings. His area of research is hydraulic engineering, computational hydraulics, and hydraulic structures.



M. Kumar is currently a Professor of Civil Engineering Department at Delhi technological University, India. He obtained his B.Tech. (Civil) and M.Tech. (Hyd) degree from AMU Aligarh, and a Ph.D. degree from IIT Delhi. He has published about 20 papers in referred national and international journals and conference proceedings. His research areas are applied Fluid Mechanics, surface water quality management, computational hydraulics, and hydraulic structures.

Arabian Journal of Geosciences (2022) 15:102
https://doi.org/10.1007/s12517-021-09385-1

ORIGINAL PAPER

Fireworks induced quasi-ultrafine particle number concentration and size-resolved elemental distribution in megacity Delhi

Shailendra Kumar Yadav[1] · Rajeev Kumar Mishra[1] · Bhola Ram Gurjar[2]

Received: 20 July 2021 / Accepted: 21 December 2021
© Saudi Society for Geosciences 2022

Abstract
Delhi is known for its bad air quality, and the pollution reaches its dangerous levels during the Diwali festival. Different particle size range (quasi-ultrafine, sub-fine, fine, and coarse particles) number concentration was studied during pre, post, and Diwali firework displays in Delhi, India. Simultaneously, a personal cascade impactor was also used to enquire about the elemental size distribution in different size bin particles emitted from firecrackers. The particle number concentration within the size range of qUFP (quasi-ultrafine particle) showed their highest contribution (>95%) toward total particle number concentration (PNC) during the evening hours at the Diwali festival. There was a noticeable increase in average PNC, i.e., 4.09×10^6 #/liter (two times higher than average PNC of morning hours) during the evening hours of Diwali. Regarding elemental distribution in different size bins (qUFP, sub-fine, fine and coarse particles), the highest percentage contribution was found as (32.7%) followed by Pb (23.46%), Cu (23.45%), Al (22.9%), Mg (21.98%), V (20.81%), Cd (19.73%), K (18.11%), and Mn (14.07%) in qUFP. All these elements have high concentrations in comparison to national and international standards on Diwali day celebrations. These findings may be used by policymakers/competent authorities to analyze the effectiveness of framed festival rules at ground level and take appropriate actions to reduce firecracker-induced air pollution and associate health risks of people.

Keywords Diwali · Elemental concentration · Firecrackers · PNC · qUFP

Introduction

Delhi is ranked as the world's second most polluted capital city (WHO 2018; Guttikunda et al. 2019). Most of India's population was exposed to household air pollution in 2017 (about 60% of the total population), and 3 million people died due to $PM_{2.5}$ pollution (Health Effects Institute and Project 2019). As India's population grew, it added the most burden, with approximately 21.3 million DALYs (disability-adjusted life years) and 0.673 million deaths, both of which were attributable to $PM_{2.5}$ (Balakrishnan et al. 2019). The WHO's most lenient air quality standard (e.g., 25 µg/m^3- 24 h mean for $PM_{2.5}$) is not even met in most Indian metropolitan areas. Transportation, residential energy usage, industries, power plants, dust resuspension, and waste burnings are all important sources of criteria pollutants (PM_s, SO_x, NO_x, CO, O_3, Pb) in an urban area (Viana et al. 2008; Thurston et al. 2011; Karagulian et al. 2015; Amato et al. 2016). As a result of urbanization, health-related attention has shifted to the short- and long-term effects of air pollution exposure in recent years. Air pollution exposure and dispersion are heavily influenced by urban planning and geography, as Delhi is known as a "Landlock Area" in terms of air pollution dispersion because airborne pollutants stay in the air for a longer period during the winter due to low temperatures, low wind movements, and high relative humidity in late October to late January. In addition, Diwali, a major festival in India, is celebrated on a grand scale around the same time (October to November). The burning of stubble in Haryana and Punjab is a significant precursor to the arrival of Delhi's airborne pollution in November (Jain et al. 2014). However, the magnitude of their contribution to Delhi air quality is still being monitored.

Responsible Editor: Amjad Kallel

✉ Rajeev Kumar Mishra
rajeevkumarmishra@dtu.ac.in

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

[2] Department of Civil Engineering, IIT Roorkee, Roorkee, Uttarakhand 247667, India

Diwali (India), Lunar New Year (China), Bastile Day (France), Australia Day (Australia), New Year's Eve (Globally), Fourth of July/Independence Day (USA), and other significant events all involve the use of firecrackers/bonfires. The use of fireworks releases a variety of pollutants into the atmosphere, such as particulate matter and toxic gases, which affect human health directly or indirectly more in urban areas than in rural areas (Ravindra et al. 2003; Kulshrestha et al. 2004; Seidel and Birnbaum 2015; Hamad et al. 2016). During sudden spear of fireworks and bonfire events, severe health effects such as asthmatic attacks, lung infection, eye allergies, fever, cough, upset, and cardiovascular disease can occur (Hirai et al. 2000; Becker et al. 2000; Barman et al. 2008; Beig et al. 2013; Künzli 2015; Hu et al. 2017; Habre et al. 2018). During fireworks events worldwide, the concentrations of air pollutants (particulates matter, heavy metals, and oxides of gases) increase by a factor of 2–10 (Barman et al. 2008; Kong et al. 2015; Seidel and Birnbaum 2015). In general, the sharpest skewer concentration can be seen during the event, with levels returning to normal after 24 h. During Diwali, the highest concentration is usually observed after evening hours (20:00–01:00 h) (Kumar et al. 2016). During fireworks events, concentrations range from a few μg/m^3 to a maximum of 2000 μg/m^3 for short periods in different cities (Singh et al. 2019).

Regarding emission rates of PNC, whistling sparkles show the highest PM emission rates (3.24×10^{10} particles per gram of burnt sparkler than 1.68×10^{10} particles per gram of sparkler burnt for low smoke sparklers) and also get quickly deposited in the respiratory system (Betha and Balasubramanian 2013). To date, only a few studies have been conducted on particles of lower size (ultrafine and quasi-ultrafine particles), particle mass, and PNC (Apte et al. 2011; Saraswat et al. 2013; Gani et al. 2018, 2020; Singh et al. 2019; Arub et al. 2020a; Yadav et al. 2021). For Delhi, Kumar et al. (2011) estimate the total PNC for vehicular emissions from 2010 to 2030 based on business as usual (BAU) and best estimate scenario (BES) as 1.37×10^{25} for 2010, which are expected to increase by four times in 2030-BAU but decrease by 18 times in 2030-BES due to fuel replacement for sustainability. Saraswat et al. (2013) calculated the PNC (10 to 1000 nm) 4×10^3 #/cm^3 with a geometric mean particle diameter of 43 nm in Delhi using a spatiotemporal land-use regression model for urban background PNC. However, PNC increased ~7-fold (280×10^3 #/cm^3) in the road microenvironment of Delhi (Apte et al. 2011). Apte et al. (2011) reported results from ~180 h of real-time ultrafine particle number concentration measurements inside a common vehicle, the auto-rickshaw, in New Delhi, India. Long-term PNC monitoring in Delhi and mass conversion with chemical characterization are covered in detail by Gani et al. (2019), Hagan et al. (2019), Arub et al. (2020b) and Bhandari et al. (2020). A study found a higher PNC (10 to 350 nm) concentration in 2018 than in 2017 during autumn (mid-September to November) in Delhi (Patel et al. 2021). More information on the formation, growth rate, and classification of ultrafine (few nanometers to several nanometers size range) particles in Delhi can be found in the study conducted by Gani et al. (2020), Kanawade et al. (2020) and Yadav et al. (2021). Diwali celebration is a significant source of short-term air pollution in Delhi, so there is a need to investigate the field of PNC of lower size particles, which have a higher potential for health effects (Schraufnagel 2020).

Quasi ultrafine Particle (qUFP) is 0.25μm > dp < 0.50 μm range (Saffari et al. 2013). During such events as Diwali (firecracker celebration), PNC has increased (Drewnick et al. 2006; Wehner et al. 2000). Betha and Balasubramanian (2013) found that fireworks emit 60 to 85% ultrafine size range (below 100 nm) particles (Betha and Balasubramanian 2013). In China, during the spring festival season, the total PNC was found 42% and 57% with Aitken and accumulation mode, respectively, during firework emission (Yang et al. 2014). The average PNC of 2.5×10^4 cm^{-3} was reported during Spring festival fireworks. Wehner et al. (2000) detailed the PNC increase in the accumulation range (less than 10 μm) in the events of firework (Wehner et al. 2000). In India, Sarkar et al. (2010) found exposure to certain elements associated with firework emissions (Al, Mn, Ba) increased health effects by at least 2% during Diwali. In contrast, no such risk difference was found in Bangkok due to metal elements such as Pb, Zn, and Cr (Sarkar et al. 2010). It is unclear whether there are long-term effects along with short-term effects or not (Singh et al. 2019). Ultrafine particles and Quasi ultrafine particles have easy access to get into the bronchi and alveoli of the human respiratory system and deposit in the areas of the respiratory tract (Oberdörster et al. 2005, 2004; Dockery and Pope 1994; International Commission on Radiological Protection 1994; Oberdörster et al. 2005). The processes through which the inhaled particles get deposited in the human respiratory system are diffusion, gravitational settling, interception, and inertial compaction (Dockery and Pope 1994; Oberdörster et al. 2004, 2005; Chen et al. 2016).

The firecracker used during the Diwali festival generally consists of paper tubes filled with organic, non-metals, and metals elements (charcoal, sulfur, carbon, silicon, aluminum), oxidizing agents (nitrate, chlorates), reducing agents (sulfur, charcoal), coloring agents like strontium, copper, lead, barium, and binders (Drewnick et al. 2006; Vecchi et al. 2008). It causes pollutants' emissions, including sulfur, charcoal compounds, trace gases, trace metals, and particulate matter (Ravindra et al. 2003; Kulshrestha et al. 2004; Wang et al. 2007, 2012). Potassium (K) is a significant component during fireworks emission (Drewnick et al. 2006; Wang et al. 2007). The particulate matter emits from

fireworks, a rich source of metals like Lead (Pb), Barium (Ba), and Manganese (Mn), with some organic compounds like polycyclic aromatic hydrocarbons (PAH). Barman et al. (2008) outlined the increase in the concentration of Copper (Cu), Nickel (Ni), Cadmium (Cd), Zinc (Zn), and Chromium (Cr) during the Diwali celebration in Lucknow, India (Barman et al. 2008). There was found a significant increase in the concentration of Ba, i.e., 109 times more than the day before Diwali, along with other elements like K, Sr, Cu, Mg, Ni, V, and As with a notable increase in barium, strontium, aluminum, and potassium (Kulshrestha et al. 2004).

Because of the impact of air pollution caused by excessive bursting of firecrackers in the national capital and crop burning in neighboring states during Diwali in 2016, the Supreme Court of India issued a notification restricting the sale and bursting of firecrackers. Even though the sale of fireworks was restricted in Delhi in 2018 to reduce Diwali air pollution, no significant improvement in air quality was observed (Trivedi 2018). The ruling did not contribute to the desired outcomes because ambient air quality data comparison from Diwali 2017 revealed severe indications. Studies from various countries on how fireworks affect air quality are presented in the literature (Singh et al. 2019 and references herewith). According to the current evidence, fireworks can cause a significant increase in PNC, and future research should now focus on improving this scenario. There are no research articles on PNC and its size distribution concerning Delhi Diwali firecracker emissions, whereas the mass concentration of PMs (PM_{10}, $PM_{2.5}$, PM_1, and their chemical characterization) is numerous. Smaller particles cause more health problems than larger particles due to their high lung penetration and deposition capacity (Terzano et al. 2010; Fuller et al. 2013, 2018; Chen et al. 2016; Lane et al. 2016; Schraufnagel 2020; Brugge and Fuller 2021). These factors necessitate the current study to assess the PNC status and chemical characterization in Delhi during Diwali. This study aims to evaluate the impact of the episodic event Diwali on-air pollution-prone metropolitan cities like Delhi in terms of PNC trend and metal concentration in different Quasi-particle (qUFP) distributions in 2018.

Materials and methods

Site description

The experiment was carried out in New Delhi (28°36′36″N, 77°13′48″E), at an altitude of about 216 m above mean sea level and covering an area of 1484 km^2. The metropolis is bounded by the Himalayas to the north, the Thar Desert to the west in Rajasthan, and the south's central Indian plains known as the land lock area. Westerly or North-Westerly winds dominate Delhi during the months of Diwali (As per the Hindu calendar, Diwali is celebrated on Amavasya (or new moon)-the 15th day of the month of Kartik, every year, which comes in October and November of the Gregorian calendar.). The temperature begins to fall in the city starting in October. In November, lower temperatures are combined with foggy conditions that reduce visibility and air masses that contain many pollutants that pass through Delhi during this season. Lower temperature can be attributed to the low mixing heights observed in these months that render mixing pollutants with the upper boundary layer difficult (Tiwari et al. 2012). The monitoring of PMs was made at 'Hari Nagar- Maya Enclave Building' (urban area of West Delhi Lok Sabha constituency), surrounded by markets and residential neighborhoods at the height of ~14.5 m on the roof of a building (See Fig. S1). Solar radiation, wind speed, wind direction, and relative humidity were collected from the nearest monitoring station (Mandir Marg) of the Central Pollution Control Board (CPCB), Central Control Room (CCR) web interface. The experimental location receives considerable emissions from the road microenvironment also. However, the traffic during the celebration day of Diwali dropped drastically (>60%). Still, meanwhile, traffic congestion has taken the joy out of shopping for Diwali as Delhiites spend a considerable time on the roads during the last two evenings before Diwali. According to traffic police, commuters remain stuck in snarls in many parts of the city due to a 50% increase in traffic volume on the last evening of Diwali day (TOI 2014).

The experiment was carried out 24 h from 30th October to 13th November 2018 (total 15 days), including celebratory events (7th November 2018) to investigate the temporal effect of the pre-and post-Diwali emission. But for chemical characteristics, the monitoring was conducted only on Diwali day for different size bins, when relatively moderate temperatures as high as 18 °C, lower relative humidity (~44%), and no rainfall conditions prevailed over the study region. The mean (± standard deviation) air temperature, relative humidity, and wind speed were 18.25 ± 6.07 °C, 63.40 ± 18.88%, and 0.75 ± 0.59 m s^{-1}, respectively, mainly suggesting calm weather conditions.

Instrument and campaign

Particle number concentrations of different size bins are observed in this experiment. The chemical characterization of some fireworks-related elements has also been conducted on Diwali day. For this study's monitoring of PNC, GRIMM™ Portable Aerosol Spectrometer (1.109) is used. It works on the principle of light scattering at single particles with diode laser; detection volume is aerodynamically focused. This instrument's features are real-time monitoring of particle number distribution. This instrument has 31 equidistant size bins for particles ranging in size from 0.25

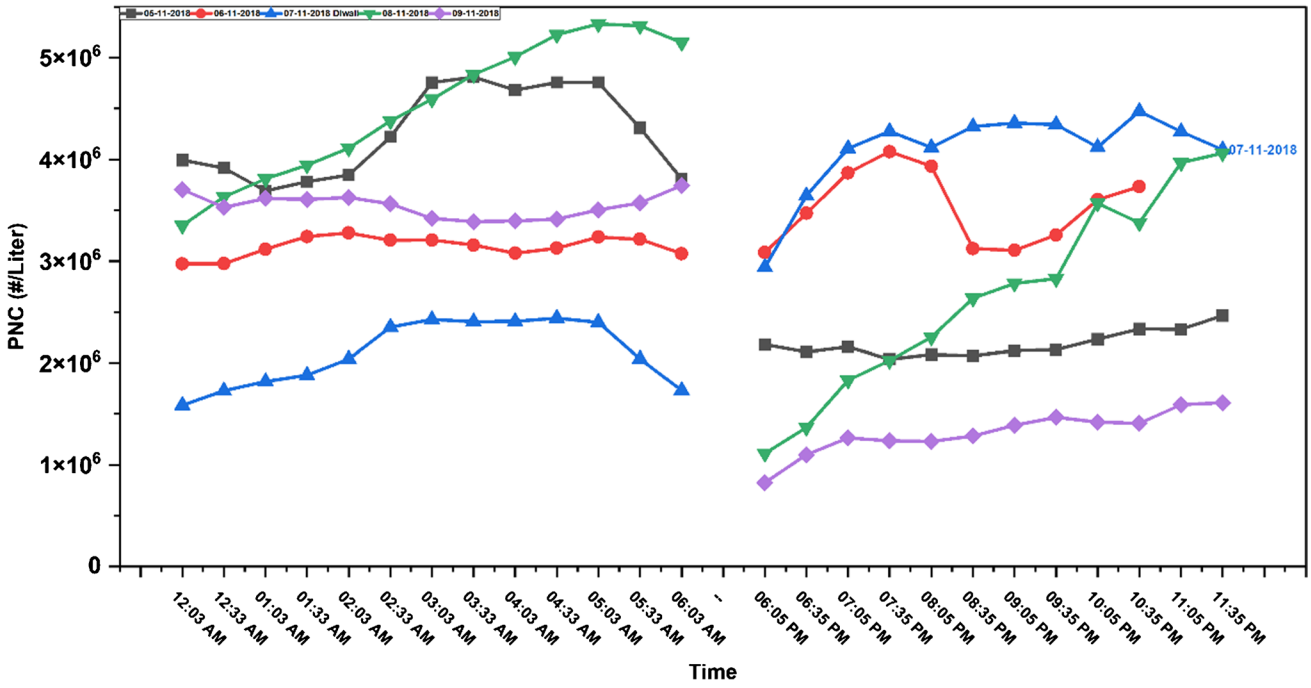


Fig. 1 Total particle number concentration of 'AM' and 'PM' time of selected pre-, post-, and Diwali day (7th Nov-2018)

to 32.0 µm. In this study, 31 size bins or channels are classified as Quasi-Ultrafine (qUFP; 0.25 > Dp < 0.50 µm), sub-fine particles (0.50 > Dp < 1.0 µm), fine particles (1.0 > Dp < 2.5 µm), coarse particles (2.5 > Dp < 10.0 µm). This study excludes > 10.0 µm particle size from the study because their contribution is minimal and beyond the scope of this study. Typically, PNC for UFPs (0.10 µm) is measured in #/cm^3. Still, #/L is used here as a better graphical and statistical presentation because larger size bins (fine and coarse particles) contribute less concentration than qUFP and sub-fine particles.

For chemical characterization of selected elements (Al, V, Mn, Mg, Cu, As, Cd, Pb, and K) at each stage of SKC™, Sioutas Cascade Impactor has also been used in this monitoring. Cascade impactor consists of four impaction stages, known as 'A' (2.5 µm > Dp < 10.0 µm), 'B' (1.0 µm > Dp < 2.5 µm), 'C' (0.50 µm > Dp < 1.0 µm), and 'D' (0.25 µm > Dp < 0.50 µm) are classified in this study as coarse particles, fine particles, sub-fine particles, and qUFP, respectively. It is used with the Leland Legacy Sample Pump to maintain a constant flow rate of 9 L/min. Particles or Dust above each cut-point are collected on a 25-mm PTFE filter in the appropriate stage ('A,' 'B,' 'C,' 'D'). More details can be found in Joshi et al. (2019) and Rohra et al. (2018). The samples were collected from the impactor on 25 mm diameter, 2mm pore size pre-weighted PTFE filter paper (the mass of the filter paper was determined using a four-digit microbalance) with the necessary blanks. Blank tests were also performed during the monitoring by exposing them to the ambient air when the sampling box was opened to remove or replace filter paper from the impactor. These blank values were typically found to be very low. The filters were handled only with tweezers coated with Teflon tapes to minimize the risk of contamination. These collected filter papers were then refrigerated at 4°C. This is essentially done to limit the loss of semi-volatile components. The collected filters were conditioned for 48 h at a stable temperature of 21°C and relative humidity of 35% before the pre-sample and post-sample weighing. The collected filters were extracted by digesting them in 8 ml of analytical grade HNO_3 and kept at a temperature of 40–60°C for 90 min. To avoid impurities and contamination, it was further diluted up to 50 ml with distilled de-ionized water and stored in polypropylene bottles before storing overnight. It was followed by analysis using coupled plasma mass spectrometry (ICP-MS) and Atomic Absorption spectrometry (AAS) techniques for the concentration of the metal (Habil et al. 2016).

Estimation of particle number size distribution and mass weight of elements

We have derived the modal parameters of the mean particle number size distributions during the event and non-event periods by fitting the observed size distributions to the analytical log-normal size distribution of the following form (Kompalli et al. 2018, 2020):

$$\frac{dN}{dlnD_p} = \sum_{j=1}^{n} \frac{N_k}{\sqrt{2\pi}\ln\sigma_{m,j}} \exp\left[-\frac{\left(\ln D_p - \ln D_{m,p,j}\right)^2}{2\ln^2\sigma_{m,k}}\right] \quad (1)$$

(Where D_p is the particle diameter, n is the total number of the modes considered, $D_{m,p,j}$ is the mode diameter of the j^{th} mode, $\sigma_{m,j}$ is the corresponding width of the mode, and N_j is the particle number concentration (amplitude) of the j^{th} mode).

Mass estimation of particulate matters on filter paper of cascade impactor follows the below-mentioned formula.

Calculate the total volume of air sampled using the following equation:

$$V = Q_a \times t \quad (2)$$

Where V = volume of air sampled; Qa = average sampler flow rate across the sampling time; t = total elapsed sampling time in minutes.

And for estimating the mass of particulate matter of each stage of cascade impactor using the following equation:

$$Mass = {W_f - W_i}/{V} \quad (3)$$

Where W_f = final weight of exposed filter; W_i = initial weight of unexposed filter; V = volume of air sampled.

Statistical tests

The dataset normality was tested using a combination of the Exploratory Data Analysis (EDA) plot and the Shapiro-Wilk (S-W) test. Statistical methods are used to examine whether statistically significant variations are found in the analyzed data or not. The Kruskal-Wallis (K-W) test was primarily used to analyze the variance equality between data collected during monitoring days. In particular, the K-W test examines whether the null hypothesis is true that the population variances are equal. In both instances, a 0.05 significance level was used. Therefore, if the p-value was smaller than the threshold of significance (<0.05), the null hypothesis was dismissed, and there was a statistically significant difference in the variance of the results.

Results and discussions

Time series analysis of particle number concentration

Diurnal variation of total particle number concentration (PNC) was carried out during the entire 15 days (30th Oct to 6th Nov 2018, Pre-Diwali, 7th Nov 2018 on Diwali day, 8th Nov to 13th Nov 2018 post-Diwali) for a continuous duration of 24 h (see in Fig. S 2a, b, c). This graph represents the temporal variation of particle number concentration (>0.25 to 10 µm) during the entire monitoring period (See Fig. S 2b). The effect of event episodes due to fire bursting can be easily observed in this graph. Because other days do not show significant variation in Diwali emission in Delhi 2018, the date from 5 to 9 November (1:00 am to 6:00 am and 6:00 pm to 12:00 pm) has been chosen for further insightful evaluation as depicted in Fig. 1 (for more detail, see in the Fig. S 2c). Since firecrackers bursting and lighting are the main components of the Diwali festival in India, it is celebrated from evening to late at night. It can be seen in Fig. 1 that the event celebration affects air quality even till the morning of the next day, in which the meteorological conditions also played a very significant role (Singh et al. 2010). Although, pre-Diwali activity during the evening period increases manifold the traffic emissions (TOI 2014). Therefore, one day before pre-Diwali (on November 6th), during 'PM' period (6:00 pm to 8:30 pm), a higher concentration of particulate matter was observed due to shopping and bursting firecrackers on Choti Diwali. However, during late-night, PNC falls and continues to fall till early morning of 7th November. Interestingly, on 5th November as 'Dhanteras shopping'-one day before Diwali at evening time, famous for traffic jams till late at night (India.com 2018), the 'PM' period does not show a higher or similar trend of PNC like on 6th November, the 'PM' period. The reason behind it may be the firecracker bursting frequency and amount, which is generally lower at 'Danteras' evening compared to Choti Diwali. On the other hand, traffic emissions increase the PNC (60 to 80%) in the range of ultrafine particles (<.1 µm) in the total PNC of the road-microenvironment (Harris and Maricq 2001; Kittelson et al. 2004; Knibbs et al. 2011; Kumar et al. 2014; Strasser et al. 2018; Rönkkö and Timonen 2019; Yadav et al. 2019; Lv et al. 2020; Brugge and Fuller 2021; Zheng et al. 2021). Therefore, being traffic jam, 5th November's evening has less PNC than on 6th November, as this study cover PNC > 0.25 µm. On Diwali, i.e., 7th November 2018, the PNC is higher than on other days at 'PM' period (6:00 pm to 12:00 pm), while it is lower at 'AM' period (1:00 am to 6:00 am).

On the Diwali morning period (12:03 am to 06:03 am), the average particle number concentration was 2.10×10^6 #/liter and maximum and minimum 2.44×10^6 #/liter and 1.58×10^6 #/liter, respectively (See in Table S1 and Fig. 2). But during the 'AM' period, the coefficient of variance (15%) was more than the evening period (6:03 pm to 11:30 pm) coefficient of variance (9.80%) with an average PNC of 4.09×10^6 #/liter along with maximum and minimum PNC of 4.47×10^6 #/liter and 2.94×10^6 #/liter, respectively. The mean of PNC at evening time was approximately double compared with morning time on Diwali, which is similar to the study of Chatterjee et al. (2013) for PMs (mass

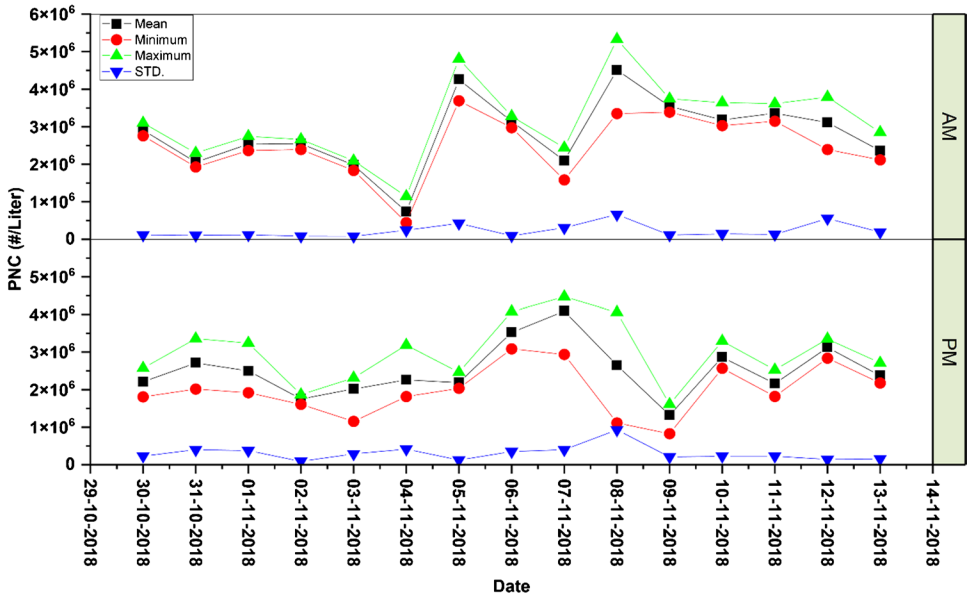


Fig. 2 Statistical graphical representation of total PNC for 15 days

concentration). Diwali day has a high PNC in the evening compared to another day because the evening is the most appropriate time for fire busting and lightning during the Diwali celebration.

Generally, in North India, youth burst firecrackers even before and after 2 or 3 days of Diwali (Thomas 2018). It can be seen in Figs. 1 and 2 that during these days, the average PNC of the period ('AM,' 'PM') is much higher than the other days. The average PNC for 30th October to 5th November is 7.34×10^5 to 4.26×10^6 #/liter, 1.7×10^6 to 2.19×10^6 #/liter for 'AM' and 'PM' periods, respectively. But average PNC shows a significant increment during the firecracker celebration period ('PM' period) as 3.53×10^6 #/liter for 'Choti Diwali (6th November) and 4.09×10^6 #/liter for Diwali day (7th November), which presents the effect of fire bursting because post-Diwali (8th to 13th November) average PNC is less (1.32×10^6 to 3.13×10^6 #/liter for 'PM' period (see more detail in Table S 2). However, wind speed has also affected the particle concentration as the wind speeds from 6th to 7th November was much lower than earlier (Fig. 3 and Fig. S 3 A, B, C). Primarily wind speed, and relative humidity played a crucial role in the concentration distribution of PNC (Grundström et al. 2015; Rivas et al. 2017; Zhou et al. 2018). Polar plots (Fig. S 3 A) show that the wind direction was south-southwest during all 5 days of monitoring, with a higher concentration of PNC with low wind speed and a lower concentration with high wind speed. Polar plots depicted the variation in particle size dynamics caused by the bursting of firecrackers and the planetary boundary condition during the 'PM' and 'AM' periods, respectively (Fig. S 3 B and C). With low wind speed, Diwali and Choti Diwali have a higher concentration of smaller particles. However, on other days, such as on 5th November 'AM' Period, a higher PNC was observed than the average Diwali day in which the meteorological condition played (planetary boundary layer condition) a significant role (Kompalli et al. 2014, 2018). Figure 3 summarizes the meteorological data (SR, Tem., RH and WS) of pre, post, and Diwali day. It is depicted from Fig. 3 that low wind speed (7th to 10th November) and low mixing height favor the relatively high particle concentrations on Diwali day and post-Diwali day. On the very next day of Diwali (8th November), during 'AM' period's PNC was found comparatively higher than the Diwali day 'AM' period's PNC (Table S 1 and 2). It may be due to lower temperature and wind speed by which the pollutant did not get sufficient time and space to get dispersed and not get significantly diluted into the ambient air. Instead, it remains in ambient air and gets mixed with the pollutant emitted from the other sources. Here, the maximum concentration may be due to two major regions, i.e., the particle emitted from burning firecrackers and the day's moderate wind speed.

Low mixing height and low wind speed directly correlate with high particle concentrations (Singh et al. 2010; Chatterjee et al. 2013; Zhou et al. 2018). Particle number increases due to firecrackers' burning and meteorological conditions (Yadav et al. 2019a). It gives a better view of the applicability and efficiency of the rule set out by the Supreme Court of India on Diwali and the traditional display of fireworks. The entire temporal distribution of particle number concentration on Diwali's day is depicted in Fig. 1 and Table S 1 and 2. On Diwali (see Fig. 1), during the first 1 h of sampling (6:50 pm to 7:05 pm), the average peak value is 4.27×10^6 #/liter recorded. A similar pattern was observed with a peak of 4.47×10^6 #/liter, intensified gradually between 10:05 p.m. to

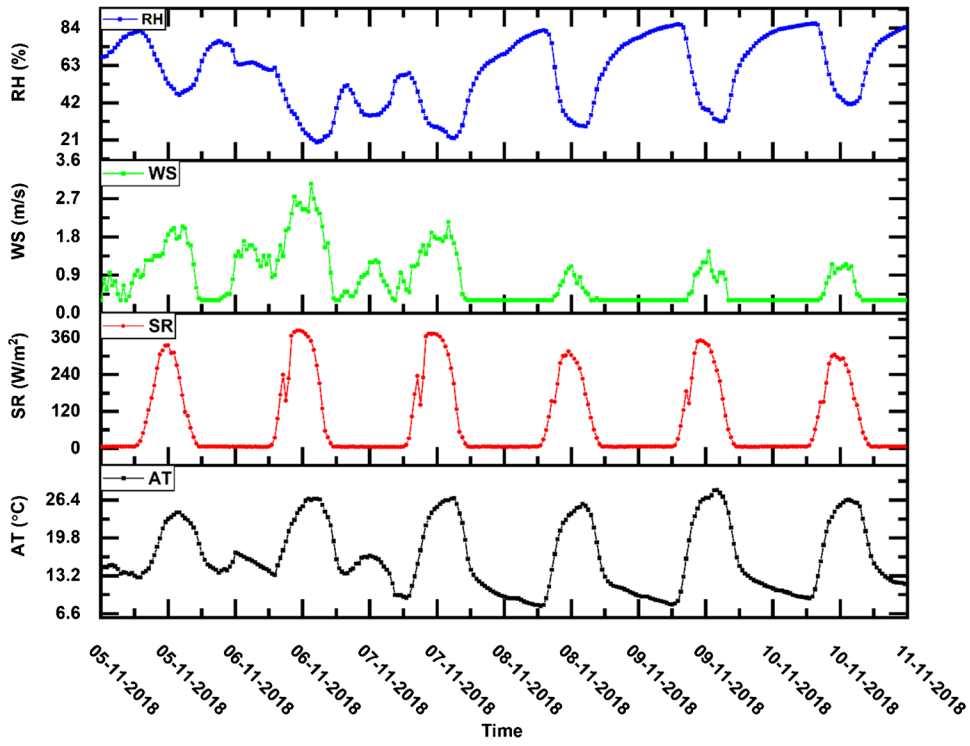


Fig. 3 Temporal variation (24 h of each day) of wind speed and solar radiation (30 min average)

10:35 p.m., and again falls up to ~3.35 × 10^6 #/liter toward the beginning of the next day.

Again, PNC escalates exponentially with a peak of 5.33 × 10^6 #/liter toward the sampling pre-sunshine, but after post-sunshine, it reduces. In concern of fire bursting rule 2018, the public started crackers bursting before 8:00 p.m., not following the provision made by India's government very strictly (Trivedi 2018). Although the pattern of particle number concentration of 6th November was found similar to the pattern of Diwali day till 8:00 p.m., only a slight difference was observed in the concentration.

Particle number size distribution (PNSD)

Total PNC categorized into 5 particle size bins; classified as Quasi-Ultrafine (qUFP; 0.25 µm > Dp < 0.50 µm), sub-fine particles (0.50 µm > Dp < 1.0 µm), fine particles (1.0 µm > Dp < 2.5 µm), coarse particles (2.5 µm > Dp < 10.0 µm). According to the literature, conceptually, the high-size particles are less significant in contributing to particles' number concentration (Yadav et al. 2019).

It is clear from the SI 4 figure that the individual contribution of a larger particle size such as fine and coarse particles size bin has less than 1% (regarding total PNC). Therefore, to make the graph more understandable, these two-size bins have been omitted from SI 5 (A to E) and Table 1. In concern of qUFP, Table 1 shows two sizes bins; 1–0.25 to 0.28 µm; 2–0.28 µm to 0.50 µm as lower qUFP (0.25 µm > dm < 0.28 µm) have significant contribution in PNC with low size range compared to higher qUFP (0.28 µm > dm < 0.50 µm).

It shows that qUFP consists of >95 % PNC of total PNC, whereas lower qUFP's participation of total PNC is found as 16.0 to 26.0 % and 17.0 to 30.0 % at 'AM' and 'PM' period, respectively, and higher qUFP shows 72.0 to 81.0 % and 69.0 to 81.0 % at 'AM' and 'PM' period, respectively. Diwali day's PNC was found higher during the 'AM' and 'PM' among all the monitored days. Although concerning sub-fine, the 'PM' period duration (3.0%) on Diwali showed higher PNC than the other day's 'PM' period like (2.9%) of November 6th and the 'AM' period (3.4%) of November 8th, respectively. This change indicates the effect of fire bursting on particle number concentration of Pre Diwali, Diwali, and Post Diwali (Table 1). Joshi et al. (2016) have also found a higher total PNC than background concentration on Diwali (Mumbai). According to Joshi et al. (2016), no conclusive inferences (an expected increase of number concentration in 0.30 to 0.50 µm) could be drawn from the size distribution data measured by OPC. Most of the emissions due to firework bursting are expected to be fine and ultrafine sizes (Yadav et al. 2019a). According to Joshi et al. (2016) and Yadav et al. (2019), it is not possible to distinguish an episodic event effectively (like Diwali) through the particle size distribution or dynamics. Because the peak of PNC in a different mode (nucleation, Aitken mode, small Aitken mode, accumulation mode) is not similar, like the

Table 1 Percentage contribution of PNC in different size bin

Date	Percentage contribution of different size bin in PNC					
	qUFP				Subfine	
	0.250 to 0.280 µm		0.280 to 0.500 µm		0.500 to 1.00 µm	
	AM	PM	AM	PM	AM	PM
5th Nov. 2018	16.2%	19.5%	80.4%	77.9%	3.2%	2.4%
6th Nov. 2018	16.6%	17.3%	80.5%	79.6%	2.7%	2.9%
7th Nov. 2018 Diwali Day	25.4%	16.2%	72.6%	80.5%	1.7%	3%
8th Nov. 2018	12.4%	25%	84%	71.8%	3.4%	2.1%
9th Nov. 2018	16.4%	29%	81.1%	68.1%	2.4%	2.2%
'AM' period—12:03:01 am to 06:03:01 am			'PM' period—06:05:01 pm to 11:35:01 pm			

pre- and post-Diwali's PNC peak, to make any definite conclusion about the mode shifting. But in the higher size range (qUFP), the peak of PNC is not shifting into other modes (different size bin), although only showing the increment in particle number concentration, which can be seen in the table or graph (See SI 6 (5th November to 9th November 2018)).

To get more insight into size distribution, particle number size distribution is presented as supplementary information (SI-6) that shows that less than 0.500-µm diameter particle is higher in the total PNC. Except for particle concentration, no significant difference is observed in the trend of particle concentration between 'AM' and 'PM' periods. Since PNC (qUFP) started increasing from 6:30 pm on Diwali day, people did not strictly follow the Supreme Court rule and started bursting firecrackers shortly before 8:00 pm. It leads the concentration growth significantly and starts to affect human health (Trivedi 2018). Several studies have reported that small particles are more harmful than large particles and have many effects on human health, such as respiratory, cardiovascular, inflammatory, etc. (Chen et al. 2016; Rohra et al. 2018).

Percentage distribution of elements in different particle size bin

During the Diwali celebration, the emission of firecrackers contains heavy metals in higher concentrations than other days. This study includes analyzing nine elements (Al, V, Mn, Mg, Cu, As, Cd, Pb, and K) as lower-sized diameter particulate matter concentrations. Particles belonging to UFP and qUFP range cannot be filtered out by general mechanisms of the human body, i.e., nasal hairs or mucous membranes, and end up passing the throat to the lungs, posing a wide range of health hazards (Chen et al. 2016) (Table 2).

Concerning the toxicity on human health, these element concentrations correspond to qUFP, sub-fine, fine, and coarse, which were evaluated further in this section as lower-sized particles with high penetration efficiency in the human lung.

As earlier discussed in the instrumentation section, the SKC™ cascade impactor has four stages, collecting the particulate matters simultaneously on each stage. In that condition, after taking all stages (A, B, C, D) collectively, 33%, 23%, 22%, and 21% particulate matter contributed by qUFP, sub-fine, fine, and coarse particle size range, respectively.

Table 2 Distribution of element concentration in each particle bin

Particle size/metal	Distribution of element concentration in each particle bin								
	Aluminum (ppm)	Vanadium (ppm)	Manganese (ppm)	Copper (ppm)	Arsenic (ppm)	Cadmium (ppm)	Lead (ppm)	Magnesium (ppm)	Potassium (ppm)
qUFP	155.43	0.20	3.45	3.09	0.45	0.16	8.05	1.16	3.44
Subfine	100.25	0.10	1.99	1.57	0.17	0.09	4.88	0.96	4.54
Fine	253.47	0.14	7.35	4.36	0.41	0.45	16.82	1.15	7.95
Coarse	194.45	0.54	11.71	4.16	0.34	0.10	4.55	2.02	3.05
Total	703.60	0.98	24.49	13.18	1.38	0.80	34.30	5.30	18.98
Average	175.90	0.25	6.12	3.29	0.34	0.20	8.57	1.32	4.74

Interestingly, it reveals that lower size particulate matter has a higher percentage contribution than larger size particulate matter on Diwali day (SI 7). The chemical characterization of each stage shows the concentration of Al > Pb > Mn > K > Cu > Mg > As > V > Cd in average concentration because burning of firecrackers produce heavy metals emission (Liu et al. 1997; Dutcher et al. 1999; Gaeggeler et al. 2008; Shi et al. 2014; Li et al. 2017). Aluminum is found with the highest concentrations among all elements as it is generally used in most firecrackers (Average 703.6 ppm). Kulshrestha et al. 2004 and Drewnick et al. 2006 reported a high concentration of Al and K during Diwali. The metal does not have ambient air standards set by WHO but does have national standards set by US EPA. According to these standards, the Al level should not exceed 10 mg/m^3 for an 8-h interval. This roughly translates to 9.4 ppm for an 8-h interval. Thus, 112.8 ppm gravimetric collection for 48 h, and an average below 2.3 ppm. The observations are ~7 times higher than the standard for the ideal situation, which is very harmful. Pb in firecrackers is banned ("Supreme Court bans the use of heavy metals in firecrackers to prevent air pollution - Education Today News," 2017). Thus, it is a relevant observation that Lead is over 34 ppm, way beyond the acceptable levels in potable water itself, which is 1.5 ppb. In ideal conditions of 25°C and 1 atm, the acceptable level of Lead (Pb) is 0.06 ppb in Indian Standards (24 h) and 0.017 ppb rolling average of 3 months according to US EPA. The results show the levels to be over 5.0×10^5 times higher than the set limits according to Indian Standards and 3.0×10^6 times the US EPA standards. Apart from aluminum, manganese (Mn) and potassium (K) are generally used in firecrackers. Potassium is one of the key components of burning firecrackers. High concentrations of these metals are a matter of concern but not surprising. The concerning fact is that over 13.18 ppm and 34.30 ppm of Pb and Cu, respectively, cannot be filtered by the human body (Elbagermi et al. 2012). It cannot be filtered by ordinary masks either. Copper (Cu) is responsible for the blue color found in firecrackers (Liu et al. 2018). Magnesium (Mg) is the substitute for aluminum in most firecrackers ("CSIR developed Less Polluting Firecrackers named – SWAS, SAFAL, and STAR | Current Affairs Today," 2018). Magnalium, i.e., magnesium-aluminum alloy, is used to inherit properties of both Mg and Al. In this regard, the level of Mg is found very low in comparison to Al. It indicates the use of primitive firecrackers being burst (Harrison and Shallcross 2011; ChemViews 2014). Metals like Al and Mg are responsible for making firecrackers more colorful and brighter (Tsai et al. 2012). Arsenic (As), another toxic metal, does not have an average daily limit but has an annual average limit of 1.98 ppb. The observed value is found over 1.38 ppm, which is ~600 times higher than the standard. Thus, a human being breathing this air cannot save himself from inhaling this high quantity of As. The metals cadmium (Cd) and vanadium (V) present in lower amounts in comparison to other metals but found well above the observed generic levels (Tchounwou et al. 2012). These metals do not have an ambient air standard, nor do they pose any danger in scarce amounts. Although, they are not present in scarce amounts (Suvarapu and Baek 2016). They are known for damaging both human bodies and the environment when present in such quantities.

Al, Cu, Cd, Pb were found with the highest concentration in the fine particle range, and vanadium (V), manganese (Mn), magnesium (Mg) particles were found in the coarse particle range. In contrast, the highest concentration of arsenic (As) was observed within the qUFP particle range (Fig. 4). Figure 4 depicted that the sub-fine particulate matter size range shows the lesser contribution of the elements compared to other size ranges. While qUFP shows as ~33% as highest and Mg (14%) as lowest. Mostly, fine and coarse particle range having a higher % of elements.

Interestingly Cd and Pb having less % contribution in the coarse particle range but V and Mn have higher % contributions. Similarly, Cd and Pb are having a higher % contribution in the fine particle range, but V and Mg having less % contribution. However, these toxic elements are significantly higher than national and international standards in each particle range, significantly lower size range (qUFP and Subfine). qUFP and subfine smaller particles are more harmful than larger particles because their penetration and retention efficiency are much higher than larger particles (Kumar et al. 2011; Rohra et al. 2018).

The implementation of green crackers by India's government in concern to reduce the impact of fire bursting emission into ambient air might affect the element's composition ("https://indiankanoon.org/doc/63024516/Supreme Court of India; Arjun Gopal vs. Union Of India on 23 October 2018," n.d.). Only green-crackers are allowed in Delhi and the NCR region to celebrate the Diwali festival because it contains significantly less toxic elements than conventional firecrackers ("CSIR developed Less Polluting Firecrackers named – SWAS, SAFAL, and STAR | Current Affairs Today," 2018; TV Venkateswaran 2019).

Conclusion

The study clearly stated that people did not follow very strictly the rules made by the Supreme Court, whereas, before 8:00 pm, people started celebrating firecrackers. The number concentration of particles in the size bin of qUFP was higher than other size bins on different days, including Diwali. But the PNC peak of Diwali did not shift to any other size bin mode. Regarding element concentration distribution, the size bin of qUFP diameter particles showed the highest contribution of As (32.7%) followed

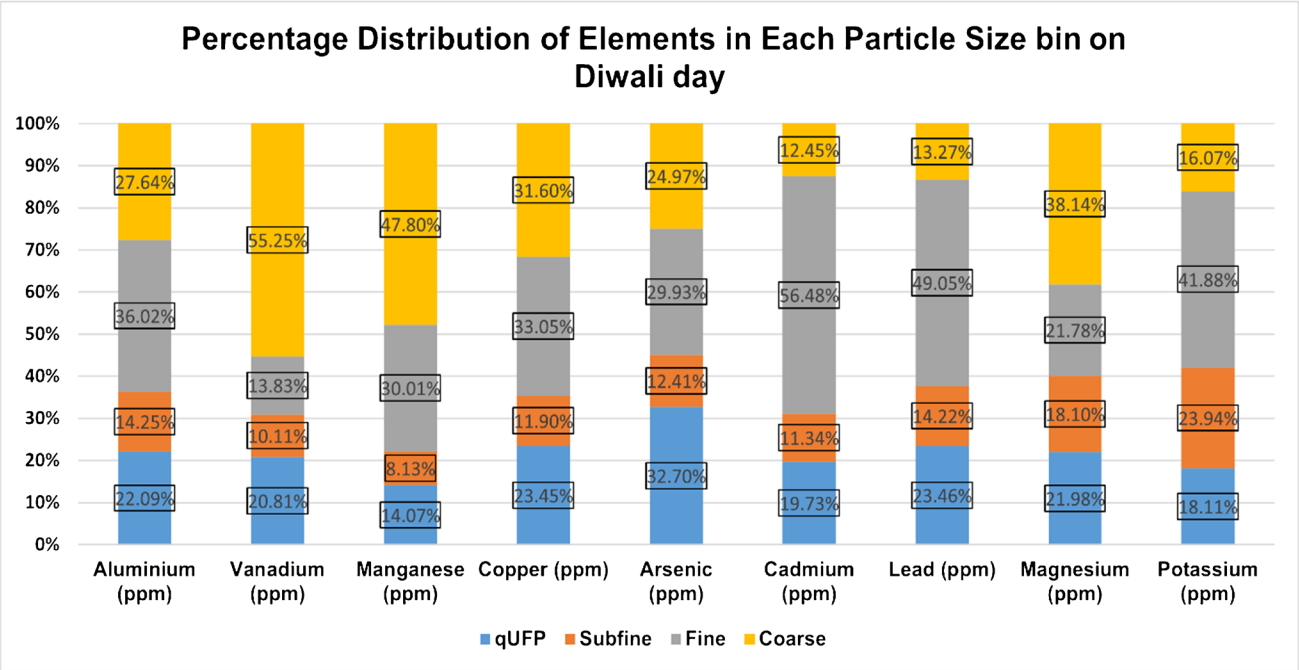


Fig. 4 Percentage distribution of elements in each particle size (qUFP, subfine, fine and coarse particle range)

by Pb (23.46%), Cu (23.45%), Al (22.9%), Mg (21.98%), V (20.81%), Cd (19.73%), K (18.11%), and Mn (14.07%). Meanwhile, the toxicity of these elements also increases due to their presence in finer particle size. This study's outcomes indicate that there is still an urgent requirement of comprehensive awareness among the public regarding the latest government guidelines related to firecrackers bursting during the Diwali festival, which may further reduce the associated health risks among people. The implementation of green crackers by India's government in concern to reduce the impact of fire bursting emission into ambient air might affect the element's composition ("https://indiankanoon.org/doc/63024516/Supreme Court of India; Arjun Gopal vs. Union Of India on 23 October 2018," n.d.). Only green-crackers should be allowed in Delhi and NCR region in celebrating the Diwali festival because it contains significantly less toxic elements than conventional firecrackers ("CSIR developed Less Polluting Firecrackers named – SWAS, SAFAL, and STAR | Current Affairs Today," 2018; TV Venkateswaran 2019). However, except for source-identified events such as Diwali, the main limitation of this type of study is that it does not tell us much about the sources' categorization of PNC on normal monitoring days and being a single monitoring location does not represent the "PNC" status across the entire city of Delhi. In the future, multilocation and long-duration studies (especially for chemical characterization) for PMs in the context of PNSD should be conducted, so that the PNC of Diwali period in Delhi and its influencing factors can be distinguished on a larger scale from the pre- and post-Diwali periods.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12517-021-09385-1.

Acknowledgements The authors are very thankful to Advance Air and Acoustics Research Lab, Delhi Technological University, for providing instrumental facilities to conduct this experiment inside the campus. The first author is grateful to the University Grand Commission (UGC), India, for providing a Doctoral Scholarship.

Data availability The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Author contribution Methodology, Monitoring, Formal analysis and investigation, draft preparation: Shailendra Kumar Yadav; Conceptualization, supervision, review & editing: Rajeev Kumar Mishra; Investigation, resources, data curation, review & editing: B.R. Gurjar

Declarations

Ethics approval and consent to participate This article does not contain any studies with human participants or animals performed by any authors.

Consent for publication The authors confirm that the final version of the manuscript has been reviewed, approved, and consented for publication by all authors.

Conflict of interest The authors declare that they have no competing interests.

References

Amato F, Lucarelli F, Nava S, et al (2016) Case studies of source apportionment and suggested measures at southern European cities. Issues Environ Sci Technol 2016-Janua:168–263. https://doi.org/10.1039/9781782626589-00168

Apte JS, Kirchstetter TW, Reich AH et al (2011) Concentrations of fine, ultrafine, and black carbon particles in auto-rickshaws in New Delhi, India. Atmos Environ 45:4470–4480. https://doi.org/10.1016/j.atmosenv.2011.05.028

Arub Z, Bhandari S, Gani S et al (2020a) Air mass physiochemical characteristics over New Delhi: impacts on aerosol hygroscopicity and cloud condensation nuclei (CCN) formation. Atmos Chem Phys 20:6953–6971. https://doi.org/10.5194/acp-20-6953-2020

Arub Z, Bhandari S, Gani S et al (2020b) Air mass physiochemical characteristics over New Delhi: impacts on aerosol hygroscopicity and cloud condensation nuclei (CCN) formation. Atmos Chem Phys 20:6953–6971. https://doi.org/10.5194/acp-20-6953-2020

Balakrishnan K, Dey S, Gupta T et al (2019) The impact of air pollution on deaths, disease burden, and life expectancy across the states of India: the Global Burden of Disease Study 2017. Lancet Planet Heal 3:e26–e39. https://doi.org/10.1016/S2542-5196(18)30261-4

Barman SC, Singh R, Negi MPS, Bhargava SK (2008) Ambient air quality of Lucknow City (India) during use of fireworks on Diwali Festival. Environ Monit Assess 137:495–504. https://doi.org/10.1007/s10661-007-9784-1

Becker JM, Iskandrian S, Conkling J (2000) Fatal and near-fatal asthma in children exposed to fireworks. Ann Allergy, Asthma Immunol 85:512–513. https://doi.org/10.1016/S1081-1206(10)62581-1

Beig G, Chate DM, Ghude SD et al (2013) Evaluating population exposure to environmental pollutants during Deepavali fireworks displays using air quality measurements of the SAFAR network. Chemosphere 92:116–124. https://doi.org/10.1016/j.chemosphere.2013.02.043

Betha R, Balasubramanian R (2013) Particulate emissions from commercial handheld sparklers: evaluation of physical characteristics and emission rates. Aerosol Air Qual Res 13:301–307. https://doi.org/10.4209/aaqr.2012.08.0208

Bhandari S, Gani S, Patel K et al (2020) Sources and atmospheric dynamics of organic aerosol in New Delhi, India: insights from receptor modeling. Atmos Chem Phys 20:735–752. https://doi.org/10.5194/acp-20-735-2020

Brugge D, Fuller CH (2021) Ambient combustion ultrafine particles and health

Chatterjee A, Sarkar C, Adak A et al (2013) Ambient air quality during Diwali Festival over Kolkata — a mega-city in India. Aerosol Air Qual Res 13:1133–1144. https://doi.org/10.4209/aaqr.2012.03.0062

ChemViews (2014) The Chemistry of Fireworks. ChemViews. https://doi.org/10.1002/chemv.201400134

Chen R, Hu B, Liu Y et al (2016) Beyond PM2.5: The role of ultrafine particles on adverse health effects of air pollution. Biochim Biophys Acta - Gen Subj 1860:2844–2855. https://doi.org/10.1016/j.bbagen.2016.03.019

CSIR developed Less Polluting Firecrackers named – SWAS, SAFAL and STAR | Current Affairs Today (2018) In: CSIR. https://pib.gov.in/newsite/PrintRelease.aspx?relid=184462. Accessed 13 May 2020

Dockery DW, Pope CA (1994) Acute respiratory effects of particulate air pollution. Annu Rev Public Health 15:107–132. https://doi.org/10.1146/annurev.pu.15.050194.000543

Drewnick F, Hings SS, Curtius J et al (2006) Measurement of fine particulate and gas-phase species during the New Year's fireworks 2005 in Mainz, Germany. Atmos Environ 40:4316–4327. https://doi.org/10.1016/j.atmosenv.2006.03.040

Dutcher DD, Perry KD, Cahill TA, Copeland SA (1999) Effects of indoor pyrotechnic displays on the air quality in the houston astrodome. J Air Waste Manag Assoc 49:156–160. https://doi.org/10.1080/10473289.1999.10463790

Elbagermi MA, Edwards HGM, Alajtal AI (2012) Monitoring of heavy metal content in fruits and vegetables collected from production and market sites in the Misurata area of Libya. ISRN Anal Chem 2012:1–5. https://doi.org/10.5402/2012/827645

Fuller CH, Brugge D, Williams PL et al (2013) Indoor and outdoor measurements of particle number concentration in near-highway homes. J Expo Sci Environ Epidemiol 23:506–512. https://doi.org/10.1038/jes.2012.116

Fuller CH, O'Neill MS, Sarnat JA et al (2018) Short-and medium-term associations of particle number concentration with cardiovascular markers in a Puerto Rican cohort. Environ Res 166:595–601. https://doi.org/10.1016/j.envres.2018.06.042

Gaeggeler K, Prevot ASH, Dommen J et al (2008) Residential wood burning in an Alpine valley as a source for oxygenated volatile organic compounds, hydrocarbons and organic acids. Atmos Environ 42:8278–8287. https://doi.org/10.1016/j.atmosenv.2008.07.038

Gani S, Bhandari S, Seraj S et al (2018) Submicron aerosol composition in the world's most polluted megacity: the Delhi Aerosol Supersite campaign. Atmos Chem Phys Discuss 5:1–33. https://doi.org/10.5194/acp-2018-1066

Gani S, Bhandari S, Seraj S et al (2019) Submicron aerosol composition in the world's most polluted megacity: the Delhi Aerosol Supersite study. Atmos Chem Phys 19:6843–6859. https://doi.org/10.5194/acp-19-6843-2019

Gani S, Bhandari S, Patel K, et al (2020) Particle number concentrations and size distribution in a polluted megacity: the Delhi Aerosol Supersite study. Atmos Chem Phys 1–30. https://doi.org/10.5194/acp-2020-6

Grundström M, Hak C, Chen D et al (2015) Variation and co-variation of PM10, particle number concentration, NOx and NO2 in the urban air — relationships with wind speed, vertical temperature gradient and weather type. Atmos Environ 120:317–327. https://doi.org/10.1016/j.atmosenv.2015.08.057

Guttikunda SK, Nishadh KA, Jawahar P (2019) Air pollution knowledge assessments (APnA) for 20 Indian cities. Urban Clim 27:124–141. https://doi.org/10.1016/j.uclim.2018.11.005

Habil M, Massey DD, Taneja A (2016) Personal and ambient PM2.5 exposure assessment in the city of Agra. Data Br 6:495–502. https://doi.org/10.1016/j.dib.2015.12.040

Habre R, Zhou H, Eckel SP et al (2018) Short-term effects of airport-associated ultrafine particle exposure on lung function and inflammation in adults with asthma. Environ Int 118:48–59. https://doi.org/10.1016/j.envint.2018.05.031

Hagan DH, Gani S, Bhandari S et al (2019) Inferring aerosol sources from low-cost air quality sensor measurements: a case study in Delhi, India. Environ Sci Technol Lett 6:467–472. https://doi.org/10.1021/acs.estlett.9b00393

Hamad S, Green D, Heo J (2016) Evaluation of health risk associated with fireworks activity at Central London. Air Qual Atmos Heal 9:735–741. https://doi.org/10.1007/s11869-015-0384-x

Harris SJ, Maricq MM (2001) Signature size distributions for diesel and gasoline engine exhaust particulate matter. J Aerosol Sci 32:749–764. https://doi.org/10.1016/S0021-8502(00)00111-7

Harrison T, Shallcross DE (2011) Smoke is in the air : how fireworks affect air quality. Did you realise that fireworks cause measurable air pollution? Sci Sch 47–51

Health Effects Institute, Project. and the I for HM and EGB of D (2019) A Special Report on Global Exposure to Air Pollution and its Disease Burden What is the State of Global Air? Who

is it for? How can I explore the data? MAHealth Eff Institute, Boston, p 24
Hirai K, Yamazaki Y, Okada K et al (2000) Acute eosinophilic pneumonia associated with smoke from fireworks. Intern Med 39:401–403. https://doi.org/10.2169/internalmedicine.39.401
https://indiankanoon.org/doc/63024516/Supreme Court of India (n.d.); Arjun Gopal vs Union Of India on 23 October, 2018. https://indiankanoon.org/doc/63024516/. Accessed 13 May 2020
Hu J, Jathar S, Zhang H et al (2017) Long-term particulate matter modeling for health effect studies in California — Part 2: concentrations and sources of ultrafine organic aerosols. Atmos Chem Phys 17:5379–5391. https://doi.org/10.5194/acp-17-5379-2017
India.com (2018) Dhanteras Evening Sees Massive Traffic Jam on Delhi-Gurugram Route; 6 km-bottleneck From Mahipalpur to Udyog Vihar | India.com
International Commission on Radiological Protection (1994) Human respiratory tract model for radiological protection. A report of a task group of the international commission on radiological protection. Ann. ICRP 24:1–482
Jain N, Bhatia A, Pathak H (2014) Emission of air pollutants from crop residue burning in India. Aerosol Air Qual Res 14:422–430. https://doi.org/10.4209/aaqr.2013.01.0031
Joshi M, Khan A, Anand S, Sapra BK (2016) Size evolution of ultrafine particles: differential signatures of normal and episodic events. Environ Pollut 208:354–360. https://doi.org/10.1016/j.envpol.2015.10.001
Joshi M, Nakhwa A, Khandare P et al (2019) Simultaneous measurements of mass, chemical compositional and number characteristics of aerosol particles emitted during fireworks. Atmos Environ 217:116925. https://doi.org/10.1016/j.atmosenv.2019.116925
Kanawade VP, Srivastava AK, Ram K et al (2020) What caused severe air pollution episode of November 2016 in New Delhi? Atmos Environ 222:117125. https://doi.org/10.1016/j.atmosenv.2019.117125
Karagulian F, Belis CA, Dora CFC et al (2015) Contributions to cities' ambient particulate matter (PM): a systematic review of local source contributions at global level. Atmos Environ 120:475–483
Kittelson DB, Watts WF, Johnson JP (2004) Nanoparticle emissions on Minnesota highways. Atmos Environ 38:9–19. https://doi.org/10.1016/j.atmosenv.2003.09.037
Knibbs LD, Cole-hunter T, Morawska L (2011) A review of commuter exposure to ultrafine particles and its health effects. Atmos Environ 45:2611–2622. https://doi.org/10.1016/j.atmosenv.2011.02.065
Kompalli SK, Suresh Babu S, Krishna Moorthy K et al (2014) The formation and growth of ultrafine particles in two contrasting environments: a case study. Ann Geophys 32:817–830. https://doi.org/10.5194/angeo-32-817-2014
Kompalli SK, Babu SS, Udayasoorian C, Jayabalakrishnan RM (2018) Role of anthropogenic emissions and meteorology on ultrafine particle bursts over a high altitude site in Western Ghats during pre-monsoon. J Atmos Solar-Terrestrial Phys 179:378–388. https://doi.org/10.1016/j.jastp.2018.09.001
Kompalli SK, Nair VS, Jayachandran V et al (2020) Particle number size distributions and new particle formation events over the northern Indian Ocean during continental outflow. Atmos Environ 238:117719. https://doi.org/10.1016/j.atmosenv.2020.117719
Kong S, Li X, Li L et al (2015) Variation of polycyclic aromatic hydrocarbons in atmospheric PM2.5 during winter haze period around 2014 Chinese Spring Festival at Nanjing: insights of source changes, air mass direction and firework particle injection. Sci Total Environ 520:59–72. https://doi.org/10.1016/j.scitotenv.2015.03.001
Kulshrestha UC, Nageswara Rao T, Azhaguvel S, Kulshrestha MJ (2004) Emissions and accumulation of metals in the atmosphere due to crackers and sparkles during Diwali festival in India. Atmos Environ 38:4421–4425. https://doi.org/10.1016/j.atmosenv.2004.05.044
Kumar P, Gurjar BR, Nagpure AS, Harrison RM (2011) Preliminary estimates of nanoparticle number emissions from road vehicles in megacity Delhi and associated health impacts. Environ Sci Technol 45:5514–5521. https://doi.org/10.1021/es2003183
Kumar P, Morawska L, Birmili W et al (2014) Ultrafine particles in cities. Environ Int 66:1–10. https://doi.org/10.1016/j.envint.2014.01.013
Kumar M, Singh RKS, Murari V et al (2016) Fireworks induced particle pollution: a spatio-temporal analysis. Atmos Res 180:78–91. https://doi.org/10.1016/j.atmosres.2016.05.014
Künzli PN (2015) Long-term health effects of ambient ultrafine particles (UFP). 0–5
Lane KJ, Levy JI, Scammell MK et al (2016) Association of modeled long-term personal exposure to ultrafine particles with inflammatory and coagulation biomarkers. Environ Int 92–93:173–182. https://doi.org/10.1016/j.envint.2016.03.013
Li J, Xu T, Lu X et al (2017) Online single particle measurement of fireworks pollution during Chinese New Year in Nanning. J Environ Sci (China) 53:184–195. https://doi.org/10.1016/j.jes.2016.04.021
Liu DY, Rutherford D, Kinsey M, Prather KA (1997) Real-time monitoring of pyrotechnically derived aerosol particles in the troposphere. Anal Chem 69:1808–1814. https://doi.org/10.1021/ac9612988
Liu Y, Li S, Sun C et al (2018) Pollution level and health risk assessment of PM2.5-bound metals in baoding city before and after the heating period. Int J Environ Res Public Health 15 https://doi.org/10.3390/ijerph15102286
Lv Y, Chen X, Wei S et al (2020) Sources, concentrations, and transport models of ultrafine particles near highways: a literature review. Build Environ 186:107325
Oberdörster G, Sharp Z, Atudorei V, et al (2004) Translocation of inhaled ultrafine particles to the brain. In: Inhalation Toxicol. pp 437–445
Oberdörster G, Oberdörster E, Oberdörster J (2005) Nanotoxicology: an emerging discipline evolving from studies of ultrafine particles. Environ Health Perspect 113:823–839
Patel K, Bhandari S, Gani S et al (2021) Sources and dynamics of submicron aerosol during the autumn onset of the air pollution season in Delhi, India. ACS Earth Sp Chem 5:118–128. https://doi.org/10.1021/acsearthspacechem.0c00340
Ravindra K, Mor S, Kaushik CP (2003) Short-term variation in air quality associated with firework events: a case study. J Environ Monit 5:260–264. https://doi.org/10.1039/b211943a
Rivas I, Kumar P, Hagen-Zanker A et al (2017) Determinants of black carbon, particle mass and number concentrations in London transport microenvironments. Atmos Environ 161:247–262. https://doi.org/10.1016/j.atmosenv.2017.05.004
Rohra H, Tiwari R, Khare P, Taneja A (2018) Indoor-outdoor association of particulate matter and bounded elemental composition within coarse, quasi-accumulation and quasi-ultrafine ranges in residential areas of northern India. Sci Total Environ 631–632:1383–1397. https://doi.org/10.1016/j.scitotenv.2018.03.095
Rönkkö T, Timonen H (2019) Overview of sources and characteristics of nanoparticles in urban traffic-influenced areas. J Alzheimer's Dis 72:15–28. https://doi.org/10.3233/JAD-190170
Saffari A, Daher N, Shafer MM et al (2013) Seasonal and spatial variation of trace elements and metals in quasi-ultrafine (PM0.25) particles in the Los Angeles metropolitan area and characterization of their sources. Environ Pollut 181:14–23. https://doi.org/10.1016/j.envpol.2013.06.001
Saraswat A, Apte JS, Kandlikar M et al (2013) Spatiotemporal land use regression models of fine, ultrafine, and black carbon

particulate matter in New Delhi, India. Environ Sci Technol 47:12903–12911. https://doi.org/10.1021/es401489h
Sarkar S, Khillare PS, Jyethi DS et al (2010) Chemical speciation of respirable suspended particulate matter during a major firework festival in India. J Hazard Mater 184:321–330. https://doi.org/10.1016/j.jhazmat.2010.08.039
Schraufnagel DE (2020) The health effects of ultrafine particles. Exp Mol Med 52:311–317. https://doi.org/10.1038/s12276-020-0403-3
Seidel DJ, Birnbaum AN (2015) Effects of Independence Day fireworks on atmospheric concentrations offine particulate matter in the United States. Atmos Environ 115:192–198. https://doi.org/10.1016/j.atmosenv.2015.05.065
Shi GL, Liu GR, Tian YZ et al (2014) Chemical characteristic and toxicity assessment of particle associated PAHs for the short-term anthropogenic activity event: during the Chinese New Year's Festival in 2013. Sci Total Environ 482–483:8–14. https://doi.org/10.1016/j.scitotenv.2014.02.107
Singh DP, Gadi R, Mandal TK et al (2010) Study of temporal variation in ambient air quality during Diwali festival in India. Environ Monit Assess 169:1–13. https://doi.org/10.1007/s10661-009-1145-9
Singh A, Pant P, Pope FD (2019) Air quality during and after festivals: aerosol concentrations, composition and health effects. Atmos Res 227:220–232
Strasser G, Hiebaum S, Neuberger M (2018) Commuter exposure to fine and ultrafine particulate matter in Vienna. Wien Klin Wochenschr 130:62–69. https://doi.org/10.1007/s00508-017-1274-z
Supreme Court bans use of heavy metals in firecrackers to prevent air pollution - Education Today News (2017). In: India Today. https://www.indiatoday.in/education-today/gk-current-affairs/story/supreme-court-metals-ban-firecrackers-1027387-2017-08-01. Accessed 13 May 2020
Suvarapu LN, Baek SO (2016) Determination of heavy metals in the ambient atmosphere: a review. Toxicol Ind Health 33:79–96
Tchounwou PB, Yedjou CG, Patlolla AK, Sutton DJ (2012) Heavy metal toxicity and the environment. EXS 101:133–164
Terzano C, Di Stefano F, Conti V et al (2010) Air pollution ultrafine particles: toxicity beyond the lung. Eur Rev Med Pharmacol Sci 14:809–821
Thomas M (2018) Even before Diwali fireworks, Delhi's air is already deadly — Quartz India. In: Quartz India. https://qz.com/india/1451841/even-before-diwali-fireworks-delhis-air-is-already-deadly/. Accessed 13 May 2020
Thurston GD, Ito K, Lall R (2011) A source apportionment of U.S. fine particulate matter air pollution. Atmos Environ 45:3924–3936. https://doi.org/10.1016/j.atmosenv.2011.04.070
Time India (2014) Diwali rush: Delhi traffic up by 50% | Delhi News - Times of India
Tiwari S, Chate DM, Srivastava MK et al (2012) Statistical evaluation of PM 10 and distribution of PM 1, PM 2.5, and PM 10 in ambient air due to extreme fireworks episodes (Deepawali festivals) in megacity Delhi. Nat Hazards 61:521–531. https://doi.org/10.1007/s11069-011-9931-4
Trivedi U (2018) India Court Says No Nationwide Firecracker Ban as Smog Worsens - Bloomberg. In: Bloomberg. https://www.bloomberg.com/news/articles/2018-10-23/india-court-says-no-nationwide-firecracker-ban-as-smog-worsens. Accessed 12 May 2020
Tsai HH, Chien LH, Yuan CS et al (2012) Influences of fireworks on chemical characteristics of atmospheric fine and coarse particles during Taiwan's Lantern Festival. Atmos Environ 62:256–264. https://doi.org/10.1016/j.atmosenv.2012.08.012
Vecchi R, Bernardoni V, Cricchio D et al (2008) The impact of fireworks on airborne particles. Atmos Environ 42:1121–1132. https://doi.org/10.1016/j.atmosenv.2007.10.047
TV Venkateswaran (2019) Eco-friendly crackers: how green is your Diwali? - The Federal. In: Fed. https://thefederal.com/the-eighth-column/eco-friendly-crackers-how-green-is-your-diwali/. Accessed 13 May 2020
Viana M, Kuhlbusch TAJ, Querol X et al (2008) Source apportionment of particulate matter in Europe: a review of methods and results. J Aerosol Sci 39:827–849
Wang Y, Zhuang G, Xu C, An Z (2007) The air pollution caused by the burning of fireworks during the lantern festival in Beijing. Atmos Environ 41:417–431. https://doi.org/10.1016/j.atmosenv.2006.07.043
Wang Y, Hopke PK, Rattigan OV (2012) A new indicator of fireworks emissions in Rochester, New York. Environ Monit Assess 184:7293–7297. https://doi.org/10.1007/s10661-011-2497-5
Wehner B, Wiedensohler A, Heintzenberg J (2000) Submicrometer aerosol size distributions and mass concentration of the Millennium fireworks 2000 in Leipzig, Germany. J Aerosol Sci 31:1489–1493. https://doi.org/10.1016/S0021-8502(00)00039-2
WHO (2018) WHO Global Ambient Air Quality Database (update 2018). In: Ambient Air Qual. Database (update 2018). https://www.who.int/airpollution/data/cities/en/. Accessed 12 May 2020
Yadav SK, Mishra RK, Gurjar BR (2019) Ultrafine Particles in Concern of Vehicular Exhaust—An Overview. In: Agarwal AK, Dhar A, Sharma N, Shukla PC (eds) Engine Exhaust Particulates, 1st edn. Springer, Singapore, pp 7–38
Yadav SK, Kumar M, Sharma Y et al (2019a) Temporal evolution of submicron particles during extreme fireworks. Environ Monit Assess 191:1–12. https://doi.org/10.1007/s10661-019-7735-2
Yadav SK, Kompalli SK, Gurjar BR, Mishra RK (2021) Aerosol number concentrations and new particle formation events over a polluted megacity during the COVID-19 lockdown. Atmos Environ 118526. https://doi.org/10.1016/j.atmosenv.2021.118526
Yang L, Gao X, Wang X et al (2014) Impacts of firecracker burning on aerosol chemical characteristics and human health risk levels during the chinese new year celebration in Jinan, China. Sci Total Environ 476–477:57–64. https://doi.org/10.1016/j.scitotenv.2013.12.110
Zheng T, Wang HW, Li XB et al (2021) Impacts of traffic on roadside particle variations in varied temporal scales. Atmos Environ 253:118354. https://doi.org/10.1016/j.atmosenv.2021.118354
Zhou G, Xu J, Gao W et al (2018) Characteristics of PM1 over Shanghai, relationships with precursors and meteorological variables and impacts on visibility. Atmos Environ 184:224–232. https://doi.org/10.1016/j.atmosenv.2018.04.041

Folksonomy Based Fuzzy Filtering Recommender System

Vinod Kumar Ayush Tanwar Arhan Relan Chitraksh Grover

Computer Science & Engineering Department
Delhi Technological University
New Delhi, India
vinod_k@dtu.ac.in, ayushtw@gmail.com, arhanrelan@gmail.com, chitrakshgrover@gmail.com

***Abstract*—As the choice of items on the World Wide Web increases, the difficulty to choose the right item also increases. Fast and effective recommender systems help to overcome this difficulty by automatically using algorithms and data to recommend items like movies, webpages, videos, etc. Conventional collaborative filtering recommender systems use the rating data to find the neighborhood of users in terms of their preferences. However, these recommender systems do not incorporate the domain knowledge in their recommendations. In this paper we propose a hybrid folksonomy-based fuzzy collaborative filtering recommender system which uses domain knowledge of the items and combines them with the collaborative process. To further enhance the performance of the system we use fuzzy clusters to represent the neighborhood in a probabilistic manner which makes the recommendation process more accurate as verified by our experimental results**

Keywords— ***Recommender System, Folksonomy, Collaborative Filtering, Fuzzy Clustering, Rating Matrix.***

I. INTRODUCTION

In the age of the Internet, the exponential increase in the amount of information available on the Web has led to the problem of overfeeding information to the users. Recommender Systems are one particular type of informational filtering systems which tackle this problem. They are used to forecast a set of relevant items for a user on the basis of his/her 'liking'. Thus, they are an asset for e-commerce and other online platforms as they help customers to find the required product. Since the majority of the OTT (over-the-top) platforms are driven by impulse-based watching from the recommended items[1], recommender systems pose a great business value. Recommender systems have been widely used in movies[2], e- commerce[3], tourism[4], news[5] etc.

Recommender systems are categorized as content-based recommender systems[6] and collaborative filtering recommender systems[7]. Content-based recommender systems build item profiles and recommend similar items. However, this may lead to limited items being recommended repeatedly. Thus, Collaborative filtering is more widely used as the user-item rating matrix provides users' feedback for earlier chosen items. This matrix is then exploited to evaluate similarity in rating patterns of users and items and provide suitable recommendations.

Major shortcomings of collaborative filtering recommender systems are the cold start problem and sparse rating matrix. For sparsity, various dimensionality reduction techniques[8] are used to minimise it. Moreover, various knowledge-based techniques such as ontology and folksonomy are used to eliminate the cold start problem and enhance the recommendation process. Folksonomy[9] is defined as a semantic social tagging network. In recommender systems users can allocate social tags to items that are used to link them in the item domain. These tags represent users' choices and preferences and can be utilized to classify items accordingly. Items similar to the considered item can be then obtained for recommendation.

We propose a hybrid recommendation model[4] that utilizes folksonomy along with collaborative filtering as explicit user feedback and thus, provides personalized and intuitive recommendations. The major contributions of our research are as follows:

- Probabilistic clustering algorithm - We use the Fuzzy C-means algorithm which considers the dynamicity of each individual's preferences. Since each user belongs probabilistically to multiple clusters, the recommendations are personalized.
- Folksonomy based on explicit user tags - Since users' opinions are of utmost importance, the explicit tags in addition to the ratings given by the users provide a deeper insight of their vivid choices and enhance the recommendation process.
- Simple and intuitive approach - The hybrid approach conglomerates the advantages of the individual approaches and efficiently incorporates new users or items as well. This aids in fulfilling the desired aim of personalized and dynamic recommendations.

The rest of the paper is organized as follows: Section 2 elaborates the approaches previously used in the recommendation system domain. Section 3 explains our model. Section 4 demonstrates the experimental results and observations. Section 5 concludes our study.

II. RELATED WORK

Movie recommender systems [2] are used widely in our streaming platforms (Netflix, YouTube, Prime Video, etc.) They recommend the right movie for a particular user by using various algorithms and data. One of the most popular approaches in a movie recommender system is Collaborative filtering. A Collaborative filtering-based recommender system [10][11] uses "User access patterns" or "User Behavior" and "Item Correlations" to recommend items.

2021 IEEE Symposium Series on Computational Intelligence (SSCI) | 978-1-7281-9048-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/SSCI50451.2021.9660032

978-1-7281-9048-8/21/$31.00 ©2021 IEEE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 31,2022 at 08:49:38 UTC from IEEE Xplore. Restrictions apply.

Collaborative Filtering recommender system uses User-Item Rating Matrix [7]. Collaborative Filtering recommender systems suffer from the "Cold-Start" problem [12] which is the inability of a system to perform accurately because of insufficient data at the initial stages of a system.

Netflix [1] initially asks its users to select a minimum of 3 movies or TV-shows to resolve the cold start problem. It then categorizes its users accordingly and recommends them according to these selected Movies/TV-shows. Jaschke et al. [9] used a Tag-based system called Folksonomy system to resolve the Cold-Start problem. According to Semeraro [13], folksonomies can be employed to provide domain knowledge to the recommender system. Folksonomy gives descriptions and categories to various users and items [14]. These descriptions/tags can be harvested from Movie databases like IMDB.

Hybrid recommender systems combine two or more recommender system approaches. Tran [3] was one of the first to employ a hybrid recommender system, he combined the collaborative filtering approach with a knowledge-based approach to devise a dynamic e-commerce recommender system. Logesh [4] created a hybrid recommender system for the tourism industry by combining Collaborative filtering, Content-based and Knowledge-based recommendation approaches. In our approach, we have used a Hybrid approach which combines the data-centric Collaborative filtering approach with the domain-centric Folksonomy approach to provide accurate and quick recommendations.

III. Our Approach

In this section we present our hybrid approach for a recommender system which uses folksonomy along with collaborative filtering for predicting ratings and providing recommendations. In a conventional collaborative filtering system, a user- item rating matrix represents users' views regarding the items. In addition to this, a folksonomy of users and the public tags allocated by them to various items is exploited to determine relevant and semantically similar items.

Fig. 1 shows the architecture of our model. The model is divided into four modules where each fulfils a specific purpose.

1. Preprocessing

This module extracts two types of information sources, the user-item rating matrix and the folksonomy, from the database. The rating matrix is defined as follows.

Rating Matrix: It is the matrix containing ratings of M users by N items where the users are represented by the row headers and the column headers represent the items. Let U_m be the m^{th} user and I_n be the n^{th} item, then r_{mn} is the rating given to I_n by U_m.

$$\mathbf{0 < r_{mn} < 5,\ r_{mn} \in R}$$

User ratings lie between 0 to 5 with increments of 0.5. An item not rated by a user is represented by a default value 0.

Folksonomy, as given in Figure 2, is a semantic network of users, items and the public tags allocated by the users to the items. Since these tags are given by users, they do not follow any specific format and may differ for the same item as perceived by different users. However, they can be exploited to provide personalized recommendations for different users tagging an item similarly or a single user tagging different items similarly.

2. Item similarity

This module utilizes the folksonomy to compute the collaborative item similarity. Initially, the folksonomy is pre-processed to remove any unusual tags for instance tags that do not start with a letter. Lexically similar and misspelt tags such as car and cars, are grouped together using the Levenshtein similarity method [15]. Similarity between tags

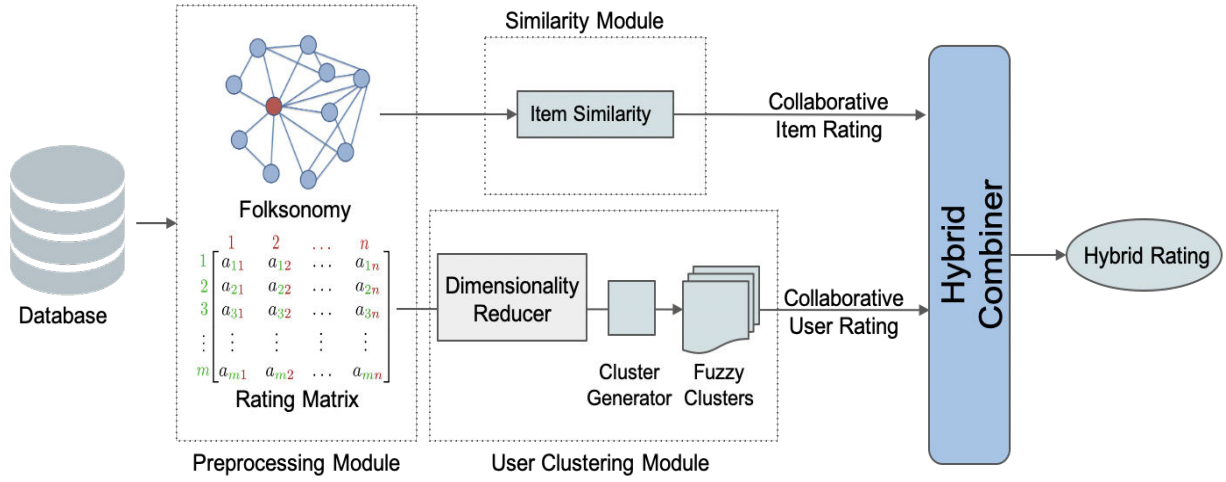


Fig. 1. Illustrating the Architecture of our approach

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 31,2022 at 08:49:38 UTC from IEEE Xplore. Restrictions apply.

t_i and t_j is computed using Lin's Information Theory [16] as given in the following formula.

$$TS = \frac{\log(pr(t_k))^2}{\log(pr(t_i))+\log(pr(t_j))} \quad (1)$$

Where t_k is the nearest tag that subsumes t_i and t_j and $pr(t_i)$ gives the probability of allocation of t_i to an item.

$$pr(t_i) = \frac{n_i}{N} \quad (2)$$

Where n_i gives the number of items associated with tag i and N gives the total number of items.

The folksonomy similarity, FS, between two items m and n for each tag i is computed using the cosine similarity coefficient as:

$$FS(m,n) = \frac{(\sum_i TS_{mi} \cdot TS_{ni})}{\sqrt{\sum_i TS^2_{mi}} \sqrt{\sum_i TS^2_{ni}}} \quad (3)$$

Finally, the collaborative item rating of item j for a particular user is computed as the weighted mean of top-K similar items' ratings. Each item is weighted by the similarity between that item and the item under consideration.

$$CIR(ij) = \frac{\sum_{k=1}^{K} FS(j,k) \cdot r_{ij}}{\sum_{k=1}^{K} FS(j,k)} \quad (4)$$

Where CIR gives the Collaborative Item Rating and r_{ij} gives the existing ratings.

3. *User clustering*

This module uses the user item rating matrix to cluster the users according to their preferences. Each user vector in the rating matrix is considered as a data point in n-dimensional space. In a traditional collaborative filtering system, these data points are clustered to form groups of users. Users belonging to each group are considered to have similar preferences since their user vectors lie in close proximity to each other in the n-dimensional feature space (here items) owing to their similar rating patterns.

However, one of the drawbacks of clustering in recommender systems is that mostly few popular items are rated by maximum number of users and some items have hardly any ratings given to them. In addition to this, the highly sparse data leads to the majority of points being present in a densely populated area. As a consequence, the clusters formed are not efficient with subtle preference of users being subdued by the popular preferences.

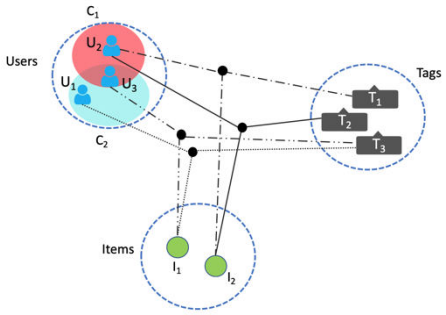


Fig. 2. Folksonomy

In our approach, unlike the conventional collaborative filtering, we assign soft clusters to the users. For this, we use Fuzzy c-means clustering [17] to assign probabilities to each user for a cluster. These probabilities represent the likelihood of the user being a part of the cluster. This alleviates the problem of clustering, since each user rather being present in any single cluster is now probabilistically assigned to many clusters. Moreover, these clusters have soft boundaries with well-defined centroids. Thus, during the collaborative filtering task system can now look for the tenuous preferences that may belong to different clusters and assign them probabilistically to the user.

The shortcoming of the fuzzy clustering is that it is not suitable for data with high dimensionality which is the intrinsic property of User-Item rating matrix. We use dimensionality reduction technique to project the high dimensional data into lower dimensions. The working of the module is as follows.

The rating matrix is passed to the dimensionality reducer which uses Principal Component Analysis to reduce the data into lower dimensions.

Let R_{mxn} be the rating matrix with U_i , i ∈[1,m] be a n dimensional user vector, then using the principal component analysis [18] is projected into a k dimension vector(k<n) given by $LDU_i = [\rho_1, \rho_2, ...\rho_k]$, i ∈[1, m]

Where ρ_j is the projection of U_i on the j^{th} principal axes of subspace defined by R_{mxn}.

These vectors are then passed to the cluster generator which uses the Fuzzy c-means algorithm to generate soft clusters.

Let $\gamma = [\gamma_1, \gamma_2...\gamma_c]$ be the k-dimensional centers of the C clusters, where C is predetermined and let μ_{iy} be the membership coefficient of i^{th} user in y^{th} cluster. Then γ and μ_{iy} are given by

$$\gamma_y = \frac{\sum_{i=1}^{m} \mu^{\eta}_{iy} \times LDU_i}{\sum_{i=1}^{m} \mu^{\eta}_{iy}} \quad (5)$$

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 31,2022 at 08:49:38 UTC from IEEE Xplore. Restrictions apply.

$$\mu_{iy} = \frac{1}{\sum_{l=1}^{C} \left(\theta(LDU_i,\gamma_y)/\theta(LDU_i,\gamma_l)\right)^{\eta}}, 1 < \eta < \infty \quad (6)$$

Where the objective function to be minimized is given by

$$F = \sum_{i=1}^{m} \sum_{j=1}^{c} {\mu^{\eta}}_{iy} \times \theta(LDU_i,\gamma_y)^2 \quad (7)$$

$$\theta(LDU_i,\gamma_y) = \frac{\left||LDU_i - \gamma_y|\right|^2}{\left||LDU_i|\right|^2 + \left|\left|\gamma_y\right|\right|^2 - LDU_i \cdot \gamma_y} \quad (8)$$

γ_y, is the distance between two k-dimensional vectors.

Finally, the collaborative user rating (CUR) for a particular item by user i is given by

$$CUR_{ij} = \sum_{y=1}^{C} \mu_{iy} \times \frac{\sum_{x=1}^{m} \mu_{xy} \times R_{xj}}{\sum_{x=1}^{m} \mu_{xy}}$$

4. *Hybrid combiner*

This module combines the Collaborative Item Rating obtained from the Similarity Module and the Collaborative User Rating obtained from the User Clustering Module to give a final Hybrid Recommendation.

We calculate the predicted rating by using the following formula.

$$HR = \beta \cdot CIR + (1 - \beta) \cdot CUR \quad (10)$$

Where HR = Hybrid Rating, CIR = Collaborative Item Rating, CUR = Collaborative User Rating, β = Hybrid Weighing factor

IV. Experimentation & Evaluation

1. *Dataset description*

We have used MovieLens Dataset [19] to evaluate our model. This dataset includes various movie features like budget, the cast, plot keywords, crew, countries released in, revenue, release dates, languages, posters, TMDB vote counts, voted means of different movies and production companies involved. It also contains 26 million ratings and 750,000 tag applications from 270,000 users for 45,000 movies. Ratings are on a scale of 1-5. The model is trained on a subset of the MovieLens dataset and contains 100,000 ratings from 700 users on 9,000 movies along with the relevant tags that are used to construct the folksonomy.

2. *Evaluation*

In the User clustering module, the number of clusters (C) needs to be predetermined. We use the Fuzzy Partition Coefficient (FPC) to determine the appropriate number of clusters. It evaluates how well the data is described by a certain model. FPC is given by

$$FPC = F(\mu) - \frac{\frac{1}{k}}{1 - \frac{1}{k}} \quad (11)$$

μ is the fuzzy user-cluster membership coefficient matrix and

$$F(\mu) = \frac{1}{C} \times \sum_{c=i}^{C} \sum_{i=1}^{m} {\mu^2}_{ic} \quad (12)$$

the coefficient varies from 0 (complete fuzzy clusters) i.e. users belonging to each group equally, to 1(hard clusters) i.e. each user belonging to a single cluster. Since we use PCA as a precursor to clustering, a plot of the number of principal components selected and the number of clusters (C) with fuzzy partition coefficient is shown in fig 3. As mentioned earlier, Fuzzy C-means clustering doesn't perform well for high dimensional data. This can be observed from the given plot as with the increase in the number of selected principal components, the fuzzy partition coefficient decreases rapidly. However, we lose a considerable amount of information if we reduce the number of components. The plot of variance ratio

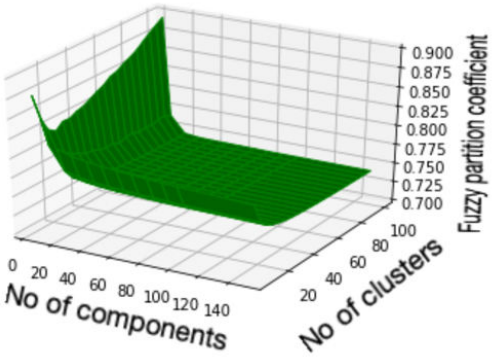


Fig. 3. Variation of FPC with number of Clusters and number of Principal Components for Fuzzy C-Means clustering.

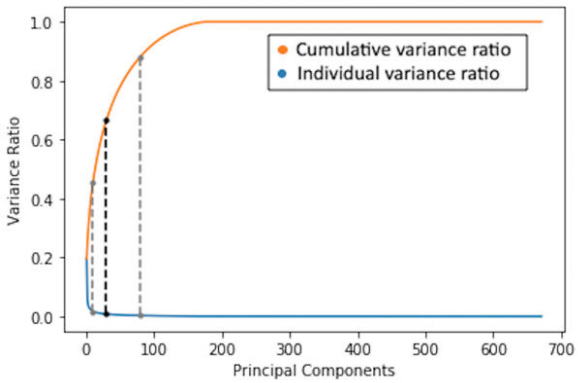


Figure 4. Plot of the individual variance ratio and cumulative various ratio explained by each principal component.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 31,2022 at 08:49:38 UTC from IEEE Xplore. Restrictions apply.

explained by each component and the cumulative variance ratio is shown in fig 4. Observing the plot carefully we see that an elbow occurs at the 10th principal component after which the variance explained by each component is not significant. However around 80 components give close to 90% of the information with low quality clusters. Table 1 shows the optimal parameters selected for our model.

TABLE I. OPTIMAL MODEL PARAMETERS

Parameters	**Value**
Principal Components	30
Number of Clusters (**C**)	10
Fuzzy Partition Coefficient (**FPC**)	0.79
Top K Similar Movies (**K**)	50
Hybrid weighing Factor (**β**)	0.4
Top-N Recommendations (**N**)	100

Accuracy metrics, Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) are first calculated as follows.

$$MAE = \frac{\sum_{i=1}^{N} |PreRat_{u,i} - ActRat_{u,i}|}{N} \quad (13)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (PreRat_{u,i} - ActRat_{u,i})^2}{N}} \quad (14)$$

PreRat, ActRat and N give the predicted rating, actual rating and the number of items rated by user u respectively.

The predictions given on the test set by fuzzy collaborative filtering and folksonomy are evaluated for the Accuracy metrics individually. The K (most similar movies) is varied and results are shown in fig 5. Since K doesn't affect the Collaborative User Rating (CUR), the value of Root Mean Square error and Mean Absolute error decreases with increasing K, with the optimal values occurring at K = 50. Further the hybrid approach is also evaluated against these metrics by changing the hybrid weighing factor, β, value as seen in fig 6. Clearly the Hybrid approach performs better

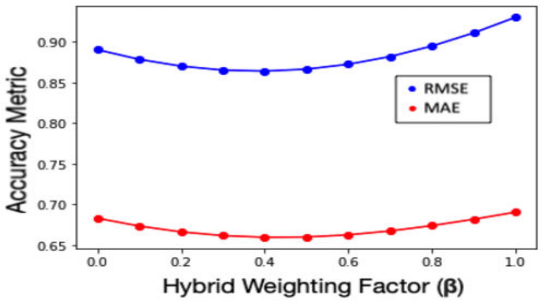


Fig. 5. Accuracy Metrics evaluation for hybrid approach.

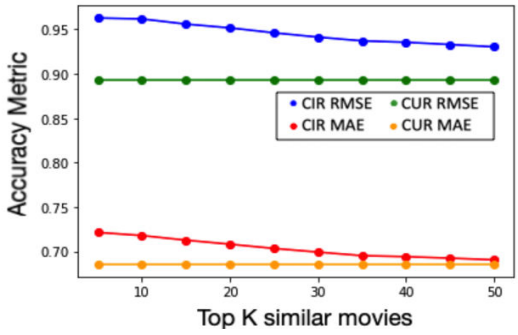


Fig. 6. Accuracy Metrics evaluation for constituent approaches.

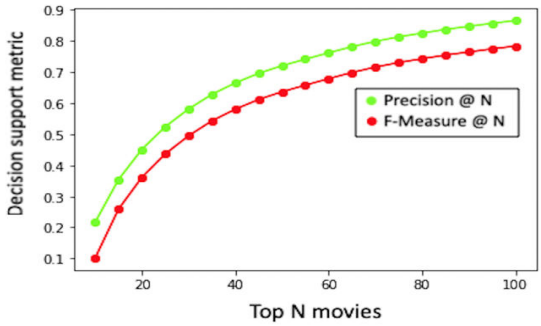


Fig. 7. Decision Support Metrics evaluation for hybrid approach.

than the individual approaches with the Collaborative User approach giving slightly better results than its counterpart. The best results are obtained for a β value of 0.4, weighted towards the better performing fuzzy clustering based approach.

TABLE II. MODEL EVALUATION

Approach	**F-measure**
Collaborative User Rating	0.6853
Collaborative Item Rating	0.6073
[20]Walek.Hybrid Approach	0.6912
[21] DNNRec Hybrid Approach	0.7256
[22] Geetha G,Hybrid Approach	0.7134
Our Hybrid Approach	**0.7829**

As accuracy metrics may lead to the imbalanced classification problem, the model is further evaluated for decision support metrics like Precision and F-measure. Rather than predicting the exact value, this help evaluate the quality of recommendations with respect to the items that are relevant to the users. These are calculated as follows

$$Recall = \frac{|RecIt \wedge RelIt|}{|RelIt|} \quad (15)$$

$$Precision = \frac{|RecIt \wedge RelIt|}{|RecIt|} \quad (16)$$

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 31,2022 at 08:49:38 UTC from IEEE Xplore. Restrictions apply.

RecIt and RelIt give the list of recommended and relevant items respectively.

A balance between recall and precision is obtained by computing the F1 score as the harmonic mean of precision and recall. It is given by:

$$F1 = \frac{2 \cdot Recall \cdot Precision}{Precision + Recall} \quad (17)$$

The F1 scores of various other RS approaches and Our RS approach are given in Table II. The table shows that performance can be improved by using hybrid approach[20][22] instead of using individual collaborative approaches. DNNRec[21] used Neural networks which showed an improvement over traditional hybrid methods. The table shows that our approach had better F1 score than DNNRec because in calculating latent relations for item and user while creating Neural networks, we lose the contextual structure of the domain. However, The Folksonomy based approach retains the contextual structure of the domain and provides better performance.

V. CONCLUSION & FUTURE WORK

Previous works have shown that incorporating knowledge-based and content-based approaches to collaborative filtering improve the performance of the system. However, the biased rating of popular items hampers the generation of quality clusters. In turn this affects the collaborative rating process.

In our model we approach this problem by forming soft clusters that assigns users to more than one cluster thus making clusters with fuzzy boundaries. This models the recommendation environment and thus helps achieve our goal. We further enhance our model by adding folksonomy which gives an idea of how users view different movies, thus identifying similar movies. Empirical evaluation of our approach confirms that combining the individual approaches improves the recommendation process. In terms of RMSE and MAE evaluations the hybrid approach gives us a better result. We achieve over 78% F1 score in our recommendation process which fulfils our primary aim of recommending relevant items to the user.

In our future work we will try to approach the clustering process in the presence of noise in the rating data and attempt to provide stable recommendations. Since the approach suggested is independent of the item domain, we assume that our approach would be helpful in other areas such as books, music, e-commerce, etc. Hence, in our future research we will explore domains other than movies.

REFERENCES

[1] Bennett J, Lanning S. The netflix prize. InProceedings of KDD cup and workshop 2007 Aug 12 (Vol. 2007, p. 35).

[2] Choudhury SS, Mohanty SN, Jagadev AK. Multimodal trust based recommender system with machine learning approaches for movie recommendation. International Journal of Information Technology. 2021 Apr;13(2):475-82.

[3] Soares M, Viana P. Tuning metadata for better movie content-based recommenda- tion systems. Multimedia Tools and Applications. 2015 Sep 1;74(17):7015-36.

[4] Luo X, Zhou M, Xia Y, Zhu Q. An efficient non-negative matrix-factorization-based approach to collaborative filtering for recommender systems. IEEE Transactions on Industrial Informatics. 2014 Feb 26;10(2):1273-84.

[5] Kabra A, Agarwal A, Parihar AS. Potent Real-Time Recommendations Using Multi-Model Contextual Reinforcement Learning. IEEE Transactions on Computational Social Systems. 2021 Aug 5.

[6] Tran T, Cohen R. Hybrid recommender systems for electronic commerce. InProc. Knowledge-Based Electronic Markets, Papers from the AAAI Workshop, Technical Report WS-00-04, AAAI Press 2000 Jul (Vol. 40).

[7] Logesh R, Subramaniyaswamy V. Exploring hybrid recommender systems for per- sonalized travel applications. InCognitive informatics and soft computing 2019 (pp. 535-544). Springer, Singapore.

[8] Nilashi M, bin Ibrahim O, Ithnin N, Sarmin NH. A multi-criteria collaborative filter- ing recommender system for the tourism domain using Expectation Maximization (EM) and PCA–ANFIS. Electronic Commerce Research and Applications. 2015 Oct 1;14(6):542-62.

[9] J¨aschke R, Marinho L, Hotho A, Schmidt-Thieme L, Stumme G. Tag recommen- dations in folksonomies. InEuropean conference on principles of data mining and knowledge discovery 2007 Sep 17 (pp. 506-514). Springer, Berlin, Heidelberg.

[10] Luo X, Xia Y, Zhu Q. Applying the learning rate adaptation to the matrix factoriza- tion based collaborative filtering. Knowledge-Based Systems. 2013 Jan 1;37:154-64.

[11] Neysiani BS, Soltani N, Mofidi R, Nadimi-Shahraki MH. Improve performance of association rule-based collaborative filtering recommendation systems using genetic algorithm. Int. J. Inf Technol. Comput. Sci.. 2019;2:48-55.

[12] Safoury L, Salah A. Exploiting user demographic attributes for solving cold-start problem in recommender system. Lecture Notes on Software Engineering. 2013 Aug 1;1(3):303-7.

[13] Semeraro G, Lops P, De Gemmis M, Musto C, Narducci F. A folksonomy-based recommender system for personalized access to digital artworks. Journal on Com- puting and Cultural Heritage (JOCCH). 2012 Oct 29;5(3):1-22.

[14] Said A, Kille B, De Luca EW, Albayrak S. Personalizing tags: a folksonomy-like approach for recommending movies. InProceedings of the 2nd International Work- shop on Information Heterogeneity and Fusion in Recommender Systems 2011 Oct 27 (pp. 53-56).

[15] Heymann P, Garcia-Molina H. Collaborative creation of communal hierarchical taxonomies in social tagging systems. Stanford; 2006.

[16] Ratsaby J. An algorithmic complexity interpretation of Lin's third law of informa- tion theory. Entropy. 2008 Mar;10(1):6-14.

[17] Bezdek JC, Ehrlich R, Full W. FCM: The fuzzy c-means clustering algorithm. Computers Geosciences. 1984 Jan 1;10(2-3):191-203.

[18] Wold S, Esbensen K, Geladi P. Principal component analysis. Chemometrics and intelligent laboratory systems. 1987 Aug 1;2(1-3):37-52.

[19] Harper FM, Konstan JA. The movielens datasets: History and context. Acm trans- actions on interactive intelligent systems (tiis). 2015 Dec 22;5(4):1-9

[20] Walek B, Fojtik V. A hybrid recommender system for recommending relevant movies using an expert system. Expert Systems with Applications. 2020 Nov 15;158:113452.

[21] Kiran R, Kumar P, Bhasker B. DNNRec: A novel deep learning based hybrid recommender system. Expert Systems with Applications. 2020 Apr 15;144:113054.

[22] Geetha G, Safa M, Fancy C, Saranya D. A hybrid approach using collaborative filtering and content based filtering for recommender system. InJournal of Physics: Conference Series 2018 Apr 1 (Vol. 1000, No. 1, p. 012101). IOP Publishing.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 31,2022 at 08:49:38 UTC from IEEE Xplore. Restrictions apply.

Afrika Matematika (2022) 33:9
https://doi.org/10.1007/s13370-021-00949-8

Generalized Durrmeyer operators based on inverse Pólya–Eggenberger distribution

Minakshi Dhamija[1] · Naokant Deo[2] · Ram Pratap[2] · Ana Maria Acu[3]

Received: 7 August 2019 / Accepted: 26 October 2021
© African Mathematical Union and Springer-Verlag GmbH Deutschland, ein Teil von Springer Nature 2021

Abstract
The paper aims on Durrmeyer variant of operators made by using inverse Pólya–Eggenberger distribution. For the proposed operators, we establish many approximation properties including asymptotic formula, degree of approximation and weighted approximation.

Keywords Stancu operators · Baskakov operator · Inverse Pólya–Eggenberger distribution

Mathematics Subject Classification 41A25 · 41A36

1 Introduction

In 1970, Stancu [19] came out with novel idea of inverse Pólya–Eggenberger distribution [8] and introduced generalized Baskakov operators $V_n^{(\nu)}$, for each real valued bounded function

Project financed by Lucian Blaga University of Sibiu & Hasso Plattner Foundation research grants LBUS-IRG-2021-07.

✉ Ana Maria Acu
acuana77@yahoo.com

Minakshi Dhamija
minakshidhamija11@gmail.com

Naokant Deo
naokantdeo@dce.ac.in; dr_naokant_deo@yahoo.com

Ram Pratap
rampratapiitr@gmail.com

[1] Department of Mathematics, Shaheed Rajguru College of Applied Sciences for Women, University of Delhi, Vasundhara Enclave, Delhi 110096, India

[2] Department of Applied Mathematics, Formerly Delhi College of Engineering, Delhi Technological University, Bawana Road, Delhi 110042, India

[3] Department of Mathematics and Informatics, Lucian Blaga University of Sibiu, Str. Dr. I. Ratiu, No.5-7, 550012 Sibiu, Romania

Published online: 24 January 2022

f defined on $I = [0, \infty)$:

$$V_n^{(\nu)}(f;x) = \sum_{s=0}^{\infty} w_{n,s}(x;\nu) f\left(\frac{s}{n}\right), \qquad x \geq 0,\ n \in \mathbb{N},\ \nu \geq 0, \tag{1.1}$$

where

$$w_{n,s}(x;\nu) = \binom{n+s-1}{s} \frac{\prod_{i=0}^{s-1}(x+i\nu)\prod_{j=0}^{n-1}(1+j\nu)}{\prod_{r=0}^{n+s-1}(1+x+r\nu)}.$$

The operators (1.1) have a special case for $\nu = 0$, which is well-known classical Baskakov operators [1]:

$$V_n(f;x) = \sum_{s=0}^{\infty} \binom{n+s-1}{s} \frac{x^s}{(1+x)^{n+s}} f\left(\frac{s}{n}\right). \tag{1.2}$$

Inspired from Stancu's work, in 2016, Deo et al. [3] considered Stancu–Kantorovich variant of operators (1.1) and discussed their convergence properties. In the same year, the integral variant of Stancu operators known as Jain–Durrmeyer operators taken in to account by Dhamija et al. [5]. For more detail about Durrmeyer-type operators we refer the readers to [9], [10]. The concept of Pólya and inverse Pólya–Eggenberger distribution has been used by many researchers in the field of approximation theory (see [4,6,11,14–17]).

In 2003, Srivastava and Gupta [18] proposed the sequence of linear positive operators:

$$L_{n,c}(f;x) = n\sum_{s=0}^{\infty} p_{n,s}(x,c) \int_0^{\infty} p_{n+c,s-1}(u,c) f(u)du + p_{n,0}(x,c) f(0), \tag{1.3}$$

where $p_{n,s}(x;c) = \frac{(-x)^s}{s!}\phi_{n,c}^{(s)}(x)$ and

$$\phi_{n,c}(x) = \begin{cases} e^{-nx} & c = 0, \\ (1+cx)^{-n/c}, & c = 1, 2, 3, \ldots \end{cases}$$

The special cases of (1.3) have been mentioned in [18]. Several researchers have shown interest in the different forms of the operators (1.3) (see [2,13]) and established the various convergence properties.

Motivated from the above stated work, in this paper, we consider a new kind of generalized Durrmeyer operators:

$$U_n^{(\nu)}(f;x) = (n-c)\sum_{s=0}^{\infty} w_{n,s}(x;\nu) \int_0^{\infty} p_{n,s}(u;c) f(u)du,\ x \in I,\ c = 0, 1, \ldots, \tag{1.4}$$

where the weight functions $p_{n,k}$ are defined as:

(i) $p_{n,s}(u;c) = \frac{1}{s!}\left\{\prod_{l=0}^{s-1}(n+cl)\right\} u^s (1+cu)^{-\frac{n+sc}{c}}$, $c > 0$. If $\nu = 0$, then operators (1.4) reduce to Baskakov–Durrmeyer operators.

(ii) $p_{n,s}(u;c) = e^{-nu}\frac{(nu)^s}{s!}$, $c = 0$. If $\nu = 0$, then operators (1.4) reduce to Baskakov–Szász–Durrmeyer operators.

2 Preliminary results

Suppose $\nu = \nu(n) \to 0$ as $n \to \infty$. Next, we give some results in order to prove our main results.

Lemma 2.1 [19] *The Stancu–Baskakov operators* (1.1) *verify*

$$V_n^{(\nu)}(1;x) = 1, \qquad V_n^{(\nu)}(u;x) = \frac{x}{1-\nu},$$

and

$$V_n^{(\nu)}\left(u^2;x\right) = \frac{1}{(1-\nu)(1-2\nu)}\left[x^2 + \frac{x(x+1)}{n} + \nu\left(1-\frac{1}{n}\right)x\right].$$

Lemma 2.2 *The proposed generalized Durrmeyer operators satisfy:*

$$U_n^{(\nu)}(1;x) = 1, \qquad U_n^{(\nu)}(u;x) = \frac{1}{n-2c} + \frac{nx}{(n-2c)(1-\nu)},$$

and

$$U_n^{(\nu)}(u^2;x) = \frac{1}{(n-2c)(n-3c)}\left\{\frac{n^2}{(1-\nu)(1-2\nu)}\left[x^2 + \frac{x(x+1)}{n} + \nu\left(1-\frac{1}{n}\right)x\right] + \frac{3nx}{(1-\nu)} + 2\right\}.$$

Lemma 2.3 *For each $n \in \mathbb{N}$, $x \in I$, we have*

(i) $U_n^{(\nu)}(u-x;x) = \frac{n\nu+2c(1-\nu)}{(n-2c)(1-\nu)}x + \frac{1}{n-2c}$.

(ii) $U_n^{(\nu)}((u-x)^2;x) = \frac{1}{(n-2c)(n-3c)}\left\{\frac{n^2}{(1-\nu)(1-2\nu)}\left[x^2 + \frac{x(x+1)}{n} + \nu\left(1-\frac{1}{n}\right)x\right] + \frac{3nx}{(1-\nu)} + 2\right\} + x^2 - 2x\left[\frac{1}{n-2c} + \frac{n}{n-2c}\cdot\frac{x}{1-\nu}\right]$.

Proof We can write

$$\begin{aligned} U_n^{(\nu)}(u-x;x) &= U_n^{(\nu)}(u;x) - x\,U_n^{(\nu)}(1;x) \\ &= \frac{1}{n-2c} + \frac{nx}{(n-2c)(1-\nu)} - x = \frac{1}{n-2c} + \left[\frac{n}{(n-2c)(1-\nu)} - 1\right]x \\ &= \frac{1}{n-2c} + \frac{n-(n-2c)(1-\nu)}{(n-2c)(1-\nu)}x = \frac{1}{n-2c} + \frac{n\nu+2c(1-\nu)}{(n-2c)(1-\nu)}x. \end{aligned}$$

The second central moment can be estimate by using the linearity property of the operators and Lemma 2.2, so here we skip the proof. □

Next, we develop a recurrence relation to find the higher order moments of the operators $U_n^{(\nu)}$ for which we shall use hypergeometric series:

$${}_2F_1(a,b;c;z) = \sum_{s=0}^{\infty}\frac{z^s}{s!}\cdot\frac{a^{[s,-1]}\cdot b^{[s,-1]}}{c^{[s,-1]}}, \tag{2.1}$$

where $a, b > 0$ and $a + b < c$.

Lemma 2.4 *Let $M_{n,r}^{(\nu)}(x) = U_n^{(\nu)}(u^r; x)$. Then, we have*

$$M_{n,r+1}^{(\nu)}(x) = \frac{r^2\left(r+1-\frac{x}{\nu}-\frac{1}{\nu}-n\right)}{c^2\left(\frac{n}{c}-r-1\right)\left(\frac{n}{c}-r-2\right)\left(\frac{1}{\nu}-r-1\right)} M_{n,r-1}^{(\nu)}(x) + \frac{rn+\frac{2r}{\nu}+\frac{rx}{\nu}+\frac{1}{\nu}+\frac{nx}{\nu}-2r^2-3r-1}{c\left(\frac{n}{c}-r-2\right)\left(\frac{1}{\nu}-r-1\right)} M_{n,r}^{(\nu)}(x).$$

Proof We have

$$(n-c)\int_0^\infty p_{n,s}(u;c)u^r\,du = \frac{(r+1)_s\Gamma(r+1)\Gamma\left(\frac{n}{c}-r-1\right)}{c^r s!\Gamma\left(\frac{n}{c}-1\right)}.$$

Therefore

$$\begin{aligned} M_{n,r}^{(\nu)}(x) &= \sum_{s=0}^{\infty} w_{n,s}(x;\nu)\frac{(r+1)_s\Gamma(r+1)\Gamma\left(\frac{n}{c}-r-1\right)}{c^r s!\Gamma\left(\frac{n}{c}-1\right)} \\ &= \frac{\left(\frac{1}{\nu}\right)_n\Gamma(r+1)\Gamma\left(\frac{n}{c}-r-1\right)}{c^r\Gamma\left(\frac{n}{c}-1\right)\left(\frac{x}{\nu}+\frac{1}{\nu}\right)_n}\sum_{s=0}^{\infty}\frac{(n)_s(r+1)_s\left(\frac{x}{\nu}\right)_s}{s!\left(\frac{x}{\nu}+\frac{1}{\nu}+n\right)_s(1)_s} \\ &= \frac{\left(\frac{1}{\nu}\right)_n\Gamma(r+1)\Gamma\left(\frac{n}{c}-r-1\right)}{c^r\Gamma\left(\frac{n}{c}-1\right)\left(\frac{x}{\nu}+\frac{1}{\nu}\right)_n}\,{}_3F_2\left(r+1,n,\frac{x}{\nu};\frac{x}{\nu}+\frac{1}{\nu}+n,1;1\right). \end{aligned}$$

Thus,

$${}_3F_2\left(r+1,n,\frac{x}{\nu};\frac{x}{\nu}+\frac{1}{\nu}+n,1;1\right) = \frac{c^r\Gamma\left(\frac{n}{c}-1\right)\left(\frac{x}{\nu}+\frac{1}{\nu}\right)_n}{\left(\frac{1}{\nu}\right)_n\Gamma(r+1)\Gamma\left(\frac{n}{c}-r-1\right)}M_{n,r}^{(\nu)}(x). \quad (2.2)$$

Let $p_1 = r+1$, $p_2 = n$, $p_3 = \frac{x}{\nu}$, $q_1 = \frac{x}{\nu}+\frac{1}{\nu}+n$, $q_2 = 1$, $z = 1$. Using (2.2) in the below equality

$$\begin{aligned} &p_1(p_1+1)(1-z){}_3F_2(p_1+2,p_2,p_3;q_1,q_2;z) \\ &\quad + p_1(q_1+q_2-3p_1-2+z(2p_1-p_2-p_3+1)){}_3F_2(p_1+1,p_2,p_3;q_1,q_2;z) \\ &\quad + \left[(2p_1-q_1)(2p_1-q_2)+p_1-p_1^2-z(p_1-p_2)(p_1-p_3)\right]{}_3F_2(p_1,p_2,p_3;q_1,q_2;z) \\ &\quad - (p_1-q_1)(p_1-q_2){}_3F_2(p_1-1,p_2,p_3;q_1,q_2;z) = 0, \end{aligned}$$

we get the desired result. □

Lemma 2.5 *For the proposed operators (1.4), the higher order moments are as follows:*

i)

$$\begin{aligned} U_n^{(\nu)}(u^3;x) &= -\frac{1}{(1-3\nu)(1-2\nu)(1-\nu)(n-2c)(n-3c)(n-4c)} \\ &\cdot\left\{-x(\nu+x)(2\nu+x)n^3+3x(\nu+x)(7\nu-x-3)n^2-x(18-21\nu x-81\nu\right. \\ &\left.+2x^2+85\nu^2+9x)n+6(\nu-1)(3\nu-1)(2\nu-1)\right\}; \end{aligned}$$

ii)

$$U_n^{(\nu)}(u^4;x) = \frac{1}{(1-4\nu)(1-3\nu)(1-2\nu)(1-\nu)(n-2c)(n-3c)(n-4c)(n-5c)}$$

$$\left\{x(3\nu+x)(2\nu+x)(\nu+x)n^4-2x(\nu+x)(-3x-8+23\nu)(2\nu+x)n^3+x(\nu+x)\right.$$
$$(546\nu^2-480\nu-137\nu x+11x^2+72+48x)n^2-2x(-16x^2-1028\nu^2+204\nu x$$
$$+396\nu-3x^3+46\nu x^2-48+830\nu^3-273\nu^2x-36x)n+24(4\nu-1)$$
$$\left.(3\nu-1)(2\nu-1)(\nu-1)\right\}.$$

Proof Here we find the higher order moments with the help of Lemma 2.2 and the recurrence relation which is stated in Lemma 2.4. □

Lemma 2.6 *For the operators* (1.4)*, the fourth central moment is given as:*

$$U_n^{(\nu)}((u-x)^4;x)=\frac{1}{(1-4\nu)(1-3\nu)(1-2\nu)(1-\nu)(n-2c)(n-3c)(n-4c)(n-5c)}$$
$$\left\{\nu^2x\left(6x^2+72\nu x^2+6\nu+24x^3\nu^2+32\nu x+3x^2+46x^3\nu+3x\right)n^4\right.$$
$$-2\nu x(-3x^3c-36x^3\nu-3x^3+226x^3\nu^2c-55x^3\nu c+168x^3\nu^3c-69\nu x^2c$$
$$+48x^2\nu^3+152x^2\nu^2-9x^2-3x^2c-77\nu x^2-20x\nu c-51\nu x-6x$$
$$\left.+324x^2\nu^2c+168\nu^2x+80x\nu^2c-16\nu+46\nu^2)n^3+O(n^2)\right\}.$$

Proof We find the fourth central moment by using the linearity property of the operators and Lemmas 2.2, 2.5. □

Lemma 2.7 *For each* $n\in\mathbb{N}$, $x\in I$, *we have*

$$\left|U_n^{(\nu)}(f;x)\right|\le\|f\|.$$

Proof From (1.4), we have

$$\left|U_n^{(\nu)}(f;x)\right|=\left|(n-c)\sum_{s=0}^{\infty}w_{n,s}(x;\nu)\int_0^{\infty}p_{n,s}(u;c)f(u)du\right|$$
$$\le(n-c)\sum_{s=0}^{\infty}w_{n,s}(x;\nu)\int_0^{\infty}\left|p_{n,s}(u;c)\right||f(u)|\,du$$
$$\le U_n^{(\nu)}(1;x)\,\|f\|=\|f\|.$$

□

3 Local approximation

Let $C_B(I)$ be the space of all real valued continuous and bounded functions on I and $\|f\|=\sup_{x\in I}|f(x)|$. The Peetre's K−functional is defined as

$$K_2(f;\delta):=\inf_{x\in C_B^2(I)}\left\{\|f-g\|+\delta\left\|g''\right\|\right\},\ \delta>0,$$

where $f\in C_B(I)$ and $C_B^2(I)=\left\{g\in C_B(I):g',g''\in C_B(I)\right\}$. It is well known that there exists an absolute constant $C>0$ such that (see [7, p.177. Theorem 2.4])

$$K_2(f;\delta)\le C\omega_2\left(f;\sqrt{\delta}\right),\tag{3.1}$$

where $\omega_2\left(f;\sqrt{\delta}\right) = \sup_{0<h<\sqrt{\delta}} \sup_{x\in I} |f(x+2h) - 2f(x+h) + f(x)|$ is second order modulus of continuity.

Theorem 3.1 *Let $f \in C_B(I)$. Then*

$$\left|U_n^{(\nu)}(f;x) - f(x)\right| \leq 2\omega(f;\delta),$$

where $\delta = \left[U_n^{(\nu)}\left((u-x)^2;x\right)\right]^{\frac{1}{2}}$ and $\omega(f;\delta)$ is the usual modulus of continuity.

Proof Since

$$|(f(x) - f(y)| \leq 2\omega(f;|x-y|) \leq \left(1 + \frac{1}{\delta}|x-y|\right)\omega(f,\delta).$$

Therefore

$$\begin{aligned}\left|U_n^{(\nu)}(f;x) - f(x)\right| &\leq (n-c)\sum_{s=0}^{\infty} w_{n,s}(x;\nu)\int_0^{\infty} p_{n,s}(u;c)\,|f(u)-f(x)|du \\ &\leq \left[1 + \frac{1}{\delta}(n-c)\sum_{s=0}^{\infty} w_{n,s}(x;\nu)\int_0^{\infty} p_{n,s}(u;c)\,|u-x|du\right]\omega(f;\delta).\end{aligned}$$

Applying Cauchy–Schwarz inequality for integral, we get

$$\begin{aligned}\left|U_n^{(\nu)}(f;x) - f(x)\right| \leq \Bigg[1 + \frac{1}{\delta}(n-c)\sum_{s=0}^{\infty} w_{n,s}(x;\nu)\left(\int_0^{\infty} p_{n,s}(u;c)du\right)^{\frac{1}{2}} \\ \times \left(\int_0^{\infty} p_{n,s}(u;c)(u-x)^2dt\right)^{\frac{1}{2}}\Bigg]\omega(f;\delta).\end{aligned}$$

Again, applying Cauchy–Schwarz inequality for sum, it follows

$$\begin{aligned}\left|U_n^{(\nu)}(f;x) - f(x)\right| &\leq \Bigg[1 + \frac{1}{\delta}\Bigg((n-c)\sum_{s=0}^{\infty} w_{n,s}(x;\nu)\int_0^{\infty} p_{n,s}(u;c)du\Bigg)^{\frac{1}{2}} \\ &\quad \times \left((n-c)\sum_{s=0}^{\infty} w_{n,s}(x;\nu)\int_0^{\infty} p_{n,s}(u;c)(u-x)^2du\right)^{\frac{1}{2}}\Bigg]\omega(f;\delta) \\ &\leq \left[1 + \frac{1}{\delta}\left(U_n^{(\nu)}(1;x)\right)^{\frac{1}{2}}\left(U_n^{(\nu)}(u-x)^2;x\right)^{\frac{1}{2}}\right]\omega(f;\delta) = 2\omega(f;\delta),\end{aligned}$$

where

$$\delta = \left[U_n^{(\nu)}\left((u-x)^2;x\right)\right]^{1/2}.$$

□

Theorem 3.2 *Let $f \in C_B(I)$. Then*

$$\left|U_n^{(\nu)}(f;x) - f(x)\right| \leq M\omega_2\left(f; \frac{\delta_n(x,\nu)}{2}\right) + \omega(f;\delta_\omega),$$

where $M > 0$ is a constant,

$$\delta_n(x;\nu) = \left(U_n^{(\nu)}((u-x)^2;x)\right) + \left(U_n^{(\nu)}((u-x);x)^2\right)^{\frac{1}{2}}$$

and

$$\delta_\omega = U_n^{(\nu)}((u-x);x).$$

Proof For $x \in I$, define the operators

$$\tilde{U}_n^{(\nu)}(f;x) = U_n^{(\nu)}(f;x) - f\left(\frac{nx+1-\nu}{(n-2c)(1-\nu)}\right) + f(x). \tag{3.2}$$

Here

$$\tilde{U}_n^{(\nu)}(1;x) = 1, \quad \tilde{U}_n^{(\nu)}(u;x) = x$$

and

$$\tilde{U}_n^{(\nu)}(u-x;x) = 0. \tag{3.3}$$

Let $g \in C_B^2(I)$ and $x, u \in I$. Then by Taylor's formula, we get

$$\tilde{U}_n^{(\nu)}(g;x) - g(x) = g'(x)\tilde{U}_n^{(\nu)}(u-x;x) + \tilde{U}_n^{(\nu)}\left(\int_x^u (u-t)g''(t)dt;x\right).$$

From (3.2) and (3.3), we obtain

$$\begin{aligned}\tilde{U}_n^{(\nu)}(g;x) - g(x) &= U_n^{(\nu)}\left(\int_x^u (u-t)g''(t)dt;x\right)\\ &\quad - \int_x^{\frac{nx+1-\nu}{(n-2c)(1-\nu)}} \left(\frac{nx+1-\nu}{(n-2c)(1-\nu)} - t\right)g''(t)dt.\end{aligned}$$

Since

$$\left|\int_x^u (u-t)g''(t)\right| \leq (u-x)^2 \left\|g''\right\|.$$

Therefore

$$\left|\tilde{U}_n^{(\nu)}(g;x) - g(x)\right| \leq \left[U_n^{(\nu)}((u-x)^2;x) + U_n^{(\nu)}((u-x);x)^2\right]\left\|g''\right\|.$$

Now

$$\begin{aligned}\left|U_n^{(\nu)}(f;x) - f(x)\right| &= \left|\tilde{U}_n^{(\nu)}(f;x) - f(x) + f\left(\frac{nx+1-\nu}{(n-2c)(1-\nu)}\right) - f(x)\right|\\ &\leq \left|\tilde{U}_n^{(\nu)}(f-g;x)\right| + \left|\tilde{U}_n^{(\nu)}(g;x) - g(x)\right|\end{aligned}$$

$$+ |g(x) - f(x)| + \left| f\left(\frac{nx + 1 - \nu}{(n-2c)(1-\nu)}\right) - f(x) \right|$$
$$\leq 4\|f - g\| + \delta_n^2(x, \nu)\left\|g''\right\| + \omega(f, \delta_\omega). \tag{3.4}$$

Taking infimum of (3.4) over $g \in C_B^2(I)$.

$$\left|U_n^{(\nu)}((u-x); x)\right| \leq 4K_2\left(f; \frac{\delta_n^2(x;\nu)}{4}\right) + \omega(f, \delta_\omega)$$
$$\leq M\omega_2\left(f; \frac{\delta_n(x;\nu)}{2}\right) + \omega(f, \delta_\omega).$$

Hence the proved. □

4 Voronovskaja type theorem

Let $C_2(I) = \left\{f \in C(I), |f(u)| \leq M(1+u^2), \text{ for } M > 0\right\}$, provided with the norm

$$\|f\|_2 := \sup_{x \geq 0} \frac{|f(x)|}{1+x^2}.$$

Lemma 4.1 *If $\nu = \nu(n) \to 0$, as $n \to \infty$ and $\lim_{n\to\infty} n\nu(n) = l$, then*

i) $\lim_{n\to\infty} nU_n^{(\nu)}(u-x; x) = 2cx + lx + 1$;
ii) $\lim_{n\to\infty} nU_n^{(\nu)}((u-x)^2; x) = lx(x+1) + x(cx + x + 2)$;
iii) $\lim_{n\to\infty} n^2U_n^{(\nu)}((u-x)^4; x) = 6lx(x^2c + x^2 + 3x + xc + 2)$.

Theorem 4.1 *Let $f \in C_2(I)$ and ω_b be its modulus of continuity on the finite interval $[0, b]$, $b > 0$. Then*

$$||U_n^{(\nu)}(f) - f||_{C[0,b]} \leq 4M(1+b^2)\eta_n(b) + 2\omega_{b+1}(f, \sqrt{\eta_n(b)}),$$

where $\eta_n(b) = \max_{x\in[0,b]} U_n^{(\nu)}\left((u-x)^2; x\right)$ and $||\cdot||_{C[0,b]}$ denotes the sup norm on $C[0, b]$.

Proof For any $\delta > 0$, $x \in [0, b]$ and $u \geq 0$, the following inequality

$$|f(u) - f(x)| \leq 4M(1+x^2)(u-x)^2 + \left(1 + \frac{|u-x|}{\delta}\right)\omega_{b+1}(f; \delta), \tag{4.1}$$

holds (see [12]). Applying the operator $U_n^{(\nu)}(.; x)$ and using Cauchy–Schwarz inequality we get

$$|U_n^{(\nu)}(f; x) - f(x)| \leq 4M(1+b^2)U_n^{(\nu)}((u-x)^2; x) + \left(1 + \frac{1}{\delta}U_n^{(\nu)}(|u-x|; x)\right)\omega_{b+1}(f, \delta)$$
$$\leq 4M(1+b^2)\eta_n(b) + \omega_{b+1}(f, \delta)\left(1 + \frac{1}{\delta}\sqrt{\eta_n(b)}\right).$$

Choosing $\delta = \sqrt{\eta_n(b)}$, the desired result follows. □

Theorem 4.2 *Let $f \in C_2(I)$ and $\nu = \nu(n) \to 0$, as $n \to \infty$. If there exists first and second derivatives of the function f in a fixed point $x \in I$ and $\lim_{n\to\infty} n\nu(n) = l \in \mathbb{R}$, then*

$$\lim_{n\to\infty} n\left[U_n^{(\nu)}(f; x) - f(x)\right] = (2cx + lx + 1)f'(x) + [lx(x+1) + x(cx + x + 2)]f''(x),$$

uniformly with respect to $x \in [0, a]$, $a > 0$.

Proof Applying Taylor's expansion, we get

$$U_n^{(\nu)}(f;x) - f(x) = f'(x)U_n^{(\nu)}(u-x;x) + f''(x)U_n^{(\nu)}\left(\frac{(u-x)^2}{2!};x\right) + U_n^{(\nu)}\left(\varepsilon(u,x)(u-x)^2;x\right).$$

where $\varepsilon(u,x)$ is bounded function and $\lim_{u\to x}\varepsilon(u,x)=0$.

Using Lemma 4.1 we obtain

$$\lim_{n\to\infty} n\left[U_n^{(\nu)}(f;x) - f(x)\right] = (2cx+lx+1)f'(x) + [lx(x+1)+x(cx+x+2)]f''(x) + \lim_{n\to\infty} U_n^{(\nu)}\left(\varepsilon(u,x)(u-x)^2;x\right). \tag{4.2}$$

Now

$$U_n^{(\nu)}\left(\varepsilon(u,x)(u-x)^2;x\right) \le \sqrt{U_n^{(\nu)}\left(\varepsilon^2(u,x);x\right)}\sqrt{U_n^{(\nu)}\left((u-x)^4;x\right)}. \tag{4.3}$$

Since $\varepsilon(u,x)\to 0$ as $u\to x$, it follows from Theorem 4.1 that

$$\lim_{n\to\infty} U_n^{(\nu)}\left(\varepsilon^2(u,x);x\right) = \varepsilon^2(x,x) = 0, \tag{4.4}$$

uniformly with respect to $x\in[0,a]$. From (4.3) and (4.4) and Lemma 4.1, we get require result. □

5 Weighted approximation

In the case of boundless function defined on noncompact interval I, the computation of the rate of convergence is not valid for the operators (1.4). Therefore in present section, we study the approximation properties for the proposed operators using weighted space.

Let $C_2^*(I) = \{f\in C_2(I) : \lim_{x\to\infty}\frac{f(x)}{1+x^2} \text{ exists}\}$. For $f\in C_2^*(I)$, the weighted modulus of continuity is given by (see [20])

$$\Omega(f,\delta) = \sup_{x\ge 0}\sup_{0<|h|\le\delta}\frac{|f(x+h)-f(x)|}{1+(x+h)^2}.$$

The weighted modulus of continuity $\Omega(f,\delta)$ satisfies the following properties:

Lemma 5.1 [20] *If $f\in C_2^*(I)$, then*

(1) *$\Omega(f,\delta)$ is a monotone increasing function of δ;*
(2) $\lim_{\delta\to 0^+}\Omega(f,\delta)=0$;
(3) *for any $\lambda\in[0,\infty)$, $\Omega(f,\lambda\delta)\le(1+\lambda)\Omega(f,\delta)$.*

Theorem 5.1 *For each $f\in C_2^*(I)$ and $\nu>0$, we have*

$$\lim_{n\to\infty}\sup_{x\in[0,\infty)}\frac{|U_n^{(\nu)}(f;x)-f(x)|}{(1+x^2)^{1+\nu}} = 0.$$

Proof Let $x_0\in I$ be arbitrary but fixed. Then,

$$\sup_{x\in I}\frac{|U_n^{(\nu)}(f;x)-f(x)|}{(1+x^2)^{1+\nu}} \le \sup_{x\le x_0}\frac{|U_n^{(\nu)}(f;x)-f(x)|}{(1+x^2)^{1+\nu}} + \sup_{x>x_0}\frac{|U_n^{(\nu)}(f;x)-f(x)|}{(1+x^2)^{1+\nu}}$$

$$\leq ||U_n^{(\nu)}(f) - f||_{C[0,x_0]} + ||f||_2 \sup_{x>x_0} \frac{|U_n^{(\nu)}(1+t^2;x)|}{(1+x^2)^{1+\nu}} + \sup_{x>x_0} \frac{|f(x)|}{(1+x^2)^{1+\nu}}.$$

Since $|f(x)| \leq ||f||_2(1+x^2)$, we have $\sup_{x>x_0} \frac{|f(x)|}{(1+x^2)^{1+\nu}} \leq \frac{||f||_2}{(1+x_0^2)^{\nu}}$.
Let $\epsilon > 0$ be arbitrary. We can choose x_0 to be so large that

$$\frac{||f||_2}{(1+x_0^2)^{\nu}} < \frac{\epsilon}{6}. \tag{5.1}$$

Since $\lim_{n\to\infty} \sup_{x>x_0} \frac{U_n^{(\nu)}(1+u^2;x)}{1+x^2} = 1$, it follows that

$$\sup_{x>x_0} \frac{U_n^{(\nu)}(1+u^2;x)}{1+x^2} \leq \frac{(1+x_0^2)^a}{\|f\|_2}\frac{\epsilon}{3} + 1,$$

for sufficiently large n. Therefore,

$$\|f\|_2 \sup_{x>x_0} \frac{U_n^{(\nu)}(1+u^2;x)}{(1+x^2)^{\nu+1}} \leq \frac{\|f\|_2}{(1+x_0^2)^{\nu}} \sup_{x>x_0} \frac{U_n^{(\nu)}(1+u^2;x)}{(1+x^2)} \leq \frac{\epsilon}{3} + \frac{||f||_2}{(1+x_0^2)^{\nu}}. \tag{5.2}$$

Using Theorem 4.1, we can find for sufficiently large n

$$||U_n^{(\nu)}(f;x) - f(x)||_{C[0,x_0]} < \frac{\epsilon}{3}. \tag{5.3}$$

Combining the relations (5.1)–(5.3), we obtain

$$\sup_{x\in[0,\infty)} \frac{\left|U_n^{(\nu)}(f;x) - f(x)\right|}{(1+x^2)^{1+\nu}} < \epsilon.$$

□

Theorem 5.2 *Let $f \in C_2^*(I)$. Then for sufficiently large n, we have*

$$\sup_{x\in[0,\infty)} \frac{\left|U_n^{(\nu)}(f;x) - f(x)\right|}{(1+x^2)^{\frac{5}{2}}} \leq \tilde{C}\Omega\left(f; \frac{1}{\sqrt{n}}\right), \tag{5.4}$$

where $\tilde{C} > 0$ is a constant.

Proof For $x \in I$ and $\delta > 0$, by the definition of the weighted modulus of continuity and Lemma 5.1, we get

$$\begin{aligned} |f(u) - f(x)| &\leq \left(1 + (x + |x-u|)^2\right)\Omega(f; |u-x|) \\ &\leq 2(1+x^2)(1+(u-x)^2)\left(1 + \frac{|u-x|}{\delta}\right)\Omega(f;\delta). \end{aligned}$$

Applying $U_n^{(\nu)}(\cdot;x)$ both sides, we obtain

$$\begin{aligned} |U_n^{(\nu)}(f;x) - f(x)| &\leq 2(1+x^2)\Omega(f;\delta) \\ &\times\left(1 + U_n^{(\nu)}((u-x)^2;x) + U_n^{(\nu)}\left((1+(u-x)^2)\frac{|u-x|}{\delta};x\right)\right). \end{aligned} \tag{5.5}$$

From Lemma 4.1, for sufficiently large n, it follows

$$nU_n^{(\nu)}((u-x)^2;x) \leq C(1+x^2) \text{ and } n^2U_n^{(\nu)}((u-x)^4;x) \leq C(1+x^2)^2, \tag{5.6}$$

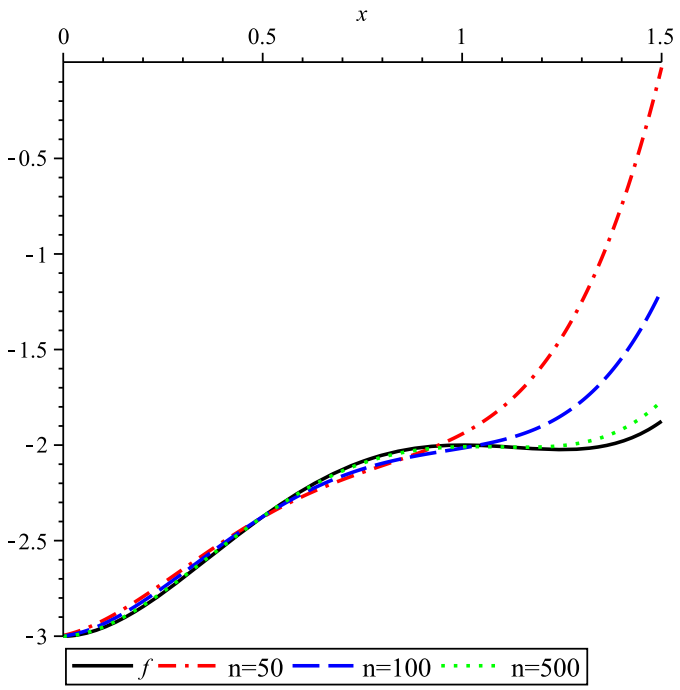


Fig. 1 Convergence of $U_n^{(\nu)}(f;x)$ to $f(x)$

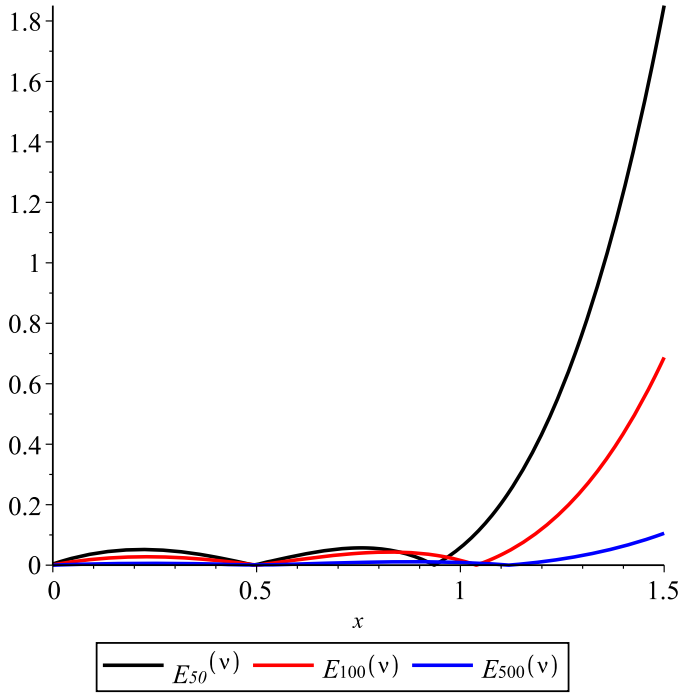


Fig. 2 Error of approximation $E_n^{(\nu)}$

where $C > 0$ is a constant.
Applying the Cauchy–Schwarz inequality in (5.5), we get

$$U_n^{(\nu)}\left((1+(u-x)^2)\frac{|u-x|}{\delta};x\right) \le \frac{1}{\delta}\left(U_n^{(\nu)}((u-x)^2;x)\right)^{1/2} + \frac{1}{\delta}\left(U_n^{(\nu)}((u-x)^4;x)\right)^{1/2}\left(U_n^{(\nu)}((u-x)^2(x);x)\right)^{1/2}. \tag{5.7}$$

Combining the estimates (5.5)–(5.7) and taking

$$\tilde{C} = 2\left(1+\sqrt{C}+2C\right) \text{ and } \delta = \frac{1}{\sqrt{n}},$$

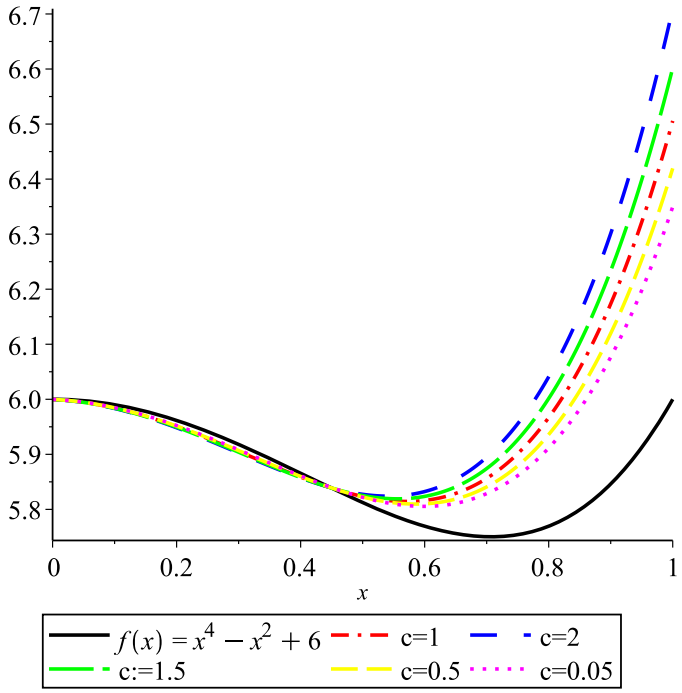


Fig. 3 Convergence of $U_n^{(\nu)}(f;x)$ to $f(x)$

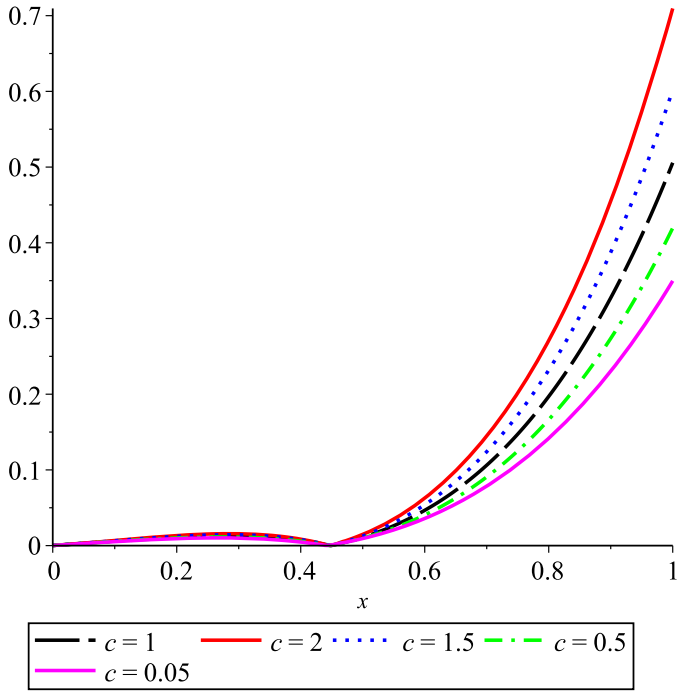


Fig. 4 Error of approximation $E_n^{(\nu)}$

we obtain the required result. □

Example 1 Set $E_n^{(\nu)}(f;x) = |U_n^{(\nu)}(f;x) - f(x)|$, the absolute error of the operators $U_n^{(\nu)}(f;x)$ with the function $f(x)$. The convergence of the operators $U_n^{(\nu)}(f;x)$ is shown in Fig. 1, where $f(x) = 2x^4 - 6x^3 + 5x^2 - 3$, $c = 0.02$, $\nu = 1/n$, $n \in \{50, 100, 500\}$. Remark that for decreasing values of ν, the operator $U_n^{(\nu)}(f;x)$ converges faster to the function f. The absolute error $E_n^{(\nu)}(f;x)$ is illustrated in Fig. 2.

Example 2 Let $n = 100$, $\nu = \dfrac{1}{n}$, $c \in \{0.05, 0.5, 1, 1.5, 2\}$. The convergence of the operators $U_n^{(\nu)}$ to the function $f(x) = x^4 - x^2 + 6$ is illustrated in Figs. 3. Remark that the operators $U_n^{(\nu)}$

presents better order of approximation for smaller values of c. The absolute error $E_n^{(\nu)}(f;x)$ is shown in Fig. 4.

References

1. Baskakov, V.A.: A sequence of linear positive operators in the space of continuous functions. Dokl. Acad. Nauk. SSSR **113**, 249–251 (1957)
2. Deo, N.: Faster rate of convergence on Srivastava–Gupta operators. Appl. Math. Comput. **218**(21), 10486–10491 (2012)
3. Deo, N., Dhamija, M., Miclăuş, D.: Stancu–Kantorovich operators based on inverse Pólya–Eggenberger distribution. Appl. Math. Comput. **273**, 281–289 (2016)
4. Deo, N., Dhamija, M.: Generalized positive linear operators based on PED and IPED. Iran J. Sci. Technol. Trans. Sci. (2018). https://doi.org/10.1007/s40995-017-0477-5
5. Dhamija, M., Deo, N.: Jain–Durrmeyer operators associated with the inverse Pólya–Eggenberger distribution. Appl. Math. Comput. **286**, 15–22 (2016)
6. Dhamija, M., Deo, N.: Approximation by generalized positive linear-Kantorovich operators. Filomat **31**(14), 4353–4368 (2017)
7. DeVore, R.A., Lorentz, G.G.: Constructive Approximation. Springer, Berlin (1993)
8. Eggenberger, F., Pólya, G.: Über die Statistik verkerter Vorgänge. Z. Angew. Math. Mech. **1**, 279–289 (1923)
9. Gupta, V., Greubel, G.C.: Moment estimations of new Szász–Mirakyan–Durrmeyer operators. Appl. Math. Comput. **271**, 540–547 (2015)
10. Gupta, V., Aral, A.: Bernstein Durrmeyer operators based on two parameters, Facta Universitatis (Niş). Ser. Math. Inform. **31**(1), 79–95 (2016)
11. Gupta, V., Acu, A.M., Sofonea, D.F.: Approximation Baskakov type Polya–Durrmeyer operators. Appl. Math. Comput. **294**(1), 318–331 (2017)
12. İbekli, E., Gadjieva, E.A.: The order of approximation of some unbounded functions by the sequences of positive linear operators. Turk. J. Math. **19**(3), 331–337 (1995)
13. Ispir, N., Yuksel, I.: On the Bezier variant of Srivastava–Gupta operators. Appl. Math. E Notes **5**, 129–137 (2005)
14. Kajla, A., Acu, A.M., Agrawal, P.N.: Baskakov–Szász-type operators based on inverse Pólya–Eggenberger distribution. Ann. Funct. Anal. **8**(1), 106–123 (2017)
15. Lupaş, L., Lupaş, A.: Polynomials of binomial type and approximation operators. Stud. Univ. Babeş Bolyai Math. **32**(4), 61–69 (1987)
16. Miclăuş, D.: The revision of some results for Bernstein–Stancu type operators. Carpathian J. Math. **28**(2), 289–300 (2012)
17. Neer, T., Acu, A.M., Agrawal, P.N.: Bezier variant of genuine-Durrmeyer type operators based on Polya distribution. Carpathian J. Math. **33**(1), 73–86 (2017)
18. Srivastava, H.M., Gupta, V.: A certain family of summation integral-type operators. Math. Comput. Model. **27**, 1307–1315 (2003)
19. Stancu, D.D.: Two classes of positive linear operators. Anal. Univ. Timişoara Ser. Mat. **8**, 213–220 (1970)
20. Yüksel, I., Ispir, N.: Weighted approximation by a certain family of summation integral-type operators. Comput. Math. Appl. **52**(10–11), 1463–1470 (2006)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Check for updates

Original conceptual / theoretical manuscript

HR factors for the successful implementation of Industry 4.0: A systematic literature review

Journal of General Management
2022, Vol. 47(2) 73–85
© The Author(s) 2021
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/03063070211019141
journals.sagepub.com/home/gem
SAGE

Anju Verma and Murugesan Venkatesan
Indian Institute of Foreign Trade, India

Abstract
As the world moves deeper into the fourth industrial revolution, organizations around the world are polarizing towards the transformation and adoption of the advanced technologies of Industry 4.0. However, recent studies suggest that organizations struggle and face several managerial challenges in their implementation. As a consequence, many studies examined the potential factors that can influence Industry 4.0 implementation. Based on a systematic literature review, this study has tried to explore, examine, and synthesize the potential human resource (HR) success factors and develop an efficient Industry 4.0 HR framework that can play a vital role in Industry 4.0 implementation. This study would also build a solid foundation for practitioners, academicians, and researchers to get a better understanding of the impact of Industry 4.0 on HR management.

Keywords
fourth industrial revolution, HR factors, HR 4.0., human resource management challenges, smart human resource management

Introduction

As fourth industrial revolution unfolds, driverless cars, self-driving forklifts, screws communicating with robots, and smart machines coordinating the production processes have become a reality. Organizations today are making supreme efforts to harness the advanced technologies to develop new and efficient processes, products, and services by blending the physical and digital realms of the world (World Economic Forum, 2016). These emerging technologies which rely heavily on automation, digitalization, and real-time information exchange include cyber-physical systems, internet of things (Kagermann et al., 2013), robotics, big data analysis, artificial intelligence (AI), augmented reality (AR) and virtual reality (VR), three-dimensional printers, cloud computing, etc (Irudayaraj, 2017; Kagermann et al., 2013; Lasi et al., 2014; Thoben et al., 2017).

The fact that Industry 4.0 requires a large number of processes to be altered and replaced by new ones in the manufacturing environments, structural transformations are expected in the global economy (AIMA-KPMG, 2018; Lorenz et al., 2015). Developed countries like Germany, United States, China, and Japan have already initiated the transformation and are formulating national strategies and policies for development (BRICS, 2016; Dhanpat et al., 2020; Kazancoglu and Ozkan-Ozen, 2018). An increasing number of developing countries like India and Indonesia to reap benefits and compete with developed countries are also polarizing towards Industry 4.0 by making comprehensive changes throughout the organizations (Irudayaraj, 2017). Even various small and medium-sized enterprises (SMEs) around the globe are in the process of initiating digitalization and automation to survive in the globalized environment (Cimini et al., 2020; Nwaiwu et al., 2020).

Industry 4.0 has also become a burgeoning topic for researchers. The academic contributions on the subject have dramatically increased in the last few years. The existing literature has interpreted Industry 4.0 from different perspectives such as *sectors* (Oesterreich and Teuteberg, 2016), *organizations* (Bokrantz et al., 2020; Li et al., 2019), and *areas* (Jerman et al., 2020; Liboni et al., 2019).

While the existing contributions offer valuable insights, they largely focus on studying the complex nature of Industry 4.0 and the technical and manufacturing aspects. The limited management literature reveals that organizations struggle and face diverse managerial challenges (Cimini

Corresponding author:
Anju Verma, Research Scholar, Indian Institute of Foreign Trade, Delhi Technological University, House No 48 Type II, Bawana Road, Delhi 110042, India.
Email: anjuvermasrcc@gmail.com

et al., 2020; Horvath and Szabo, 2019; Liboni et al., 2019; Schneider, 2018) in the transition process and sometimes even hesitate to implement Industry 4.0 (Horvath and Szabo, 2019). Some of the major challenges highlighted in the extant literature are lack of top management support (Sony and Naik, 2020), internal knowledge and skills gap (Cimini et al., 2020), availability of experts (Sony and Naik, 2020), organizational structure (Cimini et al., 2020; Li et al., 2020), leadership style (Schneider, 2018; Shamim et al., 2017), human resources (HRs) (Liboni et al., 2019; Schneider, 2018; Shaw and Varghese, 2018), data security (Kazancoglu and Ozkan-Ozen, 2018), and financial support (Horvath and Szabo, 2019).

One of the critical Industry 4.0 concerns is the effect that transformation will have on human resource management (HRM; Bajpai et al., 2019; BRICS, 2016). The advancement of digital technology and increase in automation will mean the replacement of traditionally manpower-operated processes by machines (Liboni et al., 2019). New skills and capabilities that will realistically be required in the future will need to be identified. The current employees will need to be upskilled and retrained. New talent will be required to be attracted, retained, and integrated into the business (Fareri et al., 2020).

Comprehensively, business leaders will need to focus on fronts such as effective HR management by developing new ways of working, building the next-gen workforce, fostering an organizational climate, a culture of lifelong learning (Mohelska and Sokolova, 2018), and shape innovative strategies for the future that can directly contribute to the business strategic objectives and successful implementation of Industry 4.0 (Lorenz et al., 2015; Verma et al., 2020). The findings on the numerous HR factors that potentially influence the implementation of Industry 4.0 exhibit the key role of HRM in Industry 4.0 implementation (Cimini et al., 2020; Fareri et al., 2020; Kazancoglu and Ozkan-Ozen, 2018; Llinas and Abad, 2019; Mohelska and Sokolova, 2018).

This study, therefore, sets out to present a comprehensive picture of HR success factors through a systematic review of the literature and develop an HR framework that could build a solid foundation for future research and provide orientation for practitioners and institutions to help get a better understanding of the impact of Industry 4.0 on HRM.

This article is structured as follows after the general introduction, and the second section consists of a literature review. The third section describes the methodological approach. The fourth section discusses the findings and analysis. The fifth section presents the discussion, implications, and conclusion and finally, the sixth section includes the limitations of the study.

Literature review

Stages of industrial revolution

The industrialization phenomenon has seen four great paradigm shifts since its iteration. These paradigm shifts are characterized and influenced by different technological advancements called the ''industrial revolutions'' (Miragliotta et al., 2018). *The first industrial revolution* in the late 18th century replaced animals with machines and introduced mechanical manufacturing processes (Habanik et al., 2019). Water and steam-powered machines were used for the mass production of goods (Thoben et al., 2017). In the late 19th century, came the *second industrial revolution* where electrical energy was used to run the machines. These machines were more efficient to operate, were less costly, and required less maintenance and effort, unlike the machines used in the first industrial revolution. Mass production of goods was streamlined using an assembly line (Thoben et al., 2017). The *third industrial revolution* was introduced in the early 1970s; the entire manufacturing industry was automated with the advances in electronics, computers, and IT (Kagermann et al., 2013; Thoben et al., 2017). The invention of electronic devices such as transistors and integrated circuits and software systems reduced the efforts and increased speed and accuracy. The 21st century led to the boom in the Internet and telecommunication technologies and brought major paradigm changes in the manufacturing processes. This revolution merged the boundaries of the physical and the virtual world and was called the *fourth industrial revolution or Industry 4.0*. The technological disruptions in this revolution allow man and machine to communicate in real time with no physical or geographical barriers (Jerman et al., 2020). It revolutionizes the management of logistics, production control, scheduling, quality control, and capacity utilization. The invention of smart technologies helps continuously monitor, detect, and predict faults and suggests preventive measures and remedial action.

Industry 4.0

The term ''Industry 4.0'' was initially coined in 2011 at the Hanover Fair. Later, in 2013, the German Government adopted it as a strategic initiative to transform and restructure the manufacturing industry (Nwaiwu et al., 2020). Lately, Industry 4.0 has gained increasing attention due to its diverse benefits to organizations. Several authors have given varied definitions of Industry 4.0 but most have defined it in a technocentric perspective, with technological advancements as the main trait. Kagermann et al. (2013) described Industry 4.0 as the integration of network resources, information, objects, and people in the realm of manufacturing. Oesterreich and Teuteberg (2016) defined Industry 4.0 or the fourth industrial revolution as increased automation, digitization, and interactions enabled by the creation of a digital value chain in a manufacturing environment. Szalavetz (2019) defined it as an umbrella term comprising a variety of digital enterprise technologies. According to Schwab (2016), the fourth industrial revolution (or Industry 4.0) is characterized by diverse technologies that blend the physical, digital, and biological spheres impacting all disciplines, industries, and economies. United Nations Industrial Development Organization (2017) has described Industry 4.0 by relating

technologies with human body parts. Automation and robotics, the key components, are considered the arms and legs of Industry 4.0; cameras and other sensors are the senses; data and connectivity as the nervous system; and AI is perceived as the brain of Industry 4.0.

Industry 4.0 has put the man, machine, processes, and infrastructure in a single loop leading to the highly efficient management of manufacturing processes (Miragliotta et al., 2018). It allows the organizations to be digitally interconnected with customers, suppliers, and the external environment and managed through SMART processes, network buildings, logistics, and products (Schneider, 2018). Though Industry 4.0 is still at its nascent stage and the industries are still in the transition phase of the adoption of the new technological systems, the future of Industry 4.0 seems flickering (Bag and Pretorius, 2020; Grant Thornton, 2017).

Industry 4.0 and HRM

To harness the transformative potential of Industry 4.0 business leaders across all nations and industries need to proactively plan and formulate comprehensive strategies to meet the challenges of the new era of accelerating change and innovation. CEOs are turning increasingly to the HRs professionals to drive technology absorption, foster innovation, and attract, retain, and develop the workforce of the future. By managing the workforce implications of Industry 4.0, HR leaders play a crucial role in ensuring that organizations can successfully adopt and deploy new and emerging technologies.

To facilitate Industry 4.0, the HR leaders need to become even more creative and strategic as more routine and administrative HR tasks will go away due to automation (Liboni et al., 2019). HR functions such as recruitment, training, and development will undergo massive transformations (Shaw and Varghese, 2018). According to the World Economic Forum (2018) report, human labor hours will decrease from 71% to 58% by 2025. More than 50% of all employees will require necessary reskilling by 2022. The challenges are significant. Therefore, HR professionals need to step up their game.

The existing literature has also recognized the role of HRM in Industry 4.0 implementation (Cimini et al., 2020; Fareri et al., 2020; Kiatsuranon and Suwunnamek, 2017; Li et al., 2020; Liboni et al., 2019; Mahdiraji et al., 2020). Authors suggest that HR strategies represent a key to take advantage of the opportunities of Industry 4.0 and help mitigate the attached challenges, make organizations future-ready, and adapt to the demands of the big revolution (Fareri et al., 2020; Nwaiwu et al., 2020; Park and Huh, 2018; Shaw and Varghese, 2018).

Taking into account the significance of HRM and the unprecedented potential of HRs, this study has tried to demystify and synthesize the HR approaches and success factors mentioned in the literature and develop an HR framework that can help businesses successfully navigate Industry 4.0.

Research methodology

Research approach

Literature shows that Industry 4.0 is a cross-cutting theme and several aspects of Industry 4.0 have already been studied. The technical aspects among all enjoy greater detail in the literature (Liao et al., 2017; Schneider, 2018; Shamim et al., 2017). A growing number of studies have started focusing on another important aspect, that is, implementation of Industry 4.0 (Calzavara et al., 2020; Kazancoglu and Ozkan-Ozen, 2018; Li et al., 2019). Still, much clarity is needed to build a solid bedrock for practitioners, scientists, and the Government to formulate effective strategies and policies for Industry 4.0 implementation. In light of this, a systematic literature review offers the necessary toolkit by providing a clear view of the overall problems and solutions identified in the literature. Although several studies have applied the systematic literature review methodology in the context of Industry 4.0, no study has explored the HRM success factors in a systematic manner that are important in the context of implementing it.

The methodology of systematic literature review helps practitioners to study the literature by summarizing large studies and explain the differences among them (Cook et al., 1997; Wolf et al., 2001). Over recent years, new systematic review methods have been developed to select, summarize, and report evidence (Denyer and Tranfield, 2009; Tranfield et al., 2003). It enables the analysis of the past studies and uncovers the body of literature that is relevant to a specific subject (Kitchenham, 2004). Systematic reviews require documentation of all steps of data extraction containing all general information, such as title, author, publication details, as well as the research-specific information, such as methods, gaps, and opportunities for future research (Tranfield et al., 2003). For these reasons, the same method was chosen as the research methodology for the present article. It maintains the same principles and applies the three main steps of planning, execution, and reporting suggested by Tranfield et al. (2003).

Data collection

The first phase of data collection involved identifying and extracting articles following a rigorous preestablished scheme from the scientific database Scopus-Elsevier, the database is selected because of its importance for the contributions of many international scientific publications that can be useful to perform the systematic literature review. The keywords used for article searching were *Industry 4.0 and Human Resource Management; Industry 4.0 and Smart Human Resource Management; and Industry 4.0 and HR 4.0.*

Data screening

The first search with keywords resulted in 349 publications. The inclusion and exclusion criteria listed below were then applied to screen publications, which resulted in 110 publications.

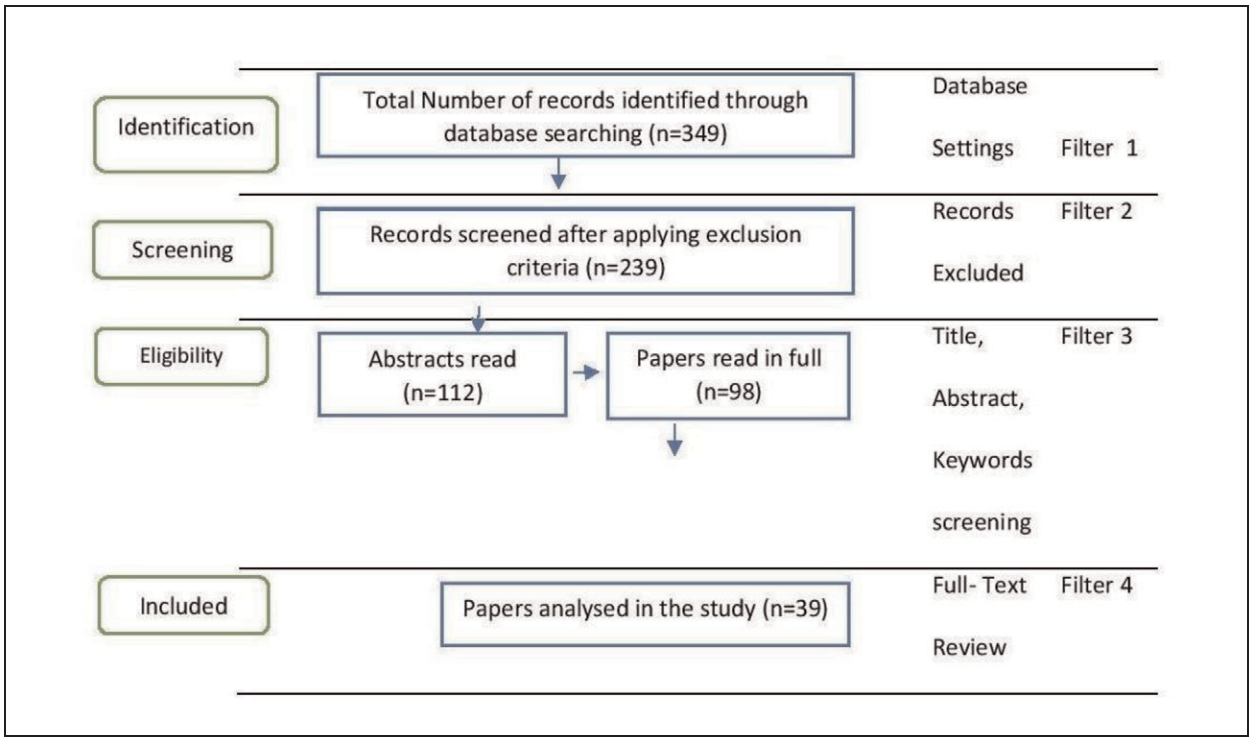


Figure 1. Systematic literature review procedure.

Inclusion criteria.

1. Studies published between 2010 and 2020.
2. Publication language is English.
3. Publication type is an article.
4. Publications in Journals only.
5. Studies that are directly related to Industry 4.0 and HRM.

Exclusion criteria.

1. Conference proceedings, monographs, book chapters, book reviews, theses, and project reports were excluded as refereed journal articles represent state-of-the-art research output with high impact (Chan and Ngai, 2011; Podsakoff et al., 2005).
2. Duplicate entries in the search output.
3. Publications where only abstract was available.
4. Articles not specifically related to HRM but Industry 4.0 in general.

Quality assessment of the reviewed literature is supremely important to a systematic literature review as the quality of conclusions entirely depends on the quality of selected literature (Yang, et al., 2021). Therefore, 110 full-text articles shortlisted from the initial search were thoroughly skimmed to evaluate the quality and eligibility of the articles. Only high-quality journal articles by reputable publishers with well-cited references were included in the review. The set was supplemented with additional five studies that were repeatedly found in the references list of selected articles. The final sample resulted in 39 articles. A memorandum with brief records of the aims and findings of each article was created for analysis in a Microsoft Excel spreadsheet. The study was completely funded by the Institution.

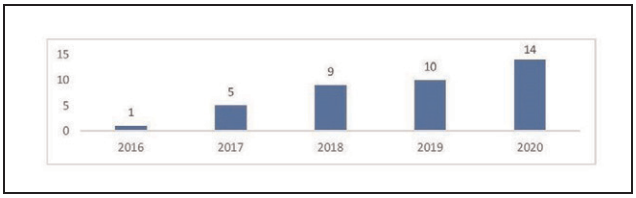


Figure 2. Number of publications per year.

The preferred reporting items for systematic review and meta-analysis protocols flowchart that reports the different phases of this systematic literature review is shown in Figure 1 (adopted from Moher et al. (2009)).

Descriptive analysis

Number of publications per year. This section shows the descriptive analysis of the articles. The timeline distribution of the articles analyzed is presented in Figure 2. Industry 4.0 and HRM perspective became a topic of interest only after 2017 as seen in the figure, with maximum studies done in the past 2 years. This depicts the accelerating interest in the topic among the researchers. Data of 2020 are still incomplete as this research was carried out in the middle of the year.

Descriptions of articles selected for the review are available in Appendix 1. It presents details about the authors and the title of the article.

Country-wise analysis of articles. The country-wise analysis of articles is conducted in Figure 3. It is depicted that most of

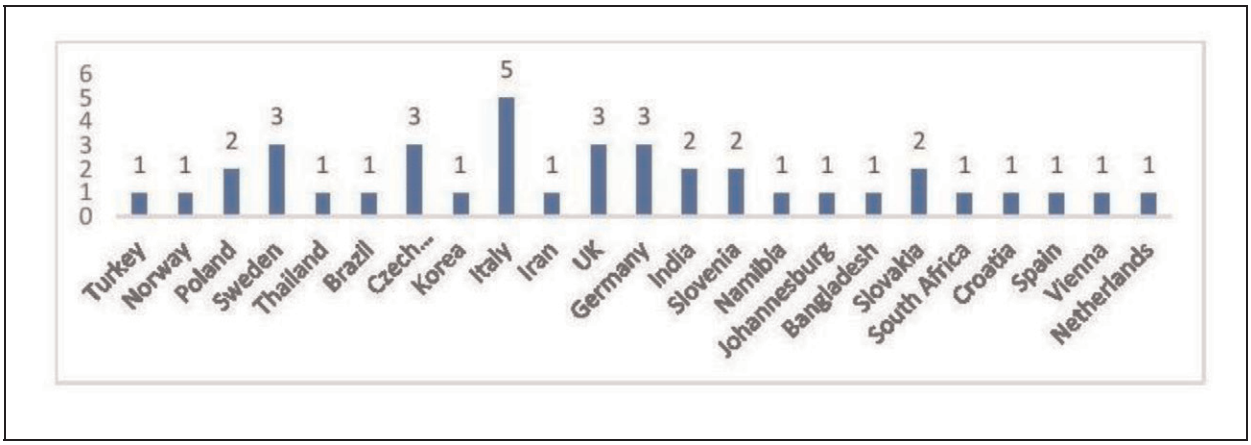


Figure 3. Country-wise analysis of articles.

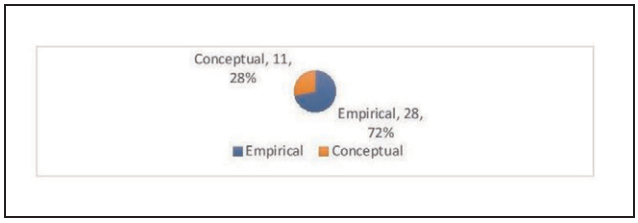


Figure 4. Approach of the study.

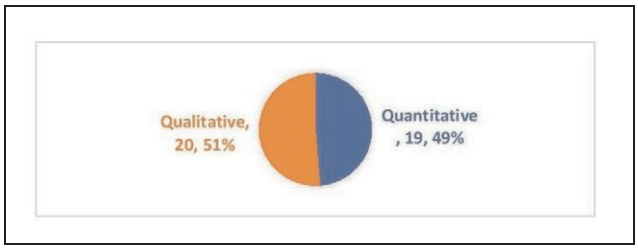


Figure 5. Method of research.

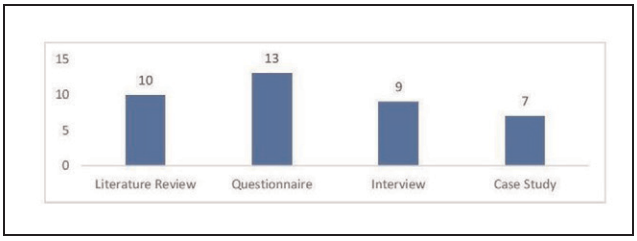


Figure 6. Techniques of research.

the research articles came from Italy, followed by Germany (where Industry 4.0 originated; Lasi et al., 2014), UK, Czech Republic, and Sweden. However, it should also be inferred that even small countries such as Namibia and Bangladesh are ushering in research on Industry 4.0 depicting the growing importance of this topic.

The approach of the study. The approach of the studies was analyzed based on conceptual and empirical research. Figure 4 shows that most studies were empirical (72%) rather than conceptual (28%). This describes a need for more conceptual research in this area.

Method of research. This category aims to investigate the methods used in the reviewed articles: qualitative or quantitative. Figure 5 depicts a somewhat equal proportion of methods used with 51% quantitative researches and 49% qualitative researches. This result may represent research opportunities for both methods in the future.

Techniques of research. This category aims to verify the research techniques used in the analyzed articles. Figure 6 shows that more than 33% of the articles used questionnaires as the main method followed by literature review (26%), interviews (23%), and case study (18%). It is relevant to highlight that there were no studies using focus groups, simulation, programming, or other different techniques.

Findings and analysis

Several studies highlight that in the Industry 4.0 manufacturing ecosystem, it becomes imperative to prepare the workforce and conduct strategic workforce planning aligned with organizational goals and policies (Liboni et al., 2019), build strategies for hiring skilled resources for new job roles, and upgrade existing roles to the new methodologies with a greater focus on compliance, regulatory, and security aspects. HR strategies formulation and implementation will have a significant impact on Industry 4.0 implementation. This section, therefore, presents a comprehensive review of the HR success factors and a framework that can help in the successful implementation of Industry 4.0.

HR success factors

These are the critical HR factors for success that have been identified from the literature.

Job design. Several authors in the literature highlighted that the digitalization and automation in the fourth industrial revolution will lead to new jobs carving and dejobbing in the organizations as robots and mechanical arms will shift

the low-skilled redundant jobs towards more high-skilled complex jobs (Benesova and Tupa, 2017; Fareri et al., 2020; Grencikova and Vojtovic, 2017; Kazancoglu and Ozkan-Ozen, 2018; Liboni et al., 2019; Pejic-Bach et al., 2020; Schneider, 2018; Shaw and Varghese, 2018; Sony and Naik, 2020). There is also a consensus that there will be a lot of dejobbing and many new professions will surface, which will require new professional skills and abilities in the employees (Fantini et al., 2020; Fareri et al., 2020; Kazancoglu and Ozkan-Ozen, 2018; Kulyk and Parmova, 2017; Whysall et al., 2019). Some of the authors also listed down new job roles and positions that will arise due to automation and digitalization (Benesova and Tupa, 2017; Jerman et al., 2020; Kulyk and Parmova, 2017).

Pointing towards the structure of the job, authors emphasized that jobs must be designed in such a manner that they include extensive job rotation, the delegation of responsibility and authority, and flexible assignments in different areas (Mahdiraji et al., 2020; Shamim et al., 2017). Further, they must also aid team-building and a collaborative work environment. Job duties and responsibilities, for instance, the number of different tasks to be performed, control over tasks and level of autonomy (Cimini et al., 2020; Wilkesmann and Wilkesmann, 2018), and specifications such as skills, knowledge, behavior, and attitude required for the job should be clearly outlined (Cimini et al., 2020).

Recruitment. Industry 4.0 will not just transform the businesses, it will also change the recruitment strategies to attract and hire people for the new jobs that will usher due to automation and digitalization. For the same reasons, many authors have suggested that the recruitment of new employees in the organizations should be strategically done according to the requirements of Industry 4.0 (Llinas and Abad, 2019; Sony and Naik, 2020). It is considered of utmost importance to develop a deep insight into the essential knowledge and skills for the rapidly changing era (Pejic-Bach et al., 2020). Literature indicates that big and imminent challenges in the human–machine interface will require highly qualified and smart employees to dominate sophisticated and complex technologies of Industry 4.0 (Bag and Pretorius, 2020; Grencikova and Vojtovic, 2017; Hasan et al., 2020; Johansson et al., 2020; Liboni et al., 2019; Nwaiwu et al., 2020).

Attributes such as technical skills, expertise (Sony and Naik, 2020), competencies (Cimini et al., 2020; Kazancoglu and Ozkan-Ozen, 2018; Kiatsuranon and Suwunnamek, 2017), innovativeness, learning, and knowledge management (Shamim et al., 2017) will be crucial for the success of Industry 4.0. It is also stressed that in the fourth industrial revolution, HR professionals will need to become more strategic and creative in recruiting employees for new jobs and positions (Shaw and Varghese, 2018). It is also worth noting that many findings illustrated the instrumental role of educational institutions in providing the professional skills needed to support the transformation and also proposing that organizations should work firmly and collaborate with Governments, educational institutions, and others in drafting 21st-century curriculum for students to better equip them to meet the skills requirements for future (Benesova and Tupa, 2017; Liboni et al., 2019; Shaw and Varghese, 2018).

Training. Literature reveals that to meet the epochal challenges of Industry 4.0, the most effective strategy could be the training of employees. With advanced technologies, businesses need to ensure that their employees have the required skills and competencies to stay abreast of the changing environment. Analysis of the literature also shows that how important and essential addressing potential skills gap, and retraining and upskilling of employees are for developing a successful Industry 4.0 strategy (Alba and Stucki, 2019; Benesova and Tupa, 2017; Nwaiwu et al., 2020; Shamim et al., 2017; Shaw and Varghese, 2018; Whysall et al., 2019). Additionally, imparting technical skills and developing soft skills such as learning, problem-solving skills, innovative capability, knowledge, social skills, attitude, flexibility, technology acceptance, and teamwork have been recognized as key areas of training to drive change and leverage on Industry 4.0 (Bokrantz et al., 2020; Fantini et al., 2020; Grencikova and Vojtovic, 2017; Hasan et al., 2020; Johansson et al., 2020; Liboni et al., 2019; Pejic-Bach et al., 2020; Schneider, 2018; Shamim et al., 2017; Stachova et al., 2019; Wilkesmann and Wilkesmann, 2018). It is no surprise that several authors shared this view.

Furthermore, studies by Bag and Pretorius (2020), Li et al. (2020), and Sony and Naik (2020) provided insights on the benefits that training programs can offer to organizations. They reported that harnessing skills and competencies can support service excellence and sustainable manufacturing, increases knowledge, and influences organizational setup. A list of skills and competencies has also been listed by various authors for the successful implementation of Industry 4.0 (Cimini et al., 2020; Jerman et al., 2020; Li et al., 2020).

Development. Apart from training, the most important success factor, a great many authors emphasized the pivotal role of developmental programs, appropriate management approaches, opportunities for personal development (Alba and Stucki, 2019; Kulyk and Parmova, 2017; Shamim et al., 2017), effective learning and innovation environment, and efficient communication (Shaw and Varghese, 2018) in Industry 4.0 paradigm. The authors stated that due to the shortage of skills in Industry 4.0, there is a greater need for diversification in talent management strategies, with more focus on internal development rather than external recruitment. Findings from Whysall et al. (2019) reveal that an inadequate supply and inability of organizations to realize the employees' expectations of accelerated career development can result in loss of talent. Therefore, it is extremely important to evaluate, build knowledge, skills, and abilities of employees, empower them with opportunities, and create career pathways for them to make them loyal and committed to the organization.

Compensation. Attracting qualified candidates may depend upon the company's ability to offer extremely competitive compensation. Though not discussed much in the literature, Shamim et al. (2017) believe that the compensation in Industry 4.0 must reflect the contribution of employees in the organization. The author also mentioned that performance-linked compensation and incentives, profit sharing, and rewards can facilitate creating a climate suitable for innovations, learning, and knowledge enhancement.

Performance appraisal. Analysis of the literature identified only one study by Shamim et al. (2017) that determined performance appraisal as a success factor in Industry 4.0. paradigm. The authors in the study stated that the performance appraisal system should focus on employee development. Continuous feedback should be provided to the employees. Also, management by objective is the most compatible performance appraisal method in Industry 4.0 to assess the performance of employees.

Safety and work conditions. New methods of production, faster decision-making, and self-organized multiagent work system in smart factories may significantly influence the health and safety of workers (Pejic-Bach et al., 2020). Therefore, adequate and right working conditions are essential to ensure workers' safety and efficiency. As highlighted by many authors safe, work conditions are crucial for volume production and avoiding potential dangers from fast-moving machines, vehicles, and equipment (Cocca et al., 2018; Tutak et al., 2020). Fantini et al. (2020) and Park and Huh (2018) mentioned in their studies that specific conditions of the workers may include the requirements for safety, because of age, permanent or temporary disabilities, stress, or fatigue. Bag and Pretorius (2020) suggested in their study that efforts should be made to maintain employee health and safety and work satisfaction for sustainable manufacturing in Industry 4.0.

Smart HRM practices. From recruitment to employee experience, from wellness to organization culture, from learning to an agile workforce, and from diversity to human capital, the role of the HR function is shifting from being a steward of employment to a steward of work. All the hardware, software, and data of HRM are expected to change in the age of Industry 4.0 (Strohmeier, 2018). HR leaders need to redefine their mandate and gear up for the new world of work which calls for smart HRM practices or HR 4.0 in organizations. Many recent studies also highlight that diverse technological advancement in Industry 4.0 will require an adapted and well-equipped strategic HRM, aligned with organizational goals and policies (Buchelt et al., 2020; Dhanpat et al., 2020; Liboni et al., 2019). Many authors believe that HR leaders, key technical personnel, and executives of the HR department must be updated, competent, trained, meticulous, skilled, and efficient to perform HR functions such as hiring, training, compensation, performance appraisal, safety, wellness, benefits, training of employees, motivation, retention, and change management while navigating Industry 4.0 (Buchelt et al., 2020; Kulyk and Parmova, 2017; Llinas and Abad, 2019; Shaw and Varghese, 2018; Sony and Naik, 2020; Strohmeier, 2018).

Studies by Buchelt et al. (2020) and Dhanpat et al. (2020) reveal that to mitigate challenges, HR professionals will need to assume the roles of strategic business leaders with competencies such as innovator, strategic partner, change agent, and people enablement. Cresnar and Nedelko (2020) mention that the personal values of HR professionals and employees will be significant in shaping the organizational environment for Industry 4.0.

Analysis of the literature also shows that innovative HR strategies should be formulated for the implementation of Industry 4.0. An environment of continuous innovation, learning, experimentation, risk-taking, and collaborations should be promoted (Schneider, 2018). Appropriate organizational culture should be developed (Li et al., 2019; Llinas and Abad, 2019; Mohelska and Sokolova, 2018). Aging workforce issues should be addressed by providing assistance devices, offering a better quality of life at work through ergonomic workstations and production processes, and giving value to their experience and maturity of older workers (Calzavara et al., 2020; Cocca et al., 2018). Efficient communication and employee development opportunities should be ensured, and employer brand should be developed and promoted to retain highly skilled workers and minimize attrition in this fast-changing age (Shaw and Varghese, 2018; Whysall et al., 2019).

Furthermore, few studies also assert that smart HRM powered by emerging technologies and new generation employees has the potential to transform all the HR processes (Bondarouk and Brewster, 2016; Sivathanu and Pillai, 2018). These technologies may be used extensively to perform various HR functions such as recruitment, selection, training, preparing job descriptions, and compensation management. Big data analytics can change the way data are collected, analyzed, and used for making the right decisions. AI can reduce the enormous amount of time and manual efforts that are spent on resume screening by automatizing the matching of a candidate's resume and job description. Training needs can be automatically tailored according to employee knowledge and skill gaps (Bondarouk and Brewster, 2016). Automated, online tests, and remote video-based interviews can predict better on-the-job performance in the future and reduce the overall recruitment cycle considerably. AI chatbots can be used to remove the interviewers' bias by real-time interpretation and validation of candidate responses. AR/VR can help in providing training to the employees anywhere at any time. It can also guide the onboarders with various office processes and layouts. Data-driven decisions like setting compensation, promotions, and performance appraisal can be facilitated through AI (Sivathanu and Pillai, 2018).

The review of the HR factors is summarized in Table 1.

Table 1. HR success factors.

HR factors	HR success factors	Authors (year)
Job design	Restructuring of jobs	Benesova and Tupa (2017); Grencikova and Vojtovic (2017); Kazancoglu and Ozkan-Ozen (2018); Pejic-Bach et al. (2020); Schneider (2018); Shaw and Varghese (2018); Sony and Naik (2020); Whysall et al. (2019)
	Job formalization (work processes standardization and regulation of workers' behavior)	Cimini et al. (2020); Fantini et al. (2020); Fareri et al. (2020); Wilkesmann and Wilkesmann (2018)
	Job specialization	Benesova and Tupa (2017); Cimini et al. (2020); Fantini et al. (2020); Fareri et al. (2020); Schneider (2018); Wilkesmann and Wilkesmann (2018)
	Horizontal (number of different tasks to be performed)	
	Vertical (control tasks and level of autonomy)	
	Extensive job rotation	Shamim et al. (2017)
	Flexible assignments	Mahdiraji et al. (2020); Shamim et al. (2017)
	Extensive delegation of responsibility and authority	Shamim et al. (2017)
	Team building	Shamim et al. (2017)
	Collaborative work environment	Shamim et al. (2017)
Recruitment	Qualification and skills	Bag and Pretorius (2020); Grencikova and Vojtovic (2017); Hasan et al. (2020); Johansson et al. (2020); Liboni et al. (2019); Nwaiwu et al. (2020); Pejic-Bach et al. (2020)
	Identification and redeployment of transferable skills	Shaw and Varghese (2018)
	Recruitment of permanent staff	Shaw and Varghese (2018)
	Recruitment of contractual staff	Shaw and Varghese (2018)
	Host/attend industry events/forums	Shaw and Varghese (2018)
	Outsourcing to special consultancies	Shaw and Varghese (2018)
	Special attention to innovativeness, learning, and knowledge while recruiting	Liboni et al. (2019); Shamim et al. (2017)
	Skill and task-based recruitment	Bag and Pretorius (2020); Benesova and Tupa (2017); Grencikova and Vojtovic (2017); Sony and Naik (2020)
	Strategic partnerships as well as collaborations with higher education institutions and Industry associations	Benesova and Tupa (2017); Liboni et al. (2019); Schneider (2018); Shaw and Varghese (2018); Wilkesmann and Wilkesmann (2018)
Training	*Technical skills*	Kiatsuranon and Suwunnamek (2017)
	State-of-the-art knowledge	Alba and Stucki (2019); Bag and Pretorius (2020); Benesova and Tupa (2017); Bokrantz et al. (2020); Cimini et al. (2020); Grencikova and Vojtovic (2017); Hasan et al. (2020); Jerman et al. (2020); Johansson et al. (2020); Kazancoglu and Ozkan-Ozen (2018); Kulyk and Parmova (2017); Liboni et al. (2019); Pejic-Bach et al. (2020); Schneider (2018); Stachova et al. (2019)
	Automation programming	Bag and Pretorius (2020); Bokrantz et al. (2020); Cimini et al. (2020); Liboni et al. (2019); Pejic-Bach et al. (2020)
	Data analysis	Bokrantz et al. (2020); Cimini et al. (2020); Kazancoglu and Ozkan-Ozen (2018)
	Awareness of IT security and data protection	Kazancoglu and Ozkan-Ozen (2018); Liboni et al. (2019)
	Soft skills	Bag and Pretorius (2020); Cimini et al. (2020)
	Conflict-solving	Cimini et al. (2020)
	Problem-solving	Benesova and Tupa (2017); Jerman et al. (2020); Kazancoglu and Ozkan-Ozen (2018); Schneider (2018); Shamim et al. (2017); Shaw and Varghese (2018)
	Creativity and innovation	Bag and Pretorius (2020); Cimini et al. (2020); Fantini et al. (2020); Grencikova and Vojtovic (2017); Hasan et al. (2020); Jerman et al. (2020); Kulyk and Parmova (2017)
	Decision-making	Bag and Pretorius (2020); Benesova and Tupa (2017); Cimini et al. (2020); Fantini et al. (2020); Hasan et al. (2020); Jerman et al. (2020)
	Organizational and processual understanding	Bokrantz et al. (2020); Kazancoglu and Ozkan-Ozen (2018)
	Sense-making	Kulyk and Parmova (2017); Shaw and Varghese (2018)
	New media literacy	Kulyk and Parmova (2017)
	Transdisciplinarity	Kulyk and Parmova (2017)
	Multitasking	Shamim et al. (2017)

(continued)

Table 1. (continued)

HR factors	HR success factors	Authors (year)
	Personal skills	Kiatsuranon and Suwunnamek (2017)
	Values	Cimini et al. (2020)
	Attitudes	Cimini et al. (2020)
	Flexibility	Benesova and Tupa (2017); Bokrantz et al. (2020); Cimini et al. (2020); Hasan et al. (2020); Jerman et al. (2020); Kazancoglu and Ozkan-Ozen (2018); Kulyk and Parmova (2017); Liboni et al. (2019); Shaw and Varghese (2018)
	Motivation to learn	Cimini et al. (2020); Johansson et al. (2020); Liboni et al. (2019)
	Ability to work under pressure	Cimini et al. (2020)
	Continuous learning and cooperation	Bag and Pretorius (2020); Bokrantz et al. (2020); Grencikova and Vojtovic (2017); Jerman et al. (2020); Kazancoglu and Ozkan-Ozen (2018)
	Trust in new technologies	Kazancoglu and Ozkan-Ozen (2018)
	Thinking in overlapping processes	Kazancoglu and Ozkan-Ozen (2018); Whysall et al. (2019)
	Mental cognition	Kulyk and Parmova (2017); Shaw and Varghese (2018)
	Attentional control	Shaw and Varghese (2018)
	Intuition	Shaw and Varghese (2018)
	Communication	Benesova and Tupa (2017); Grencikova and Vojtovic (2017); Li et al. (2020); Liboni et al. (2019); Shaw and Varghese (2018); Whysall et al. (2019)
	Reliability	Benesova and Tupa (2017)
	Sustainable mind-set	Liboni et al. (2019)
	Language skills	Benesova and Tupa (2017); Bokrantz et al. (2020); Liboni et al. (2019)
	Interpersonal skills	Kiatsuranon and Suwunnamek (2017)
	Social skills and abilities	Bokrantz et al. (2020); Cimini et al. (2020); Fantini et al. (2020); Johansson et al. (2020); Kulyk and Parmova (2017)
	Networking skills	Cimini et al. (2020); Li et al. (2020); Liboni et al. (2019)
	Leadership skills	Bokrantz et al. (2020); Cimini et al. (2020); Liboni et al. (2019)
	Ability to work in teams	Cimini et al. (2020); Fantini et al. (2020); Johansson et al. (2020); Liboni et al. (2019); Shamim et al. (2017)
	Cross-cultural competency	Kulyk and Parmova (2017); Liboni et al. (2019)
	Continuous mentoring	Shamim et al. (2017)
	Training to managers	Shamim et al. (2017)
	Retraining and upskilling of internal employees	Li et al. (2020); Nwaiwu et al. (2020); Shaw and Varghese (2018); Sony and Naik (2020); Whysall et al. (2019)
Development	Development programs	Alba and Stucki (2019); Shamim et al. (2017)
	Opportunities for personal development	Shamim et al. (2017); Whysall et al. (2019)
	Effective learning and innovation environment	Shamim et al. (2017); Shaw and Varghese (2018); Whysall et al. (2019)
Compensation	Rewards	Shamim et al. (2017)
	Profit-sharing	Shamim et al. (2017)
	Incentive-pay	Shamim et al. (2017)
Performance appraisal	Behavior-based appraisal	Shamim et al. (2017)
	Result-based appraisal	Shamim et al. (2017)
	Feedback on the employee performance	Shamim et al. (2017)
	MBO approach	Shamim et al. (2017)
Employee health and safety	Workers' safety and efficiency	Cocca et al. (2018); Park and Huh (2018)
	Work conditions	Bag and Pretorius (2020); Fantini et al. (2020); Tutak et al. (2020)
Smart HRM practices	HR professionals readiness and competencies	Buchelt et al. (2020); Cresnar and Nedelko (2020); Dhanpat et al. (2020); Kulyk and Parmova (2017); Liboni et al. (2019); Shaw and Varghese (2018); Sony and Naik (2020)
	Aligning HR practices and plans with Industry 4.0.	Buchelt et al. (2020); Dhanpat et al. (2020); Liboni et al. (2019); Llinas and Abad (2019); Sony and Naik (2020)
	Aging workforce management	Calzavara et al. (2020); Cocca et al. (2018); Schneider (2018)
	Organizational culture	Li et al. (2019); Llinas and Abad (2019); Mohelska and Sokolova (2018); Schneider (2018)
	Retention strategies	Llinas and Abad (2019); Shaw and Varghese (2018); Whysall et al. (2019)

Table 2. Industry 4.0 HR framework.

	Impact on HR practices	HR success factors	
Industry 4.0 technologies →	• Job design	• Restructuring of jobs	Successful Industry 4.0 implementation (organizational effectiveness) ←
	• Recruitment	• Attracting qualified and smart employees to dominate sophisticated and complex technologies of Industry 4.0	
	• Training	• Retraining and upskilling of employees	
	• Development	• Appropriate development programs	
	• Compensation	• Competitive compensation	
	• Performance appraisal	• Performance appraisal system focused on the employee development	
	• Safety and work conditions	• Adequate and right working conditions	
		• Smart HRM practices	

Industry 4.0 HR framework

In the fourth industrial revolution, businesses need to redefine work and the way they produce value. HR as a function also strives to redefine its mandate. The implications for HR are significant. New strategies for the management of HRs need to be developed. From the analysis of literature, the following Industry 4.0 HR framework has been developed for the successful implementation of Industry 4.0 (see Table 2).

Discussion implications and conclusion

The accelerating pace of Industry 4.0 creates an urgency to define how businesses must change and deliver within the digital transformation. Industry 4.0 also propagates significant management challenges. HR professionals will need to become smart to manage the era of Industry 4.0. They will need to continuously innovate, grow, identify challenges, barriers, and skill gaps, and revamp traditional HR methods to withstand major changes that will come their way in Industry 4.0 adoption.

The HRM approaches will also need to be changed with the change in the work environment and tasks in Industry 4.0. HR policies will need to be aligned with organizational goals and policies. Innovative HR strategies will be required for attracting, developing, evaluating, and retaining employees in the organization and enhancing their experience. Jobs will comprise less manual work and more control and supervision of the processes. Employees will no longer be associated with one particular type of job. They will need to share the space with intelligent robots. Information and data will be the key elements. Therefore, continuous learning, upskilling, and training would be necessary to develop the skills required for Industry 4.0.

The findings of the study strengthen the idea that HR will play a more strategic role in Industry 4.0 transformation. Effective HR strategies will prove to be a game-changer with the myriad of opportunities for companies. It will enable mass customization of products through agile and automated processes with a dynamic and predictive supply chain. The success metrics for businesses would be increased organizational effectiveness, profitability, and productivity, improved quality, and reduced labor costs, time, inventory and maintenance, and new value-generating services. Consequently, with the rapid obsolescence of technologies and offerings, companies will want to enjoy an early-mover advantage to Industry 4.0 to put themselves ahead of their competitors in global manufacturing. Moreover, with Governments in various countries investing resources into the Industry 4.0 mandate, the pace of transformation is increasing and future research is necessary to devise Industry 4.0 HR strategies.

Limitations

The limitations of the study include: first, the approach is explorative and requires quantitative validation. Second, the review is based on the articles extracted from a single academic source, that is, Scopus. Future research may include articles from other important databases as web of science, science direct, etc. Further, the factors unearthed may be discussed in more depth and statistically validated through questionnaires and interviews in different sectors and countries in the future.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs

Anju Verma https://orcid.org/0000-0002-2629-8007
Murugesan Venkatesan https://orcid.org/0000-0003-0507-6461

References

AIMA-KPMG (2018) *Industry 4.0—India Inc. Gearing up for Change*. India: AIMA-KPMG.

Alba JMD and Stucki V (2019) Enhancing youth employability: a public private development partnership model to acquire industrial skills. *International Journal of Economics and Business Research* 17(1): 1–17.

Bag S and Pretorius JHC (2020) Relationships between Industry 4.0, sustainable manufacturing and circular economy: proposal of a research framework. *International Journal of Organizational Analysis*. Epub ahead of print 16 July 2020. DOI: 10.1108/IJOA-04-2020-2120.

Bajpai H, Tandon A and Sinha A (2019) *The Future of Work in the Automotive Sector in India* (White Paper). India: The Centre for Internet and Society, India.

Benesova A and Tupa J (2017) Requirements for education and qualification of people in Industry 4.0. *Procedia Manufacturing* 11: 2195–2202.

Bokrantz J, Skoogh A, Berlin C, et al. (2020) Smart maintenance: an empirically grounded conceptualization. *International Journal of Production Economics* 223(107534): 1–17.

Bondarouk T and Brewster C (2016) Conceptualising the future of HRM and technology research. *The International Journal of Human Resource Management* 27: 2652–2671.

BRICS (2016) *Skill Development for Industry 4.0*. India: BRICS Business Council.

Buchelt B, Fraczkiewicz-Wronka A and Dobrowolska M (2020) The organizational aspect of human resource management as a determinant of the potential of polish hospitals to manage medical professionals in healthcare 4.0. *Sustainability* 12(5118): 1–15.

Calzavara M, Battini D, Bogataj D, et al. (2020) Ageing workforce management in manufacturing systems: state of the art and future research agenda. *International Journal of Production Research* 58(3): 729–747.

Chan YY and Ngai E (2011) Conceptualising electronic word of mouth activity an input-process-output perspective. *Marketing Intelligence & Planning* 29(5): 488–516.

Cimini C, Boffelli A, Lagorio A, et al. (2020) How do Industry 4.0 technologies influence organisational change? An empirical analysis of. *Journal of Manufacturing Technology Management* 32: 1–27.

Cocca P, Marciano F, Rossi D, et al. (2018) Business software offer for Industry 4.0: the SAP case. *IFAC PapersOnLine* 51(11): 1200–1205.

Cook DJ, Mulrow CD and Haynes B (1997) Systematic reviews: synthesis of best evidence for clinical decisions. *Academia and the Profession* 126(5): 376-380.

Cresnar R and Nedelko Z (2020) Understanding future leaders: how are personal values of generations Y and Z tailored to leadership in Industry 4.0? *Sustainability* 12(11): 4417.

Denyer D and Tranfield D (2009) *Producing a Systematic Review. The Sage Handbook of Organizational Research Methods*. Thousand Oaks, CA: Sage Publications Ltd, pp. 671–689.

Dhanpat N, Buthelezi ZP, Joe M, et al. (2020) Industry 4.0: the role of human resource professionals. *SA Journal of Human Resource Management* 18: 1–11.

Fantini P, Pinzone M and Taisch M (2020) Placing the operator at the centre of Industry 4.0 design: modelling and assessing human activities within cyber-physical systems. *Computers & Industrial Engineering* 139(105058): 1–11.

Fareri S, Fantoni G, Chiarello F, et al. (2020) Estimating Industry 4.0 impact on job profiles and skills using text mining. *Computers in Industry* 118(103222): 1–19.

Grant Thornton CII (2017) *India's Readiness for Industry 4.0—A Focus on Automotive Sector*. India: Grant Thornton.

Grencikova A and Vojtovic S (2017) Relationship of generations X, Y, Z with new communication technologies. *Problems and Perspectives in Management* 15(2–3): 557–563.

Habanik J, Grencikova A and Krajco K (2019) The impact of new technology on sustainable development. *Inzinerine Ekonomika—Engineering Economics* 30: 41–49.

Hasan MZ, Mallik A and Tsou J-C (2020) Learning method design for engineering students to be prepared for Industry 4.0: a Kaizen approach. *Higher Education, Skills and Work-Based Learning* 10: 2042–3896.

Horvath D and Szabo RZ (2019) Driving forces and barriers of Industry 4.0: do multinational and small and medium-sized companies have equal opportunities? *Technological Forecasting & Social Change* 146: 119–132.

Irudayaraj J (2017) *Industry 4.0—A Talent Perspective* (White paper). India: Rinalytics Advisors.

Jerman A, Bach MP and Aleksic A (2020) Transformation towards smart factory system: examining new job profiles and competencies. *Systems Research and Behavioral Science* 37: 388–402.

Johansson S, Kullstrom M and Bjork J (2020) Digital production innovation projects—the applicability of managerial controls under high levels of complexity and uncertainty. *Journal of Manufacturing Technology Management* 32(3): 772–794.

Kagermann H, Wahlster W and Helbig J (2013) Recommendations for implementing the strategic initiative INDUSTRIE 4.0: securing the future of German manufacturing industry. Final Report of the Industrie 4.0 Working Group. *Acatech—National Academy of Science and Engineering,* April 2013.

Kazancoglu Y and Ozkan-Ozen YD (2018) Analyzing workforce 4.0 in the fourth industrial revolution and proposing a road map from operations management perspective with fuzzy DEMATEL. *Journal of Enterprise Information Management* 31(6): 891–907.

Kiatsuranon K and Suwunnamek O (2017) Determinants of Thai information and communication technology organization performance: a structural equation model analysis. *Kasetsart Journal of Social Sciences* 40: 1–8.

Kitchenham B (2004) *Procedures for Performing Systematic Reviews*. s.l.: Keele: University Technical Report TR/SE-0401.

Kulyk V and Parmova DS (2017) E-business development: the comparative study of the Czech Republic and the Ukraine. *The Central European Journal of Regional Development and Tourism* 9(1): 80–110.

Lasi H, Fettke P, Feld T, et al. (2014) Industry 4.0. *Business & Information Systems Engineering* 6: 239–242.

Li AQ, Rich N, Found P, et al. (2020) Exploring product–service systems in the digital era: a socio-technical systems perspective. *The TQM Journal* 32(4): 897–913.

Li D, Berglund AF and Paulin D (2019) Current and future Industry 4.0 capabilities for information and knowledge sharing case of two Swedish SMEs. *The International Journal of Advanced Manufacturing Technology* 105: 3951–3963.

Liao Y, Deschamps F, Loures EdFR, et al. (2017) Past, present and future of Industry 4.0 a systematic literature review and research agenda proposal. *International Journal of Production Research* 55(12): 3609–3629.

Liboni LB, Cezarino LO, Jabbour CJC, et al. (2019) Smart industry and the pathways to HRM 4.0: implications for SCM. *Supply Chain Management: An International Journal* 24(1): 124–146.

Llinas D and Abad J (2019) The role of high-performance people management practices in Industry 4.0: the case of medium-sized Spanish firms. *Intangible Capital* 15(3): 190–207.

Lorenz M, Rüßmann M, Strack R, et al. (2015) *Man and Machine in Industry 4.0—How Will Technology Transform the Industrial Workforce Through 2025?* s.l. Boston, MA: The Boston Consulting Group.

Mahdiraji HA, Zavadskas EK, Skare M, et al. (2020) Evaluating strategies for implementing Industry 4.0: a hybrid expert oriented approach of BWM and interval valued intuitionistic fuzzy TODIM. *Economic Research—Ekonomska Istraživanja* 33(1): 1600–1620.

Miragliotta G, Sianesi A, Convertini E, et al. (2018) Data driven management in Industry 4.0: a method to measure data productivity. *IFAC Papers OnLine* 51(11): 19–24.

Mohelska H and Sokolova M (2018) Management approaches for Industry 4.0—the organizational culture perspective. *Technological and Economic Development of Economy* 24(6): 2225–2240.

Moher D, Liberati A, Tetzlaff J, et al. (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Medicine* 6(7): 1–6.

Nwaiwu F, Duduci M, Chromjakova F, et al. (2020) Industry 4.0 concepts within the Czech SME manufacturing sector: an empirical assessment of critical success factors. *Business: Theory and Practice* 21(1): 58–70.

Oesterreich TD and Teuteberg F (2016) Understanding the implications of digitisation and automation in the context of Industry 4.0: a triangulation approach and elements of a research agenda for the construction industry. *Computers in Industry* 83: 121–139.

Park S and Huh JH (2018) Effect of cooperation on manufacturing it project development and test bed for successful Industry 4.0 project: safety management for security. *Processes* 6(88): 1–31.

Pejic-Bach M, Bertoncel T, Mesko M, et al. (2020) Text mining of Industry 4.0 job advertisements. *International Journal of Information Management* 50: 416–431.

Podsakoff PM, MacKenzie SB, Bachrach DG, et al. (2005) The influence of management journals in the 1980s and 1990s. *Strategic Management* 26(5): 473–488.

Schneider P (2018) Managerial challenges of Industry 4.0: an empirically backed research agenda for a nascent field. *Review of Managerial Science* 12: 803–848.

Schwab K (2016) *The Fourth Industrial Revolution: What It Means, How to Respond*, s.l. Cologny: World Economic Forum.

Shamim S, Cang S, Yu H, et al. (2017) Examining the feasibilities of Industry 4.0 for the hospitality sector with the lens of management practice. *Energies* 10(499): 1–19.

Shaw P and Varghese RM (2018) Industry 4.0 and future of HR. *Journal of Management* 5(6): 96–103.

Sivathanu B and Pillai R (2018) Smart HR 4.0—how Industry 4.0 is disrupting HR. *Human Resource Management International Digest* 26: 7–11.

Sony M and Naik S (2020) Key ingredients for evaluating Industry 4.0 readiness for organizations a literature review. *Benchmarking: An International Journal* 22(7): 2213–2232.

Stachova K, Papula J, Stacho Z, et al. (2019) External partnerships in employee education and development as the key to facing Industry 4.0 challenges. *Sustainability* 11(345): 1–19.

Strohmeier S (2018) Smart HRM—a Delphi study on the application and consequences of the internet of things in human resource management. *International Journal of Human Resource Management* 31: 1–31.

Szalavetz A (2019) Industry 4.0 and capability development in manufacturing sub-sidiaries. *Technological Forecasting & Social Change* 145: 384–395.

Thoben K-D, Wiesner S and Wuest T (2017) Industrie 4.0 and smart manufacturing—a review of research issues and application examples. *International Journal of Automation Technology* 11(1): 1–13.

Tranfield D, Denyer D and Smart P (2003) Towards a methodology for developing evidence-informed management knowledge by means of systematic review. *British Journal of Management* 14: 207–222.

Tutak M, Brodny J and Dobrowolska M (2020) Assessment of work conditions in a production enterprise—a case study. *Sustainability* 12(5390): 1–32.

United Nations Industrial Development Organization (2017) *Industry 4.0: The Opportunities Behind the Challenge*. Vienna: United Nations Industrial Development Organization.

Verma A, Bansal M and Verma J (2020) Industry 4.0: reshaping the future of HR. *Strategic Direction* 36(5): 9–11.

Whysall Z, Owtram M and Brittain S (2019) The new talent management challenges of Industry 4.0. *Journal of Management Development* 38(2): 118–129.

Wilkesmann M and Wilkesmann U (2018) Industry 4.0—organizing routines or innovations? *VINE Journal of Information and Knowledge Management Systems* 48(2): 238–254.

Wolf FM, Shea JA and Albanese MA (2001) Toward setting a research agenda for systematic reviews of evidence of the effects of medical education. *Teaching and Learning in Medicine* 13(1): 53–60.

World Economic Forum (2016) *Mastering the Fourth Industrial Revolution*. Cologny: World Economic Forum.

World Economic Forum (2018) *The Future of Jobs Report*. Cologny: World Economic Forum.

Yang L, Zhang H, Shen H, et al. (2021) Quality assessment in systematic literature reviews: a software engineering perspective. *Information and Software Technology* 130(1–24): 93–112.

Appendix 1

Authors (year)	Title
Alba and Stucki (2019)	Enhancing youth employability: a public private development partnership model to acquire industrial skills
Bag and Pretorius (2020)	Relationships between Industry 4.0, sustainable manufacturing and circular economy: proposal of a research framework
Benesova and Tupa (2017)	Requirements for education and qualification of people in Industry 4.0
Bokrantz et al. (2020)	Smart Maintenance: an empirically grounded conceptualization
Bondarouk and Brewster (2016)	Conceptualising the future of HRM and technology research
Buchelt et al. (2020)	The organizational aspect of human resource management as a determinant of the potential of polish hospitals to manage medical professionals in healthcare 4.0
Calzavara et al. (2020)	Ageing workforce management in manufacturing systems: state of the art and future research agenda
Cimini et al. (2020)	How do Industry 4.0 technologies influence organisational change? An empirical analysis of
Cocca et al. (2018)	Business Software Offer for Industry 4.0: the SAP case
Cresnar and Nedelko (2020)	Understanding future leaders: How are personal values of generations Y and Z tailored to leadership in Industry 4.0?
Dhanpat et al. (2020)	Industry 4.0: The role of human resource professionals
Fantini et al. (2020)	Placing the operator at the centre of Industry 4.0 design: Modelling and assessing human activities within cyber-physical systems
Fareri et al. (2020)	Estimating Industry 4.0 impact on job profiles and skills using text mining
Grencikova and Vojtovic (2017)	Relationship of generations X, Y, Z with new communication technologies
Hasan et al. (2020)	Learning method design for engineering students to be prepared for Industry 4.0: a Kaizen approach
Jerman et al. (2020)	Transformation towards smart factory system: Examining new job profiles and competencies
Johansson et al. (2020)	Digital production innovation projects—The applicability of managerial controls under high levels of complexity and uncertainty
Kazancoglu and Ozkan-Ozen (2018)	Analyzing Workforce 4.0 in the fourth industrial revolution and proposing a road map from operations management perspective with fuzzy DEMATEL
Kiatsuranon and Suwunnamek (2017)	Determinants of Thai information and communication technology organization performance: A structural equation model analysis
Kulyk and Parmova (2017)	E-business development: The comparative study of the Czech Republic and the Ukraine
Li et al. (2019)	Current and future Industry 4.0 capabilities for information and knowledge sharing case of two Swedish SMEs
Li et al. (2020)	Exploring product–service systems in the digital era: a socio-technical systems perspective
Liboni et al. (2019)	Smart industry and the pathways to HRM 4.0: implications for SCM
Llinas and Abad (2019)	The role of high-performance people management practices in Industry 4.0: The case of medium-sized Spanish firms
Mahdiraji et al. (2020)	Evaluating strategies for implementing Industry 4.0: a hybrid expert oriented approach of BWM and interval valued intuitionistic fuzzy TODIM
Mohelska and Sokolova (2018)	Management approaches for Industry 4.0—the organizational culture perspective
Nwaiwu et al. (2020)	Industry 4.0 concepts within the Czech SME manufacturing sector: An empirical assessment of critical success factors
Park and Huh (2018)	Effect of cooperation on manufacturing IT project development and test bed for successful Industry 4.0 project: Safety management for security
Pejic-Bach et al. (2020)	Text mining of Industry 4.0 job advertisements
Schneider (2018)	Managerial challenges of Industry 4.0: an empirically backed research agenda for a nascent field
Shamim et al. (2017)	Examining the feasibilities of Industry 4.0 for the hospitality sector with the lens of management practice
Shaw and Varghese (2018)	Industry 4.0 and future of HR
Sivathanu and Pillai (2018)	Smart HR 4.0—how Industry 4.0 is disrupting HR
Sony and Naik (2020)	Key ingredients for evaluating Industry 4.0 readiness for organizations a literature review
Stachova et al. (2019)	External partnerships in employee education and development as the key to facing Industry 4.0 challenges
Strohmeier (2018)	Smart HRM—a Delphi study on the application and consequences of the internet of things in human resource management
Tutak et al. (2020)	Assessment of work conditions in a production enterprise—a case study
Whysall et al. (2019)	The new talent management challenges of Industry 4.0
Wilkesmann and Wilkesmann (2018)	Industry 4.0—organizing routines or innovations?

Note: HR: human resource; HRM: human resource management; SME: small and medium-sized enterprise; BMW: best–worst method; DEMATEL: decision-making trial and evaluation laboratory.

HRV and GSR as Viable Physiological Markers for Mental Health Recognition

Shikha
Department of Computer Science and Engineering
Delhi Technological University
New Delhi, India
shikha_2k19phdco502@dtu.ac.in

Aryan
Department of Electronics and Communication Engineering
Delhi Technological university
New Delhi, India
aryan_2k19ec032@dtu.ac.in

Lovish Arya
Department of Electronics and Communication Engineering
Delhi Technological University
New Delhi, India
lovisharya_2k19ec102@dtu.ac.in

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

2022 14th International Conference on COMmunication Systems & NETworkS (COMSNETS) | 978-1-6654-2104-1/22/$31.00 ©2022 IEEE | DOI: 10.1109/COMSNETS53615.2022.9668439

***Abstract*—Mental stress has become a standard part of day-to-day life. However, experiencing long-term and high-level stress affects the daily life and wellness of the person. Consequently, an individual's performance and management ability degrade significantly in critical situations. Electrocardiogram (ECG), Galvanic Skin Response (GSR), Electromyogram (EMG), Skin Temperature (ST), and Respiration are essential physiological biomarkers to quantify stress effectively. This paper aims to classify the stress level with improved performance based on GSR and ECG-derived Heart Rate Variability (HRV) features using machine and deep learning algorithms. It uses the Stress Recognition in Automobile Drivers (SRAD) dataset, which contains a collection of multiparameter recordings from 17 healthy participants who drive on a prescribed route under various pressure conditions. The work uses Pearson's Correlation, RFECV, and LightGBM feature selection methods with different classifiers to reduce redundancy between features and enhance performance. The accuracy and F1-score for stress level classifications are computed and compared using machine and deep learning algorithms. For binary classification (stress vs. non-stress), Random Forest achieves the best classification accuracy of 93.96% which is higher than previous works. It also provides an accuracy of 81.41% for three-class (baseline vs. medium stress vs. high stress) stress level classification.**

***Index Terms*—Physiological signals, Machine Learning, Deep Learning, Stress Detection, Heart Rate Variability (HRV), Galvanic Skin Response (GSR)**

I. Introduction

Stress indicates a feeling of strain and pressure caused by physical or psychological pressures at school, work, or personal life. Chronic stress results in many health problems that play a crucial role in the development of a weakened immune system [1], cancer [2], cardiovascular disease [3], depression [4], diabetes [5], or asthma [6]. Thus, stress causes severe damage to physiological health and psychological wellbeing. Stress avoidance is impossible, due to this reason, it is of paramount importance to identify a robust method to detect human stress. A questionnaire is a traditional approach to detecting stress, but the entire process depends on the individual's honesty. Automatic stress detection is challenging, but it could directly help people manage stress, and healthcare experts may provide more effective treatment for stress-related illnesses. Many researchers survey the physiological information of human beings can detect physical and mental stress [7]. The most common physiological biomarkers used to detect stress are as follows [7]:

- *Galvanic skin response (GSR):* GSR measures the continuous fluctuations in skin conductance. As the stress increases, sweat secretion increases in sweating glands, leading to decreased skin resistance and increased skin resistance.
- *Electrical activity of the heart (ECG):* Electrocardiogram (ECG) has been a reliable, accurate, and non-invasive method to derive the two commonly used stress indicating parameters, i.e., Heart Rate (HR) and Heart Rate Variability (HRV).
- *Electromyogram (EMG):* An electromyogram measures the electrical activity of the muscles. During stress, changes occur in the contraction of the muscles, which is the measure to identify stress.
- *Skin temperature (ST):* Stress level is associated with the change in temperature of the human body. Body temperature increases due to psychological stress, but some studies claim finger surface temperature decreases during the high-stress level.

These physiological biomarkers are observed as an accurate methodology for stress detection since these biosignals can not be masked or conditioned by voluntary human actions. However, this work majorly focuses on HRV and GSR signals.

The ECG-derived HRV is the variation in the interval between successive normal RR intervals of the QRS signal [8]. The time elapsed between two consecutive heartbeats is known as the RR interval. HRV is the widely used signal to investigate the changes in Autonomic Nervous Systems [9]. It

Authorized licensed use limited to: DELHI TECHNOLOGICAL UNIV. Downloaded on January 25,2022 at 08:34:00 UTC from IEEE Xplore. Restrictions apply.

indicates certain variations between consecutive heartbeats and represents the balance between sympathetic and parasympathetic nerves activities. When the person is under stress, HRV decreases but on the other hand, heart variability increases during relaxing and recovering activities.

Similar to HRV, GSR, also known as Electrodermal Activation (EDA), can predict the stress level by measuring the skin surface's electrical characteristics (conductance). This concept works on the measure of continuous variations in the electrical conductance of the skin surface caused by the sweating activity of the human body [10]. GSR works on Autonomic Nervous System (ANS), where if its sympathetic branch gets highly aroused, the sweat glands follow [10]. As a result, skin resistance decreases, and skin conductance increases. In this instance, the measured responses of the human Sympathetic Nervous System (SNS) affect human emotional and behavioral regulation. Thus GSR data can be used to evaluate the mental state of an individual. However, other physiological signals are required to discriminate between positive and negative high-arousal states as GSR alone cannot provide definitive information about mental stress.

In recent years, many researchers proposed various stress level detection approaches based on Physiological signals. Healey and Picard [11] introduced a Stress Recognition in Automobile Drivers (SRAD) dataset for stress detection of drivers using Physiological signals. The database consists of 17 individual volunteers driving on predefined routes. There are different levels of stress (Low, Medium, High) associated with different regions of driving experience, including rest, highway, and city streets. Each file consists of physiological signals of 17 participants such as Electromyography, Electrocardiogram, GSR, and Respiration.

Lopez-Martinez et al. [12] have employed SRAD Dataset for classifying two stress levels using Support Vector Machine (SVM) and Logistic Regression (LR) with high accuracy of 93% and 84%, respectively. They have shown that electrodermal activity (EDA) signals play a more significant role in detecting the driver's affective state than the heart rate (HR) signal. However, the authors focus exclusively on binary classification (Low-High) and ignore intermediate (Medium) stress evaluation.

Jiménez-Limas et al. [13] Propose an approach for binary stress classification using linear and non-linear features using SRAD Dataset. Out of six classifiers, Linear Regression provides the best results. It achieves the highest accuracy of 81.39% based on four biological signals: HRV, Respiration (RSP), F-GSR, and H-GSR. Thus, the driver's mental state is likely to be detected as soon as possible to avoid potential accidents. However, practically, long-term signals such as 5 minutes do not help see an early driver's mental state in actual situations.

Pankajavalli et al. [14] develop a unified machine learning framework for determining the binary stress level using SRAD Dataset. The framework tests six machine learning algorithm's performance with the complete and selected subset of features. The authors use two feature selection algorithms named R-FS and LASSO - FS and achieve the highest accuracy of 89% using the Support vector machine.

Elgendi et al. [15] investigate longitudinal and transitional to assess stress using supervised and unsupervised ML techniques. The results of supervised Machine Learning algorithms show ECG biosignal is the most informative biosignal for stress detection compared to the other tested signals such as EMG, GSR, HR, and RSP with an overall accuracy of 75.02%. They use the combination of unsupervised machine learning approaches, named K-means clustering, principal component analysis, and connectivity-based clustering.

Dalmeida et al. [16] study the role of HRV based features as physiological markers for stress detection. The authors develop a predictive model for binary stress classification using ECG-derived HRV features from the publicly available SRAD dataset [11]. They implement different machine learning algorithms such as Support Vector Machines (SVM), Random Forest (RF), K-Nearest Neighbor (KNN), Gradient Boosting (GB), and Multilayer Perceptron (MLP). The GB and MLP classification model provides the highest AUC score of 85% and 83% with F1 scores 79% and 77% using ECG-derived HRV features, respectively. Furthermore, the authors also validated the algorithms trained on the original dataset by extracting features from apple watch in real conditions.

Table I shows the comparison between related work. None of the previous work has done classification using both GSR And HRV, which seem to affect stress levels.

Contributions: This paper aims to analyze the physiological biosignals and enhance the performance using a reduced number of features through the machine and deep learning algorithm by classifying stress into two and three levels. We have experimented on the publicly available SRAD Database collected by Healey et al. [11]. This work considers two biosignals, i.e., HRV and GSR, and each biosignal is windowed of 30sec with 15sec overlapping. Compared to the earlier proposed results, the proposed model observes the highest classification accuracy score of 93.96% and with an F1 score of 87% using the Random Forest machine learning algorithm.

The rest of the paper comprises section 2, which presents the methodology describing the model for stress classification. Then, Section 3 discusses the results, and finally, Section 4 presents the conclusion.

II. Methodology

This section discusses the methodology for the model for stress classification. Fig. 1 illustrates the details for the various steps.

A. Description of the Database

This paper uses the SRAD Database for stress level detection collected by Healey and Picard [11]. The database is a publicly available dataset, consisting of multiparameter recordings in .dat and .hea format collected from 17 healthy automobile drivers in Boston, USA. The dataset comprises

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:34:00 UTC from IEEE Xplore. Restrictions apply.

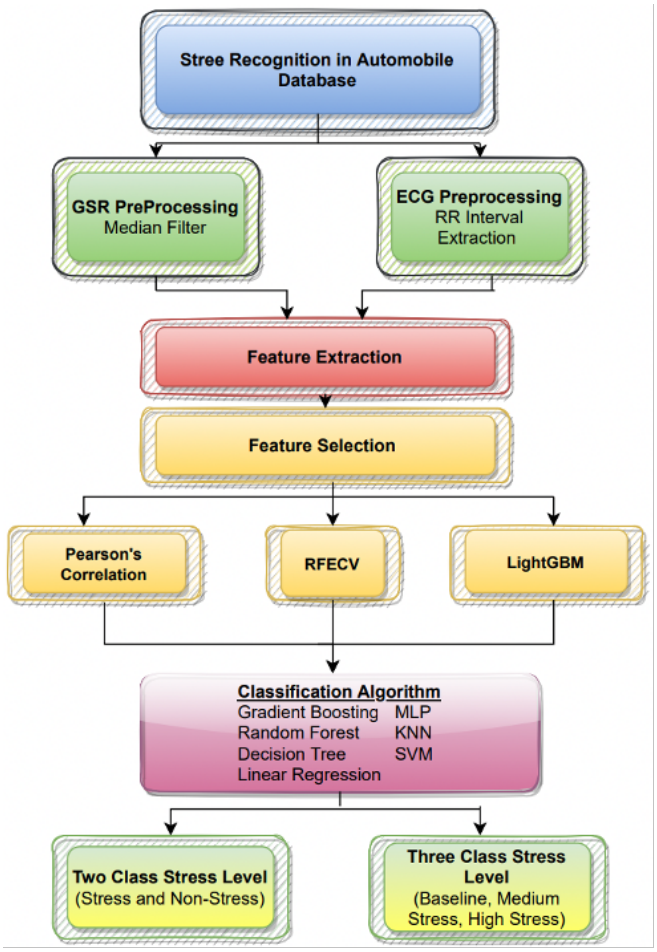


Fig. 1. Stress Classification Methodology

data collected after putting the drivers through different levels of stress. The protocol involves the period of rest with closed eyes in a relaxed position, highway driving, and city driving with traffic lights and unexpected risks generated by the cyclist, which beliefs to induce low, medium, and high stress. Each .dat file contains four physiological signals: EMG (Electromyography), ECG (Electrocardiogram), GSR, and Respiration. The sampling rate of hand GSR, foot GSR, and RESP signals is 31Hz, while ECG and EMG sampling rates are 496 and 15.5 Hz, respectively. The measured duration of the driver17 has two parts, 29 and 25 mins respectively, while other drivers have durations have 65-93 min. [11].

B. Pre-processing

The collected raw signals consist of various artifacts due to powerline interferences, movements, faulty connections. Therefore preprocessing is an important step. ECG-derived HRV features are extracted by using the correct RR intervals from the ECG signal. The median filter is applied to remove outliers from the raw GSR signal. Following is the detailed description of preprocessing pipeline.

1) Extracting RR Intervals from ECG data: An ECG waveform from the human heart consists of periodic patterns having points commonly labeled as P, Q, R, S, and T. The QRS waveform is an essential part of this signal which occurs during left and right ventricle depolarization.

A peak occurs in the middle of a QRS complex, which is generally called an R peak. Therefore, to calculate HRV, we need to find these R peaks. Preprocessing ECG signals aims to find a method to locate R peaks precisely, known as R peak detection. We can determine the RR interval from the difference between successive R peaks, the interbeat interval (IBI).

2) Peak detection using template matching: Peak detection is a two-stage process. In the first stage, we transform the signal so that the value of the output signal is around the area of our interest (QRS Complex) and is minimum at all the undesired areas. The next step is defining a threshold to filter the required components; hence a signal above the threshold is considered an R peak.

This work performs the signal transformation using a template matching algorithm as a stage, commonly used in pattern recognition, such as computer vision and particle analysis. This technique uses a filter (also known as kernel and template) with specific features. The basic algorithm works on the principle of sliding the filter simultaneously with the original signal and computing cross-correlation between the filter and signal. Signal with a similar waveform to the filter maximize the output, while a different waveform will minimize the result. The timestamp of the features that we are interested in is retrieved easily by tuning the threshold level of the signal.

Additionally, If the method uses the original signal to detect R peaks, it may miss the suppressed peaks. Therefore, the model robustly detected some suppressed R peaks similar to the given template using this technique. For selecting the template, it is better to use the template that robustly generalizes well to all the variants of signal that the model can encounter, having different sampling frequencies, different amplitudes. So sine waves are used as a template to detect QRS peaks.

3) RR interval Correction: The time elapsed between two consecutive heartbeats is known as the RR interval of the QRS signal on the ECG. By detecting RR peaks using the methods described above, some of the detected R peaks were very close, while others were far away. Hence some outliers (some very high and shallow values) were observed. To get accurate features, we must remove the outliers using a median filter. Zscore, the difference between a value and the mean of the distribution divided by standard deviation, is used as a selection criterion. The method replaces the values having a Z-score more than 2 by the median value and removes the outliers. This paper applies a median filter on GSR data as a part of preprocessing to remove the outliers; in the same way, it uses the filter on RR peaks data.

4) Stress Labeling And Windowing: The proposed model can detect three different drive phases (High, Middle, and Low Stress) using marker signals from the dataset. Based on the

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:34:00 UTC from IEEE Xplore. Restrictions apply.

dataset description, it considers 600 to 900 seconds as low-stress data, 1700 to 2000 seconds as medium-stress data, and 3100-3400 seconds as high-stress data. The model extracts balanced data from every subject with 5 minutes of signal for every stress label.

On top of that, the proposed model windows the labeled and preprocessed data. It splits every three hundred seconds of the signal for each stress level into 30-second windows with 15-second overlaps. As a result, it brings out 19 splits from every stress label of each subject. It extracts approximately 570 splits overall for two-level stress and 980 splits overall for three-level stress. Two-level stress indicates low and high-level classification, and Three-level stress indicates low, medium, and high-level classification.

C. Feature Extraction

The model extracts a total of 32 features include ECG-derived HRV features and GSR statistical features. The most relevant features are realized based on their importance and correlation and select a total of 5 features out of 32 features.

D. Feature Selection

Feature Selection is an essential step to determine a significant and relevant set of features. The feature selection advances employing three methods: Pearson's Correlation, Recursive Feature Elimination with Cross-Validation(RFECV), and LightGBM.

Pearson's correlation evaluates the correlation coefficient between objective class and each feature. The value ranges somewhere in the field of -1 and 1, where 1 indicates total positive correlation, 0 is no correlation, and -1 indicates the total negative correlation [17]. We pick the Pertinent elements in this work based on their profound correlation with the objective class. RFECV refers to the technique that uses recursive feature elimination with cross-validation to select the subset of best features for the supplied estimator. It determines and eliminates weak features using the fitted model's Coef_ and feature_importances_attribute.

After that, based on the LightGBM feature importance analysis, the model picks 5 out of 32 features. Selected features include Mean RR (ms), Mean HR (beats/min), RMSSD (ms), Mean GSR of the foot, Median GSR of the foot.

E. Parameter Tuning and Classification algorithms

This work applies Hyperparameter tuning on each algorithm to obtain the optimal parameters that would yield high performance, most efficient classification model. GridSearchCV, a scikit-learn library function, is used to achieve the best parameter combinations for high predictive performance. We apply K-Nearest Neighbor (KNN), Random Forest (RF), Multilayer Perceptron (MLP), Support Vector Machine (SVM),Gradient Boosting (GB), and Logistic Regression (LR) algorithms for different level stress classification.

1) K-Nearest Neighbour [KNN]: A supervised algorithm uses the nearest training point as neighboring points in a given region to perform the classification. We examine various neighboring point (k) values and power parameters(P) values to get the best KNN classification model. This work observes that the best parameter resulting from GridSearchCV are k = 25 and p=1.

2) Gradient Boosting [GB]: It is an ensemble-based model, which composes various decision trees. This model trains the trees to predict new data points where the individual tree dependent on each other. This model depends on the learning rate parameter that regulates the shrinkage of the contribution of each tree. This work observes the ideal value for the learning rate to be 0.75.

3) Multi-Layer Perceptron [MLP]: It is a layered feedforward artificial neural network developed to evade the constraints and weaknesses forced by the single-layer perceptron. MLP consists of a minimum of three successive layers with a specified weight. The information flows from the input layer to the output layer through one or more hidden layers. The optimal parameter values resulting from the grid search are μ = 0.1, h=0.001, and three hidden layers of 100, 10, and 400 nodes.

4) Logistic Regression [LR]: It is a statistics-based machine learning model that predicts the probability of a dependent variable whose nature is dichotomous. It signifies that the dependent variable is binary and reserves coded data as 1 (success) or 0 (failure).

5) Support Vector Machine [SVM]: It is a supervised learning algorithm that locates the best decision boundary known as a hyperplane. It segregates n-dimensional space into classes to put the new data points in the correct category of the classes. The SVM characterization model uses distinctive C parameters and functions like Linear, Polynomial, and Gaussian Radial Basis Function (RBF). RBF and polynomial function depend on the degree to find the hyperplane and γ, which defines the distance of a single point influence. Thus, the optimal parameters for the best stress classification model are the RBF function with γ = 'scale' and C = 1.

6) Random Forest [RF]: It is an ensemble of decision tree-based models comprising a mix of arbitrarily created decision tree classifiers. The forest in the model trains through bagging or bootstrap aggregating, further calculating results to acquire a more reliable training model. Thus, the number of trees, n = 100, is considered the optimal parameter for the best stress classification model.

III. Results and Discussions

This paper uses ECG-derived HRV and GSR biosignals for stress classification from the subjects who drove under various stress situations. This work aims to improve the performance of stress level classification by testing different machine and deep learning algorithms. First, we apply hyperparameter tuning to obtain optimal parameters for the different classifier models. Then to ensure the models are unbiased, the results are verified using K-fold Cross-Validation. Out of 32 features,

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:34:00 UTC from IEEE Xplore. Restrictions apply.

TABLE I
COMPARISON WITH SIMILAR WORKS

Research	Subjects	Physiological signals	Classifier	Window size	Accuracy	Classes (%)
Lopez-Martinez et al. [12]	17	EDA, HR	LR, SVM	30sec	93.00%	2 stress class (low stress-high stress)
Jiménez-Limas et al. [13]	11	FGSR, HR, RESP	LR, RF, SVM, KNN, NB	5 min	81.39%	2 stress class level (non stress-stress)
Pankajavalli et al. [14]	17	ECG, EMG, GSR	ANN, SVM, NB, DT, LR, KNN	30 sec	89.00%	2 stress class level (non stress-stress)
Elgendi et al. [15]	8	ECG, EMG, GSR, HR, RESP	KNN, LR, SVM, QDA, DT	15sec, 30sec	75.02%	2 stress class level (non stress-stress)
Dalmeida et al. [16]	17	ECG	GB, KNN, MLP, NB, RF, SVM	30 sec	$\approx$ 77%	2 stress class level (non stress-stress)
Proposed Model	**17**	**ECG and GSR**	**GB, MLP, KNN, SVM, RF, DT**	**30 sec**	**93.96%**	**2 stress class level (non stress-stress) and 3 stress class level (Baseline, medium, high)**

TABLE II
PREDICTIVE PERFORMANCE OF MACHINE LEARNING MODELS, CLASSIFYING STRESS ON 2 LEVEL AND 3-LEVEL

Classifiers	2 Level Stress		3 Level Stress	
	Accuracy (%)	F1 Score	Accuracy (%)	F1 score
MLP	80	0.85	54.4	0.53
KNN	88.69	0.78	75.02	0.78
SVM	84.82	0.81	65.01	0.67
Gradient Boosting	86	85	69.49	69
Random Forest	93.96	**0.87**	81.41	**0.84**
Decision Tree	92.57	0.84	78.53	0.83
Logistic Regression	75.22	0.69	53.45	0.49

a total of 5 features are selected based on their profound correlation. Studies have shown that changes in stress level impact HRV and GSR signals. RMSSD, Mean HR, and Mean RR as HRV metrics, and GSR mean and GSR median are the most reliable features for detecting stress levels. In light of that, they were considered the essential features for stress classification.

The proposed model focuses on accuracy and F1 score, which helps understand which algorithms would be best for experimenting on human subjects. Accuracy measures the total number of accurate predictions obtained by a model to the sum of predictions, despite incorrect and correct predictions. F1 score is tested throughout the experiment as it is one of the crucial metrics. Generally, the F1 score is the fusion of recall and precision, and an excellent F1 score intimates low false negatives and false positives. The perfect predictive model F1 score would be 1, whereas a total failure model F1 score would be 0. Equation (1) shows the formula for the F1 score.

$$F1score = 2 * \frac{precision * recall}{precision + recall} \tag{1}$$

Table I shows the comparison of the evaluated best results achieved in the recent work. Dalmeida et al. [16] label the stress level using GSR signal and uses ECG-derived signals for determining the stress states of the drivers. The authors have used the highest predictive power as a reference for developing an unsupervised stress application to classify stress using HRV measurements obtained from apple watch. However, the authors focus exclusively on binary classification (Low-High) and ignore intermediate (Medium) stress evaluation. In contrast, our work emphasizes binary and multi-level stress classification. For labeling, we consider the last five-minute segments of the rest period for low stress. The segments from highway driving and city driving periods are considered as the medium and high-stress levels.

Table II presents the obtained results from hyperparameter tuning. These results show that Random Forest outperforms other models with F1 scores of 87.88% and 84.88% for binary and three-level classification tasks imposed by this dataset. Thus, it implies that the model has at least 87.88% accuracy on classifying data instances. The Decision tree classifier also contributed comparable results to the Random Forest classifier.

TABLE III
COMPARISON OF PERFORMANCE OF MACHINE LEARNING MODELS

Classifier	**Dalmeidia et al. [16]** F1 Score	**Proposed Model** F1 score
Gradient Boosting	**0.79**	0.85
MLP	0.77	0.85
KNN	0.74	0.78
SVM	0.77	0.81
Random Forest	0.78	**0.87**
Decision Tree	—	0.84
Logistic Regression	—	0.69

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:34:00 UTC from IEEE Xplore. Restrictions apply.

Table III shows the comparison of our work with the best results achieved by Dalmeida et al. [16]. The authors report the maximum F1 score 79% based on 18 ECG-derived HRV features using Gradient Boosting for a two-level stress classification model. In contrast, the proposed model achieves the highest F1 score, 87%, and 84%, based on five HRV and GSR features for the two-level and three-level stress classification model. However, comparing the results obtained by Dalmeida et al. [16], this work obtains the highest accuracy of 93.96% and 92.57% for two-level stress classification using Random Forests and Decision tree machine learning algorithms, respectively.

IV. Conclusion

This paper uses HRV and GSR to classify the two and three-level classification, and results show that these two biosignals impact the changes in stress level. The aim is to improve the performance for two and three-class stress level classification by testing different machine, and deep learning algorithms with fewer features on the SRAD Database [11]. This work employs Pearson's Correlation, Re-cursive Feature Elimination with Cross-Validation (RFECV), and LightGBM feature selection methods to determine a set of HRV and GSR features based on their significance and correlation. The selected subset of features is input to different classifiers such as KNN, SVM, Random Forest, Decision Tree, Logistic Regression, Gradient Boosting, and MLP. In experimental results, we observe that ECG and GSR signals are effective biomarkers as stress indicators. In conclusion, Random Forest and Decision Tree outperform other models and provide the best results with accuracy of 93.9.6% and 81.41% for two-class and three-class stress level classification. Further, the F1 score of 87.88% and 84.87 for two and three-level stress classification implies that the models have the probability of more than 84% to classify the data accurately. In the future to validate the model performance on an extensive dataset, we plan to combine more driver stress dataset and create a generalized stress classification model.

References

[1] Segerstrom *et al.*, "Psychological stress and the human immune system: a meta-analytic study of 30 years of inquiry." *Psychological bulletin*, vol. 130, no. 4, p. 601, 2004.

[2] M. Moreno-Smith *et al.*, "Impact of stress on cancer metastasis," *Future oncology*, vol. 6, no. 12, pp. 1863–1881, 2010.

[3] A. Steptoe *et al.*, "Stress and cardiovascular disease: an update on current knowledge," *Annual review of public health*, vol. 34, pp. 337–354, 2013.

[4] C. Tsigos *et al.*, "Hypothalamic–pituitary–adrenal axis, neuroendocrine factors and stress," *Journal of psychosomatic research*, vol. 53, no. 4, pp. 865–871, 2002.

[5] A. M. Heraclides *et al.*, "Work stress, obesity and the risk of type 2 diabetes: Gender-specific bidirectional effect in the whitehall ii study," *Obesity*, vol. 20, no. 2, pp. 428–433, 2012.

[6] Y.-M. OH *et al.*, "Association between stress and asthma symptoms: A population-based study," *Respirology*, vol. 9, no. 3, pp. 363–368, 2004.

[7] G. Giannakakis *et al.*, "Review on psychological stress detection using biosignals," *IEEE Transactions on Affective Computing*, 2019.

[8] R. Lima *et al.*, "Heart rate variability and electrodermal activity in mental stress aloud: Predicting the outcome." in *Biosignals*, 2019, pp. 42–51.

[9] H.-G. Kim *et al.*, "Stress and heart rate variability: a meta-analysis and review of the literature," *Psychiatry investigation*, vol. 15, no. 3, p. 235, 2018.

[10] R. F. Navea *et al.*, "Stress detection using galvanic skin response: An android application," in *journal of physics: Conference series*, vol. 1372, no. 1. IOP Publishing, 2019, p. 012001.

[11] J. A. Healey *et al.*, "Detecting stress during real-world driving tasks using physiological sensors," *IEEE Transactions on intelligent transportation systems*, vol. 6, no. 2, pp. 156–166, 2005.

[12] D. Lopez-Martinez *et al.*, "Detection of real-world driving-induced affective state using physiological signals and multi-view multi-task machine learning," in *2019 8th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)*. IEEE, 2019, pp. 356–361.

[13] M. A. Jiménez-Limas *et al.*, "Feature selection for stress level classification into a physiologycal signals set," in *2018 15th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE)*. IEEE, 2018, pp. 1–5.

[14] P. Pankajavalli *et al.*, "A unified framework for stress forecasting using machine learning algorithms," *Advances in Computational Intelligence and Informatics. ICACII*, 2019.

[15] M. Elgendi *et al.*, "Machine learning ranks ecg as an optimal wearable biosignal for assessing driving stress," *IEEE Access*, vol. 8, pp. 34 362–34 374, 2020.

[16] K. M. Dalmeida *et al.*, "HRV features as viable physiological markers for stress detection using wearable devices," *Sensors*, vol. 21, no. 8, p. 2873, 2021.

[17] J. Brownlee, "A gentle introduction to the gradient boosting algorithm for machine learning. 2016."

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:34:00 UTC from IEEE Xplore. Restrictions apply.

Arabian Journal for Science and Engineering
https://doi.org/10.1007/s13369-021-06370-4

REVIEW ARTICLE–CIVIL ENGINEERING

Hydraulic Design and Analysis of Piano Key Weirs: A Review

Deepak Singh[1] · **Munendra Kumar**[1]

Received: 25 May 2021 / Accepted: 31 October 2021
© King Fahd University of Petroleum & Minerals 2021

Abstract
Dams are multipurpose structures that provide water for drinking purposes, hydroelectric power generation, navigation, industrial uses, and irrigation, in addition to flood control. The loss of storage due to siltation or submergence can be remedied by constructing an ungated spillway or weir over the dam structure. Many researchers have studied different shapes of the weir to reduce the submergence problem; some used the modified shape of the weir named Labyrinth weir in their studies, which gives the concept of a new shape of the nonlinear Piano Key Weir (PKWs). The PKW is a cost-effective solution for rehabilitation and new dam projects with a high level of constraints, for instance, limited space, small reservoir level, high specific flood discharge, etc. This paper reviews the previous studies on the historical development of the PKW, the geometry and structural integrity, the methodologies developed by researchers, and the stability of flow hydrodynamics at the PKW. The purpose of this study is to summarize the available information to date about PKW hydraulics and compare it to other alternative rectilinear or nonlinear weirs from the literature in terms of discharge capabilities, sediment transportation efficiency, aeration performance, and economic aspects. The comparison shows that the PKW is an excellent alternative for dam risk mitigation due to its high spillway capabilities, aeration performance, and economic point of view with other alternatives. Furthermore, in addition to field applicability, this study identifies future research possibilities that merit further investigation.

Keywords Spillway structures · Piano key weir and labyrinth weir · Enhancement of reservoir capacity · Aeration · Economics

1 Introduction

Dams are beneficial hydraulic structures in the development of civilization, and their safety measures are also equally essential that can be enhanced either by increasing the spillway capability or increasing the embankment crest (or increasing the gate height). Free flow spillways and weirs are the hydraulic structures used as discharge measuring devices on existing and newly constructed dam structures during flood release or open channel applications [1, 2]. The control of floods is generally assured, releasing the flood water without creating any hazardous effect on dam structures and protecting the towns and cities farther down the river. Therefore, the spillway must be appropriately designed and have structurally and hydraulically adequate capacity to release the surplus water [2, 3]. The spillway cost represents a significant part of the dam's total cost. Hence, the search for an efficient shape of free-flow spillways or weirs, which possess a high degree of performance at a low cost, is continuously going on. Many studies have been done to manage the tailwater or local submergence issue based on hydraulic characteristics, understanding of the weir, and enhancing a weir's efficiency at low head. Some researchers implemented a modified shape of a nonlinear weir named labyrinth weir.

PKW is an improved design of the traditional labyrinth concept, depicted for the first time in 2003 by Lempérière and Ouamane [4]. The first developers aimed to create a new type of labyrinth weir with a smaller footprint while maintaining a structurally economical and straightforward structure that could be easily constructed [4]. Thanks to the early fruitful collaboration between academia and industry, the concept developed fast and has been rapidly applied in the field to increase the discharge capacity at existing gravity dams in France [5] or as an impressive alternative to gated weirs for diversion structures in Vietnam [6].

✉ Munendra Kumar
mkumar66@gmail.com; munendrakumar@dtu.ac.in

Deepak Singh
kumaindeepaksingh@gmail.com;
deepak_phd2k18@dtu.ac.in

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

Published online: 16 January 2022

Since 2011, the regular organization of specific international workshops where knowledge is freely exchanged in an open and friendly environment facilitates the connection of all these actors, forming an international nonlinear weirs community, and resulted in the publication of three reference books summarizing the current state of the art [7–9]. Since the 2006 Goulours dam first Piano Key weirs commissioning in France, more than 35 Piano Key weirs have been built worldwide [10], and research continues throughout the globe. The first PKW was recently built at the Hazelmere Dam in South Africa to enhance reservoir storage and spillway discharge capacity. The installation of a PKW at the existing Tzaneen Dam is also due to start soon [11, 12]. Interestingly, in addition to the developments mentioned above focused on using PKW as a frontal weir for flood release, the concept has been recently shown interesting inside weirs applications [13].

Lempérière and Ouamane [4] were the first to show a systematic, conceptual behavior of a new shape of the nonrectilinear weirs. They named PKW tests to Types A and B. The Type A weir has a lesser efficiency than Type B [4, 14]. After that, many researchers classified the PKWs into four types (Types A, B, C, and D, as shown in Fig. 1) based on geometry. A Type A PKW has identical upstream and downstream overhangs or symmetrical about the transverse centerline axis. In Type B, the overhang portion is the upstream side only (outlet keys), whereas on the downstream side, it has vertical walls (inlet keys). Type C is the opposite of Type B, and Type D has sloping bottoms; there are no overhang portions [15]. Type A PKW has the most stable geometry; however, there is a tendency to enhance the upstream overhang length relative to the downstream one recommended by the literature to improve discharge capacity.

A combined effect of the PKW spillways with stepped chutes was presented by Bieri et al. [16], and they found that it leads to pronounced downstream energy dissipation. After installing PKW on various existing dams, Laugier et al. [17] and Vermeulen et al. [18] presented a detailed note on the lessons learned from PKWs' previous construction and design on the dam body. Also, they understood what steps they could take for further development or the creation of PKW spillways. A critical overview of the PKWs' systematic evolution from the labyrinth weirs was presented by Schleiss [19], Lempérière et al. [20], and Abhash and Pandey [21].

The published literature on debris blockage [23], scale effects [24, 25], and submergence provides guidance [26]. Noticeably, project engineers in Europe have added venting for the nappes downstream of PK weirs, despite experimental research providing no evidence of this need due to scale effects limitations. Recent EDF prototype measurements on the air demands of PK weirs [27] aid in developing appropriate design equations. All PKWs constructed to date are in concrete except one steel structure in terms of material [27]. There is potential space in this area for study and optimization.

The static hydraulic behavior of PKW is now well understood because several parametric studies have been undertaken during the past ten-to-fifteen years [28–30]. The hydrodynamics and transient flow behavior of PKWs are not so well understood and are the subject of several ongoing studies. As a result, additional research in this area is recommended.

This paper examines the use of PKW as an alternative solution for increasing spillway discharge capacity and mitigating dam risk by comparing its shape, geometry, structural integrity, discharge capacity. Moreover, other essential design aspects discussed include aeration, approach flow impact, sediment transport, scouring concept, energy dissipation, and the weir's economics.

2 Hydraulic Behavior of Piano Key Weir

PKW is a nonlinear overflow structure that induces complex and three-dimensional patterns of flow in its vicinity. It behaves similarly to a standard linear weir at the low head with a larger crest length. However, the additional velocity and longitudinal flow momentum at the inlet section at the slightly higher head cause flow over the lateral crest to deviate from this normal/average vector. Consequently, the flow efficiency over the side crest decreases marginally, and only the key to the inlet and outlet comes into play. Much literature is published to describe the hydraulic behavior; upstream structure design of the PKW. However, PKW's downstream structure design is prevented by a lack of systematic experimental studies and comprehensive data on flow conditions downstream of PKWs.

Depending on the upstream reservoir's characteristics, only a few studies have looked at the downstream components of PKWs so far. As a result, additional research in this area is recommended to supplement the preliminary work of [31–33]. Numerical models may be able to reproduce the complex flow conditions downstream of PK weirs with acceptable accuracy, allowing them to be used to validate downstream structure design or to investigate energy dissipation. Therefore, it is essential to do more research in energy dissipation and downstream scouring near the toe of PKW.

2.1 Hydraulic Analysis of PKW

A dimensional analysis between the discharge capacity and different geometrical parameters of PKWs was done by Noui and Ouamane [34]. It is defined that flow rate over the PKW has various geometrical parameters that directly or inversely affect the PKW's discharge capacity. The total discharge

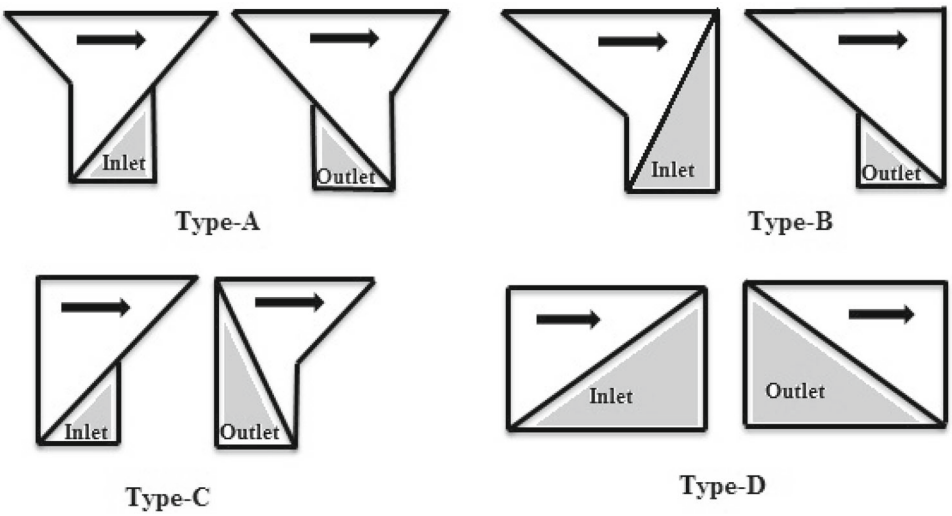


Fig. 1 PK weir types [Type A, Type B, Type C, and Type D.] (Adapted from Lempérière et al. [22])

across a PKW is a function of several parameters, as shown in Eq. (1):

$$Q = f(\rho, g, \mu, \sigma, t, H_t, L, P, P_b, W, W_i, W_o, B, S_i, S_o, \alpha, T_s, R, N) \quad (1)$$

where Q represents the total discharge flow across the PKW. ρ, μ, and σ represent the fluid properties (density, dynamic viscosity, and surface tension, respectively) of the flowing fluid. g is the acceleration of gravity; t is the temperature of flowing fluid. H_t represents the total upstream hydraulic head ($h_t + {v_t}^2/2$ g) over the PKWs, where h_t is the upstream water surface level above the weir crest and ${v_t}^2/2\ g$ average velocity head. The other geometrical parameters are related to the PKW, where L = total developed crest length; P = height of PKW, P_b = flow approach depth, W = total width of PKW; W_i and W_o = widths of inlet and outlet keys sections, respectively. $B = (B_i + B_b + B_o)$ length of side weir; where B_i and B_o represent the length of overhang portions on the inlet and outlet sides, respectively, B_b represents the base length. S_i and S_o represent the inlet and outlet key slope variations, respectively, α = angle between inlet/outlet key crest and sidewall weir of PKW, T_s = thickness of the PKW sidewall, R = height of parapet wall, and N represents the number of cycles.

Some researchers compute the discharge carrying capacity of the piano key weir under free-flow conditions based on the general discharge equation for rectangular sharp-crested weir as follows:

$$Q = \frac{2}{3} C_{dw} W \sqrt{2g} H^{3/2} \quad (2)$$

where Q represents theoretical discharge passing over the weir (m^3/s), C_{dw} is the discharge coefficient, the subscript w of the discharge coefficient refers to the discharge is being estimated uses the linear width of PKW, W is the width of the weir (m), g is the acceleration due to gravity (m/s^2), and H represents the total head over the PKW (m). [35, 36].

However, some researchers express the coefficient of discharge related to PKW developed length as:

$$Q = \frac{2}{3} C_{dL} L \sqrt{2g} H^{3/2} \quad (3)$$

where C_{dL} is the coefficient of discharge for PKW; the subscript L of the discharge coefficient refers to the discharge being estimated using the developed crest length (L) of PKW (L is in m) [37]. Concerning the hydraulics of PKWs, several design methodologies are available based on the prototype structure's model tests.

Kabiri-Samani and Javaheri [29] used Eq. (1) for the discharge calculation under the free-flow and submerged flow conditions of PKW Types A, B, and C, specific discharge ranges from 0.025 to 0.175 $m^3/s/m$. Furthermore, it suggested that Eq. (1) is only applicable to free overflow. Here, C_d is defined as a function of various relevant geometrical parameters:

$$C_d = \left[0.212 \left(\frac{H}{P}\right)^{-0.675} \times \left(\frac{L}{W}\right)^{0.377} \times \left(\frac{W_i}{W_o}\right)^{0.426} \times \left(\frac{B}{P}\right)^{0.306} \times e^{1.504 \frac{B_o}{B} + 0.093 \frac{B_i}{B}} \right] + 0.606 \quad (4)$$

where H represents the head over the PKW (m), H_d is the downstream head above the crest, P is the height of weir (m), L is the development crest length (m), W represents the total width of the weir, W_i and W_o represent the width (m) of inlet and outlet keys, respectively, B is the length of the PKW along with the flow (m), and B_i and B_o are the length (m) of overhang portion inlet and outlet sections, respectively. Then, Eq. (4) subjected the limitations as follows: $H > 30$ mm,

$0.1 \le H/P \le 0.6$, $2.5 \le L/W \le 7$, $1 \le B/P \le 2.5$, $0.33 \le W_i/W_o \le 1.22$, $0 \le B_i/B \le 0.26$, $0 \le B_o/B \le 0.26$ *and* $H_d/H \le 0.6$.

In order to calculate the discharge coefficient of PKW under submerged conditions as flows:

$$C_s = \left[1 - 0.858\left(\frac{H_d}{H}\right) + 2.628\left(\frac{H_d}{H}\right)^2 - 2.489\left(\frac{H_d}{H}\right)^3\right] \times \left(\frac{L}{W}\right)^{0.055} \tag{5}$$

Subjected to the limitations $2.5 \le L/W \le 6$, $1 \le B/P \le 2.5$, $0.33 \le W_i/W_o \le 1.22$, $0 \le Bi/B \le 0.26$, $0 \le B_o/B \le 0.26$ *and* $H_d/H > 0.6$.

Leite Ribeiro et al. [38] conducted a systematic experimental study over Type A PKW models and identified the primary and secondary parameters affecting the flow. They expressed the PKW head–discharge relation by expressing an increased discharge ratio (r) in terms of four correction factors: w, p, b, and a.

$$r = 1 + 0.24\left(\frac{(L - W)P_i}{WH}\right)^{0.9} wpba \tag{6}$$

where r represents the enhancement discharge ratio ($r = Q_{PKW}/Q_s$), Q_{PKW} represents the discharge over PKW in l/s, and Q_s is the linear sharp-crested weir discharge in l/s. The enhancement ratio r is primarily influenced by the following key parameters: overflow length L (m), weir width W (m), weir height P_i (m), and overflow depth H (m). Secondary parameters were also used to account for the contributions of key widths, overhang lengths, and parapet walls. Where $w = \left(\frac{W_i}{W_o}\right)^{0.05}$ represents the influence of the inlet and outlet key ratio, $p = \left(\frac{P_o}{P_i}\right)^{0.25}$ represents the difference in the height of the inlet and outlet keys, $b = \left(0.3 + \frac{B_i + B_o}{B}\right)^{-0.5}$ represents the influence of the overhang lengths and $a = 1 + \left(\frac{R_o}{P_o}\right)^2$ to specify the presence of any parapet walls.

Denys [39] discussed some essential standard design principles and flow-induced vibrations related to the piano key weir spillway, along with different methodologies developed by Lempérière [40]; Leite Ribeiro et al. [38] and Machiels [41]. Anderson [42] used discharge Eq. (3) for calculating the discharge coefficient over PKW and compared the results with the methodology developed by Lempérière [40]. Despite this, Anderson and Tullis [43] reported that Lempérière [40] equation is correct, but it only applies to the defined reference geometry by ICOLD [44], and not otherwise (see Denys [39] for details).

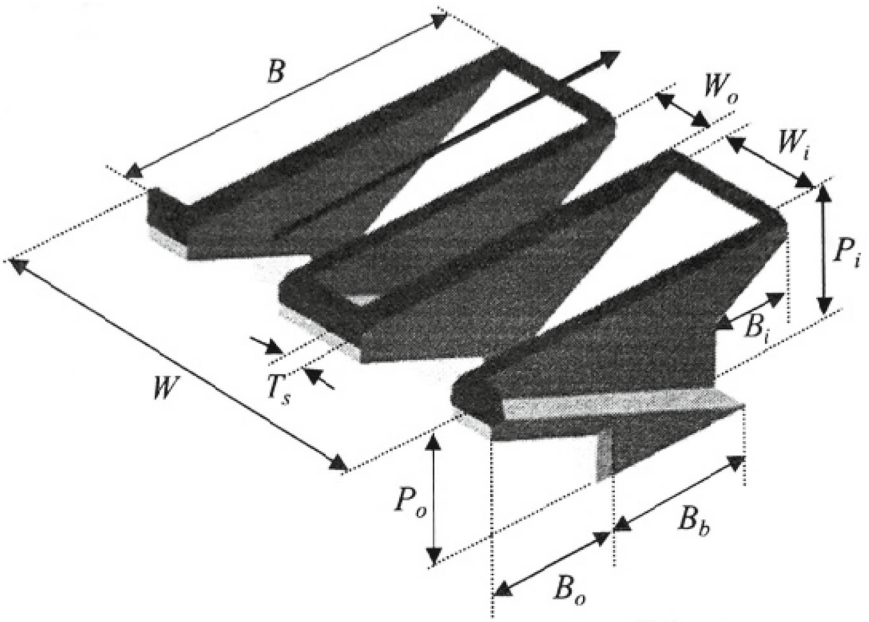
2.2 Geometry and Shape Optimization

PKW has a very complicated geometrical structure, and its efficient design depends on both aspects, geometrical and hydraulic elements. The main influencing geometrical parameters are the development crest length (L), the slope of the inlet–outlet keys (S_i and S_o), weir height (P), inlet and outlet key width (W_i and W_o), unit width (Wu), number of unit width, base length (B), lengths of inlet and outlet overhang portions (B_i and B_o), and wall thickness (T_s). The PKW unit is the fundamental element of the PKW. It represents the small-scale replica of a complete structure and is made of two sidewalls, an inlet key, and half of the outlet keys on both sides [45]. It can be written as the total width of the unit is $W_u = W_i + W_o + 2T_s$, where W_u is the width of one unit, W_i represents inlet key width, W_o represents outlet key width, and T represents wall thickness. Pralong et al. [45] described various geometrical parameters related to the PKWs (are shown in Fig. 2a, b), which are frequently used in the design analysis and are sufficient to provide comprehensive information about the PKWs. Some geometrical application limits which researchers have been used to date are presented in Table 1. There are some of the most discharge influencing parameters that are discussed in detail as below:

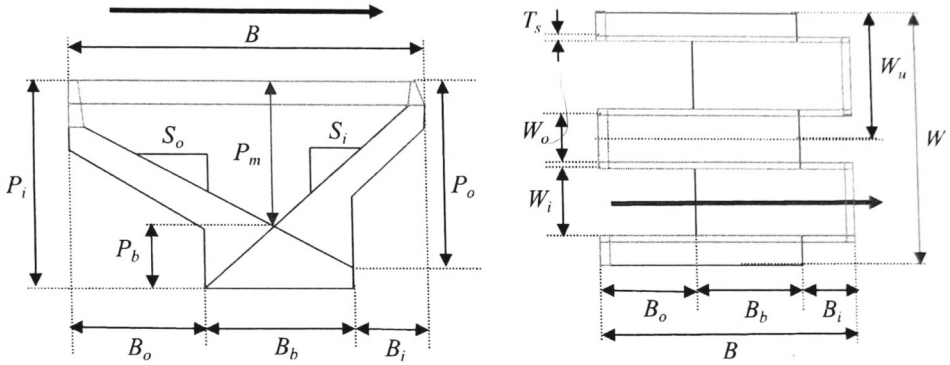
Height of weir (*P*)

The PKW's height is the most influencing parameter for the inlet and outlet key slopes, without increasing the overhang's potion length. Lempérière and Jun [51] stated that the minimum ratio of the floor slope for the inlet to outlet key should be 2:1. Further, Barcouda et al. [14] reported that with increasing the floor slope from 2:1 to 3:2, the discharge efficiency is increased by 20% for large H_t/P values, and it was found uniform for both Types A and B of PKW (where H_t head over the weir and P represents the weir height). Ouamane and Lempérière [3] tested three different PKW models with varying height P, and consequently, changes in floor slope (inlet–outlet key slope), with other parameters held constant. They found that an increasing the height by 25%, the efficiency increased by 6%. It is unclear whether the height P is more predominant for improving the efficiency or floor slope because two parameters are changing simultaneously. Lempérière [40] recommended that the inlet–outlet key slope of PKWs is 1.8:1 (horizontal: vertical). Increasing the weir height for the low head, 25%, and for the medium head, 5%, higher performance can be achieved [34].

The floor slope effect of the inlet–outlet key over the capacity of PKW was presented by Machiels et al. [35], and they suggested that as increasing the inlet key height, the evacuation capacity of the weir increases. However, Machiels et al. [52] indicate that a further rise in the slope for a slope more than 1.2 (*V: H*) increases the slope and does not significantly change discharge efficiency. Erpicum et al. [48] stated that the efficiency of PKW is influenced by various parameters such as inlet and outlet key widths and over-



(a) Fundamental parameters of a PK weir – 3D view (Adapted from Pralong *et al.* [85])



(b) Fundamental parameters of a PK weir – plan view (left) and cross-section (right) (Adapted from Pralong *et al.* [85])

Fig. 2 **a** Fundamental parameters of a PK weir—3D view (Adapted from Pralong et al. [46]), **b** fundamental parameters of a PK weir—plan view (left) and cross section (right) (Adapted from Pralong et al. [46])

hang length ratio (B_i/B_o) but highly influenced by the weir height as primary importance.

Most research on PKW height (or slope of inlet–outlet key bottoms) has used experimental and analytical methods. Only a few studies used numerical simulation. As a result, there is a need to develop numerical simulation or CFD analysis and validate the results with experimental data to understand the geometry of PKWs better.

Magnification ratio (*L/W*)

The next most significant parameter is the development crest length-to-width ratio or a magnificent ratio, which directly influenced PKW's discharge capacity. According to Machiels et al. [53], the development crest length-to-width ratio (L/W) = 5, height-to-unit width ratio (P/W_u) = 1.3, overhangs length ratio B_i/B_o = 3, and the ratio of key width (W_i/W_o) = 1.25 shows the highest discharge capacity, and this finding is steady with those of Lempérière et al. [20] and Leite Ribeiro et al. [54]. Lempérière and Jun [51], stated that the L/W ratio for effective PKW design must lie between 4 and 7. The value for $L/W = 6$ is recommended near to optimal by Barcouda et al. [14]. In order to large L/W ratio, more the PKW's efficiency [3]. Further, Lempérière [40] recommended that the L/W ratio equal to 5 was close to efficient. According to Hien et al. [55], for smaller H_t/P values, the L/W ratio equal to 7 is more effective, and for larger H_t/P values, the optimal value of the L/W ratio lies between 5 and 6. According to Laugier et al. [5], L/W ratio values of 6 or 7 will optimize discharge versus concrete quantity, thereby

Table 1 Different parametric limit of application in Design of PK weirs studied (adopted from Anderson [42], Pfister and Schleiss [47], Machiels et al. [35], Erpicum et al. [48])

Parameter → Reference ↓	L/W	H_t/P	Wi/Wo	B/P	Bi/B, Bo/B	P/Wu	Bi/P, Bo/P	S_i–S_b
Lemperiere and Ouamane [4]	6	*	1	*	*	*	*	0.75:1–1.5:1
Ouamane and Lemperiere [3]	4–8.5	*	0.67–1.49	*	*	*	*	*
Lemperiere [40]	5	*	1.25	*	*	*	*	1.8:1– 1.8:1
Anderson [42]	8	0.05–0.95	0.67–1.5	2.2–2.6	*	*	*	
Noui and Ouamane [34]	4–8	0.1–0.9	0.7–1.5	1.6–4.1	*	0.59–1.2	*	*
Kabiri and Javaheri [29]	2.5–7.0	0.1–0.6	0.33–1.22	1.0–2.5	0.0–0.26	*	*	*
Leite Ribeiro et al. [37]	3.0–7.0	0.1–2.8	0.50–2.00	1.5–4.6	0.07–1.2	0.29–0.65	0.25–4.00	*
Machiels [41]	4.2–5.0	0.1–5.0	0.50–2.00	1.0–6.0	*	0.33–2.00	0.0–2.67	*
Machiels et al. [35]	5	0.06–3.2	0.46–2.18	1–6	*	0.33–2.0	0–2.67	*
Mehboudi et al. [49]	2.8–6.5	0.07–1.11	1.33–4	1–7	0.1	0.128–0.178	0.3–2.33	*
Laiadi et al. [50]	6	0.09–0.52	1.2–1.5	1.99–2.7	0.12–.36	0.9–1.24	0.24–1.0	*

*Represents not reported

improving project economics. For maximum discharge efficiency, the development crest length of the PKW should be several times the weir width. Indeed, this ratio demonstrates how effectively a design maximizes the crest length L by maximizing the available width W. The flow control section must then be located along the crest to maximize the discharge on the developed crest length L. This means that the inlet should not constrain flow over the developed crest, and flow conditions in the outlet should not cause it to submerge. At this point, it is necessary to concentrate on the economic aspects of the PKW.

Inlet-to-outlet key width ratio (W_i/W_o)

Inlet-to-outlet key width ratio is another important parameter, which plays a significant role in influencing the PKW's discharge efficiency. The unit discharge approaching through the PKW crest directly determines the relative width of the inlet key. Most researchers presented that the optimal inlet to outlet ratio is 1.2 [3, 14, 51]. Further, Lempérière [40], Lempérière et al. [20], and Machiels et al. [53] claimed the optimal width ratio $(W_i/W_o) = 1.25$. Anderson [42] found that the maximum discharge efficiency is achieved when the width ratio (W_i/W_o) lies between 1.25 and 1.5. Leite Ribeiro et al. [54] and Machiels et al. [35] recommended that the inlet-to-outlet key width ratio is approximately 1.5.

All literature studies agreed that $W_i/W_o > 1.0$ has more significant discharge carrying efficiency than $W_i/W_o < 1.0$. With a relatively large value of the inlet key width (W_i), the flow has more space laterally, maintains the subcritical flow as possible, and reduces the losses [56]. On the other hand, the less value of the outlet key declined local submergence near the outlet section of the PKW and reduced the efficiency of the weir.

Overhangs portions (B_i, B_o)

Overhangs portions of the PKW are also equally important for developing and designing efficient PKW structures over the other main geometrical parameters. Ouamane and Lempérière [3] suggested that the overhang length ratio $B_i/B_o = 2$ is 12% less effective than Type B with $B_i/B_o = 0$ and 7% less effective than $B_i/B_o = 1$ of Type A PKWs. According to Cicero et al. [57], Type A has 15% more effective than Type C, but 5–15% less effective than Type B. Khassaf et al. [58] conducted laboratory experiments on the Type B model of PKWs to know about various geometrical parameters influencing weir's discharge capacity under free-flow conditions. They found that the Type B PKW has more efficient at low discharge. The larger overhangs on the upstream side increase the PKW efficiency because it increases the area of an inlet flow and reduces the energy losses [28]. PKW has an inclined apron with overhangs and, therefore, a smaller footprint than a labyrinth weir with the same crest length. Therefore, PKW is an efficient type of spillway from a hydraulic and economic standpoint PKW [15, 19, 59].

Parapet walls, crest shape, and nose

Parapet walls: The parapet wall is a short vertical wall mounted on the PKW's crest. It raises the height as well as the length of the weir's keys. The PKW design with parapet walls on the inlet keys can primarily improve the with relatively low heads. However, implementing PKW with parapet walls on the outlet keys has no effect on performance [60].

Crest shape: The crest shape has a significant role in the efficiency of any spillway, especially at low water levels [47]. Tests were carried out by Falvey [61] on labyrinth weir crest shape entirely applicable to PKWs. At low H_t/P values, their effect on a PKW is essential, but this

effect decreases substantially as $H_t/P > 0.3$ (where H_t is the total head over the weir and P is the height of the weir). Cicero and Delisle [62] studied the crest effect on PKW releasing capacity and noticed that a half-rounded or quarter-rounded crest on the lateral side crest had the most significant impact on a flat-topped crest.

Nose: The nose placement plays a vital role in guiding the flow to the inlet keys on either side of it, which is a valuable design feature. Their existence leads to smoother flow lines, decreased energy losses, and fewer vortex shedding, leading to improved PKW discharge efficiency [46].

3 Analytical Approach and Numerical Simulation

In the last decade, many experimental investigations on various aspects of the PKWs with rectangular plan forms have been conducted. Some researchers have carried studies based on analytical and numerical simulation and used the experimental results to develop and validate the analytical and numerical models. Machiels et al. [52] presented the effect on the releasing capacity of PKWs with the varying bottom slope of the inlet section. The experimental results were validated with an analytical approach, and the suggestion was made that the analytical method has an accuracy of 10% for the tested values. Further, Ali and Mansoor [63] compared the performance of PKW with the normal sharp-crested weir using experimental data and validating by regression analysis. They found that the PKWs perform doubly than the normal sharp-crested weir at the same head over the weir crest.

Moreover, Erpicum et al. [64] presented a 1-D numerical simulation model for the preliminary design of PKW based on the conservation of mass and momentum equations. Many other researchers have used the CFD technique to simulate the flow over PKWs reported by Oertel and Tullis [65]; Mario Oertel [30]; Cicero et al. [57]; Crookston et al., [66]; Li et al. [67]; and Al-Baghdadi [68]. Denys and Basson [69] used numerical and physical modeling to investigate unique transient hydrodynamics in the inlet and outlet keys. They collected the physical model data and compared or verified it with numerical simulation to see the pressure fluctuating upstream and downstream side of the wall.

Most studies suggested that all mathematical methods than empirical relations could more effectively simulate the PKW discharge. However, it is fair to say that the mathematical approaches for the design of PKW have not yet emerged as a reliable means. Soft computing is a collection of techniques, including artificial neural networks (ANNs), genetic programming, gene expression programming, support vector machine, group method of data handling, and adaptive neuro-fuzzy inference system (ANFIS).

Olyaie et al. [70] proposed several AI-DDM techniques (ANN, Gene Expression Programming, Extreme Learning Machine, and Least Square Support Vector Machine) to estimate and compare the discharge coefficient of the PKW. Furthermore, Kashkaki et al. [71] applied the Artificial Neural Networks (ANN) technique to their experimental data that were performed on three different circular PKW models with various angles of 45°, 60°, and 90°. They reported that the PKW spillway with a 90-degree angle has the highest efficiency, and the ANN technique has a good agreement between predicted and measured values. In order to simulate the flow passage over PKW, Zounemat-Kermani and Mahdavi-Meymand [72] used embedded artificial intelligence data-driven models (ANFIS and MLPNN) with multiple meta-heuristic algorithms (GA, PSO, FA, and MFO). Moreover, it was concluded that the hydraulic characteristics of the PKW could be simulated more accurately using analytical or mathematical approaches than empirical relations.

Generally, it is safe to assume that the mathematical model for the design of PKW has not yet emerged as a reliable means. The experts recognize the need for the code to be confirmed and accepted. By comparing the findings to experimental analyses, which are reliably the final proof, errors in mathematical methods can be minimized.

4 Efficiency of Piano Key Weir Over Its Alternatives

PKWs are the best alternative over the other weir in terms of efficiency and easy installation. Its reduced structural footprint on the top of newly constructed or existing dams, making it a cost-effective solution. It is the replacement of the linear ungated type of free-flow structures by increasing the unit discharge at the same head and same channel or spillway width. PKW is more suitable for new dams or hydraulic structures as a free flow spillway and convenient for spillway rehabilitation on the existing concrete and gravity dams already equipped with a free flow spill [4, 17, 28].

Most of the PKWs built to increase flood discharge capacity at existing dams were designed for upstream heads less than 2 m, partly due to rehabilitation project constraints to maintain existing reservoir and maximum pool levels at these dams. Furthermore, the H/P ratios for PKWs (average value of 0.46) are lower than those for labyrinth weirs, owing to the desire to maximize the discharge coefficient as well as structural stability (because the depth of PKWs is generally much less than labyrinths, thus having less overturning resistance). For new PKW projects, larger upstream heads with higher H/P ratios have been considered.

Figure 3 shows the comparison curve presented by Lempérière and Ouamane [13], among the Type A (with 4 m

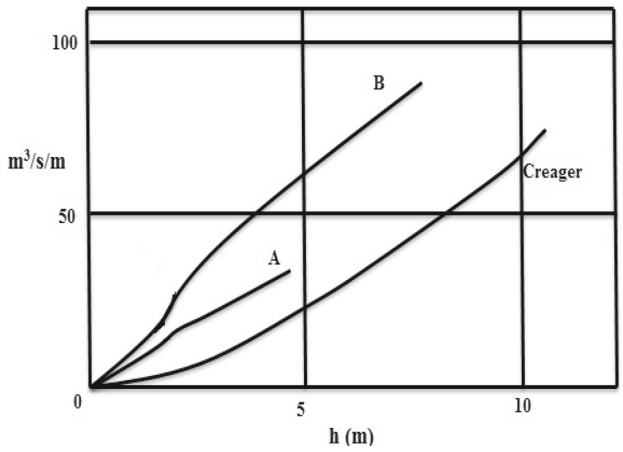


Fig. 3 Hydraulic data and saving for Type A and Type B solution of PK weir (Adapted from Lempérière and Ouamane [37])

wall height), Type B (for 8 m wall height) PKW models with Creager weir in respect of hydraulic performance and cost-saving along with the increase in live storage. The quantity of reinforcement concrete increases per meter length of the spillway, and saving is reduced significantly for reducing the nappe depth. When compared in the literature for PKW, Creager and traditional labyrinth weirs for use in existing and new structures with or without gates and fuse plugs.

Figure 4 depicts the comparative rating curves presented by Pfister and Schleiss [47] for a symmetrical Type A (prototype) PKW by reviewing various methodologies developed by the researchers [29, 41, 54]. The comparison shows that the models produce similar, but not identical, results. The differences in the results are thought to be due to differences in the crest shape profiles used in each of the methods. The methodology developed by Machiels [41] is the most conservative of the three in that it produces the lowest discharge for a given head or produces a larger head for a given discharge.

Moreover, the rating curve of a linear sharp-crested weir with the same width and height is expressed, computing in respect of Vischer and Hager [73]. It was concluded that the three studies predict comparable PKW discharge for a safe head but not equivalent effects. The second approach [54] typically predicts the highest discharge efficiency, while the analytical equation in the third approach predicts lower values [41]. The PKW's crest shape is partly the result of this disparity. Leite Ribeiro et al. [54] and Kabiri-Samani and Javaheri [29] used a half-rounded crest and a sharp-crested PKW, respectively, whereas Machiels et al. [60] used a flat-topped weir in their studies. In terms of discharge capacity, the flat-topped weir crest still falls short of the half-rounded crest and sharp-crested weir.

Figure 5 represents the comparison between discharge coefficients of various weir types by Blancher et al. [74]. The results show that for total upstream head-on weirs ranging from 0.5 to 4 m, PKW discharge performance is about 15 to 20% more efficient than an equivalent rectangular-shaped labyrinth one. These are typical discharge coefficients for a variety of overflow structures, as assessed by numerical modeling. They revealed that nonlinear weirs appear to be more efficient than linear weirs. However, this is only true for low upstream heads. The relative benefit decreases asymptotically as the heads increase over the weir in proportion to the weir's height. Under certain conditions, the labyrinth weir can achieve higher discharge efficiency than PKW at a lower cost [39]. The efficiency of the PKW to the labyrinth weir appears to be derived primarily from the spillway's sidewall. The PKW outlet part, which is more sensitive to downstream submergence, narrows the gap slightly. In contrast with the labyrinth spillway, the depth of approach significantly influences the PKW spillway. Nonetheless, the PKW is more effective.

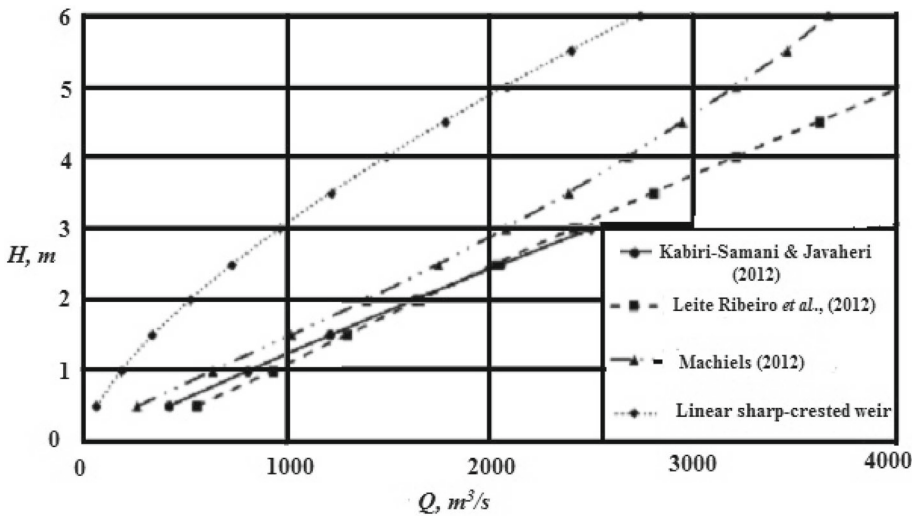


Fig. 4 Comparison of rating curves for PK weir, linear sharp-crested weir (Adapted from Pfister and Schleiss [59])

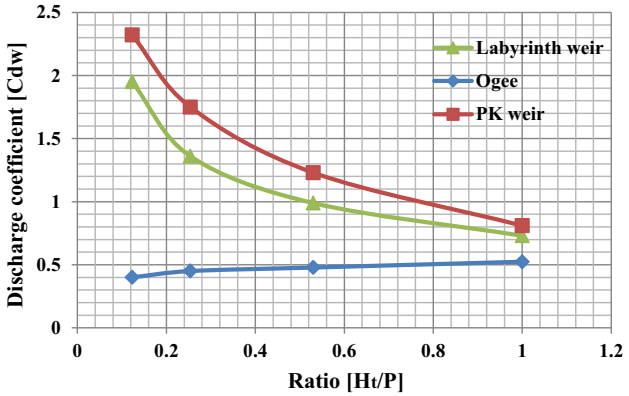


Fig. 5 Total discharge coefficients for various weir types (Adapted from Blancher et al. [74])

Figure 6 depicts a comparison of the results obtained by Erpicum et al. [75], for various types of weirs studied. They noticed that PKW (any type) has a much greater discharge per unit width than an ogee crested weir. However, to higher heads, the increment of the PKW efficiency is more diminutive than a linear weir. Because of the reduction in development length due to outlet submergence, the PKW was designed with a maximum H_t/P ratio of less than one [56]. According to hydraulic tests conducted in various countries, the PKW's discharge carrying capacity at the lower head is several times higher than the linear weir (see Fig. 6). It is primarily due to the enhancing of the developed crest. It also produces about 10% more significant discharges than similar-shaped labyrinth weirs. This is mainly due to the use of upstream overhangs.

Anderson and Tullis [66] estimate that a PKW is 10% more efficient than a traditional labyrinth weir with the same crest print for a head H_t equal to its height P_i. The coefficient of discharge of a PKW is greater than that of a comparable rectangular or trapezoidal labyrinth weir [28, 74], and the efficiency benefit can exceed 20%. According to Pfister and Schleiss [76], some previous approaches focused on physical tests performed in channels and did not consider various approach flow conditions present in the open reservoir. Moreover, none of the equations allow for economic optimization of the weir dimensions, so further mathematical, numerical, or physical modeling studies are recommended.

5 Some Important Design Aspects

The PKW is a complex structure that involves a large number of geometrical and hydraulic parameters. Several additional design aspects or factors that could change the PKW's efficiency should be considered during the design process. Several of these are explained below.

5.1 Aeration Performance of PKW

The concept of aeration was first introduced by Gameson [77]. Spillways aeration [78], side-channel elevated pool aeration [79], labyrinth weir, and PKW [4] have all been built since then. Most of the PKW's physical models have shown that the flow over the weir is naturally aerated. It is only valid for the relatively low head. The nappe's air pocket becomes isolated at the higher head, and sub-atmospheric pressure conditions may occur. This negative pressure enhances the weir's discharge efficiency; however, its influence on the structure's stability as a whole remains unclear [80, 81]. Some prototype structures, such as Malarce PKW, have installed a measurement device in the aeration network to estimate the air demand. Indeed, this PKW operates as the main spillway and is thus expected to discharge continuously

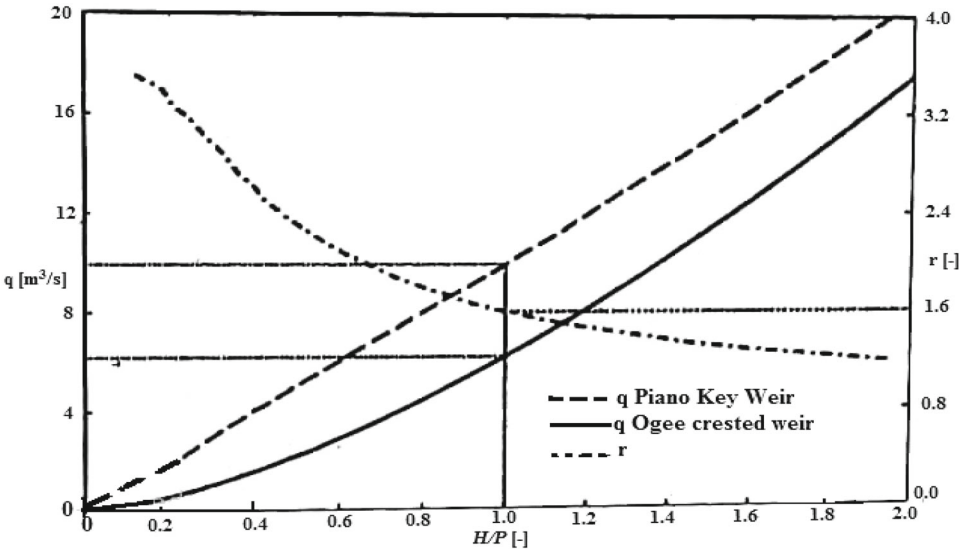


Fig. 6 Discharge per unit width q of a PK weir compared to an ogee crested-PKW, $P = P_i = P_o = 2$ m and $L/W = 5$; Ogee crested weir: design head 2 m. $C_s = 0.494$ (Adapted from Erpicum et al. [75])

[82]. According to Hien et al. [55], the cantilever or overhangs at the PKW's inlet or outlet keys perform better at nappe aeration than trapezoidal labyrinth weirs.

Vermeulen et al. [83] proposed a method of air demand emphasizing air driven by a rectangular channel stream dropping into a volume of free water surface [84, 85]. As of late, this new evaluation technique was endorsed by the degree of air interest during flooding on Malarce dam PKW**.** The empirical formula was developed by the author, which predicts the measure of airflow in a stream Qair (m^3/s) as:

$$Q_{\text{air}} = (q_w \times L_n) \times 0.26 \times \left(\frac{L_n}{P_n}\right) \times \left(\frac{d_j}{h_n}\right)^{0.446} \times \left(1 - \frac{V_m}{V_i}\right) \quad (7)$$

where, q_w = overflowing discharge through inlets or sidewalls (m^3/s), L_n = width of nappe (m), h_n = thickness of nappe (m), P_n = perimeter of nappe which is equal to P_n = 2*($L_n + h_n$) (m), d_n = height of falling jet (m), V_i = speed of jet impact (m/s), and V_m = Critical velocity required to entrain air. Nappe instabilities have been observed for fully aerated nappes [86, 87]. According to Lombaard [81], the nappe over the PKW formed a continuous water curtain over the lateral and downstream inlet key crests. There is currently no practical or standard design procedure that is available or accepted.

5.2 Energy Dissipation

The PKW is an ungated spillway structure, making it an efficient solution for increasing the evacuation carrying capacity during the rainy season. Any hydraulic system is prompting the release of water at high efficiency; all things considered, the high energy dissipation happens downstream of the weir. While designing energy dissipation structures downstream of a PKW, particular attention should be taken to avoid undesirable aspects such as scouring and cavitation [33, 88]. The first systematic experimental studies are ongoing with stepped spillway [88] or stilling basin [88, 89]. In order to promote design efforts, Eslinger and Crookston [31] were conducted an experimental study to understand the energy dissipation analysis over Type A PKW, and two equations (both are the functions of head-water ratio (H_t/P) and (W_i/W_o) are proposed to predict the relative residual energy downstream of PKWs.

Various measures are currently being investigated to improve energy dissipation and reduce downstream scour on PKW, including steps and baffles of different sizes [33, 88]. Various energy dissipators and their combinations can be studied further for effective energy dissipation at the lowest possible cost.

5.3 Scour, Sedimentation, and Debris aspects

Scour: Scouring is one of the essential design aspects downstream of the PKW. The residual flow energy downstream of a grade control structure will produce scour in an alluvial riverbed if no mitigation structure is installed. Many experimental studies have shown the scour depth or scour volume downstream of a PKW. Jüstrich et al. [32], reported a systematic experimental study on PKW toe scour, how to create the scour holes and ridges caused by PKWs with no scour defense. They found that the overall mechanism is jet-induced scour. At the toe of PKWs, scour depth and sediment removal are mainly determined by sediment characteristics, discharge, residual energy, and tailwater depth. Pfister et al. [89] examined the toe scour formation at the PKW and found that if the foundation on rock is not possible, the toe-scour occurring under flood discharges is relevant to weir stability.

Sedimentation: The influence of sediments on a PKW is not yet well specified, but it is actively studied. According to Sharma and Tiwari [90], the z-component of velocity (vertically upward in direction) at low flow rate values may help keep a remarkable amount of sediment in suspension and enhance the flushing capacity. The PKW appears to be self-cleaning, according to Tiwari and Sharma [91]; however, the effect may be too localized to enable the weir to behave as it would in its natural, sediment-free state. An experimental study was performed based on the upstream erosion and sediment passage over the PKW by Noseda et al. [92]. They stated that the PKWs are more effective than the labyrinth weir in terms of sediment ability. Gebhardt et al. [93] presented a study regarding sediment passage during floods.

Debris: Driftwood/Debris is wood washed by the action of floods, storms, and tides or waves on the shore or beach of a sea, lake, or river. Driftwood is a considerable nuisance in some waterfront areas. However, as it floats in the reservoir, the driftwood provides shelter and food for birds, fish, and other aquatic animals. The PKW's sensitivity regarding the debris and driftwood has been investigated by Pfister et al. [76], and a suggestion was made that the debris marginally affects the discharge efficiency of PKWs. Pfister et al. [94], took a closer look at the research on the impact of driftwood on PKW discharge quality in greater detail [76]. They proposed that some additional insights gained during the model investigation and prototype observations be included. PKWs are probably less sensitive to the effect of driftwood on discharge capacity than traditional weir types such as ogees. However, there is currently a scarcity of detailed observations that allow for direct comparison.

5.4 Scale effects, Approach Flow, and Hydrological Impact

Scale effects: The size effect on hydraulic structures occurs when a prototype hydraulic process is replicated in a laboratory-scale model due to disagreement with similarity laws. In this case, more than one model law plays a significant role over the free surface in experimental models. Machiels et al. [95], presented the scale effects on PKW, and they found that the scale effect creates significant variations in the model and prototype of PKW for the small discharges. To avoid these effects, Erpicum et al. [24] claim that model heads less than 0.03 m underestimate the discharge capacity of PKWs. That flow characteristics such as nappe formation and jet geometry must be accurately reproduced by 0.06 m. However, according to Tullis et al. [25], minimum heads to prevent scale effects differed for model scale and actual weir size. When considering scale effects for the PKW model, merely limiting the model geometric scale ratio is inadequate. Indeed, important considerations are the actual sizes of the prototype and model. The minimum dimensionless head (H/P) below which scale effects occur typically decreases with increasing the model size.

Approach flow: There is no substantial literature available on the approach flow, but most other dam spillways, whether they are ungated linear overflows or gated structures, can have a remarkable impact on the PKW's efficiency. Therefore, either computational modeling or physical modeling should explore its effect on flow patterns.

Hydrological impact: A PKW has high specific discharge efficiency at relatively low heads than a linear weir. Therefore, its response can be faster than a linear weir to an incoming hydrograph. Nonetheless, the effects of a PKW on the downstream flood regime have not yet been thoroughly studied, depending on the upstream reservoir or water body's characteristics [80].

5.5 Economics Aspect

At a relatively low head, PKW has a higher hydraulic efficiency than linear or any ungated weir. So, it is crucial to find out the optimal balance between efficiency and design cost. Paxson et al. [59] carried out a general exercise to determine the construction cost as low as possible at higher discharge efficiency. They stated that PKW is always the simplest and most appropriate solution. However, labyrinth spillways often offer an optimal solution depending on site conditions and operational constraints. To better understand the costs involved, they compared the cost of a 4-keys PKW and a 2-cycle labyrinth weir with a side angle ($\alpha = 15°$) for identical head-discharge characteristics. Due to the substantially reduced upstream and downstream footprints, the PKW needs 35 percent less concrete volume. They also observed that using a gated spillway would be more expensive than using a PKW due to the higher construction and furnishing costs. Since each project has its own set of opportunities and constraints, the suitability of a particular spillway form can be determined on a case-by-case basis. Lempérière and Vigny [22] presented a cost comparison between the PKW and Creager sill on the gravity dam. The volume of reinforced concrete is around half the volume of a typical labyrinth for the same savings. PKW requires approximately 0.5 m^3 of reinforced concrete and saves 1 m^3 of ordinary concrete to increase the discharge by up to one cubic meter per second compared to a Creager sill. PKW can be made of reinforced concrete or steel, even though the steel weir is limited in height (height up to 2–3 m). According to their research, the relative quantity of reinforced concrete and ordinary concrete is about one-third for a PKW of 2 m and one-fifth for a PKW of six meters.

Figure 7 shows the economic analysis presented by Lempérière and Vigny [22] between the PKW and Creager sill on the gravity dam. This economic analysis was based on cost and gained with a "K" value at constant water level and constant spillway length. The "K" reflects the PKW discharge coefficient ratio instead of the Creager sill discharge coefficient for an equivalent upstream water head. According to Lempérière et al. [20], the construction cost of a PKW per meter in a run-of-the-river scheme is approximately 20% of the cost per meter gated dam, even when the length is twice or three times that of the gated dam. As a result, the PKW has the potential to cut the cost of many run-of-river barriers in half.

6 Field Applicability

PKW is also used as a side weir in a channel or irrigation canal for flow measurement, and it shows greater discharge efficiency than other free flow alternatives [13] and [96]. PKWs can be used as a flow measuring device in a run-of-the-river scheme, off-taking canals, water treatment plants, sewage carrying tunnels, etc.

The application of PKWs offers several advantages, such as increasing discharge capacity up to three times compared with straight linear spillways, construction capability on existing and new gravity dams, economic efficiency, fewer maintenance costs, simple structure, and ease of construction [3].

7 Further Scope

From the above descriptive study, it is clear that many experimental studies have been conducted on the PKW based on their excellent understanding of hydraulics and parametric

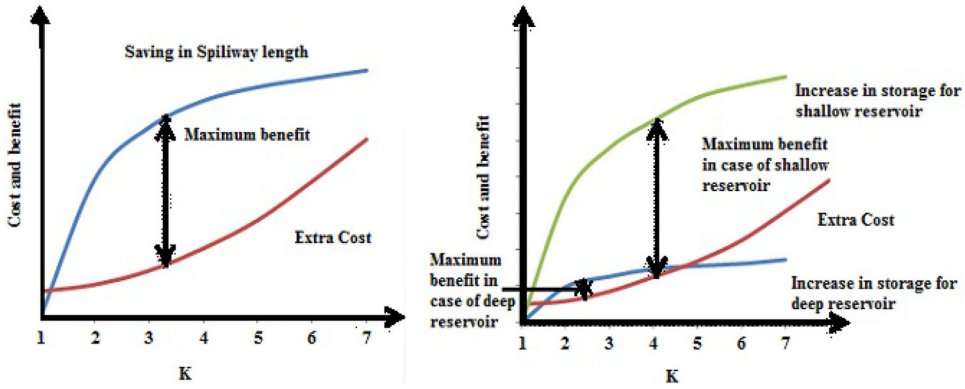


Fig. 7 Economic analysis with constant water level (left) and constant spillway length (right) (Adapted from Lempérière and Vigny [40])

geometry optimization. It can be exciting to use numerical modeling to study or even replace physical models, often more time-consuming and expensive.

The guidance on scouring processes specific to these types of weirs, the experiments with changing sill geometry are conducted to obtain a higher discharge capability and are under study. In addition, how do these weirs react to siltation in the upstream reservoir or channel, and the unit discharges needed to flush these sediments downstream, it can be an area of future work.

Structural aspects: construction of PKW in-situ depends on the availability of raw material near the site. The quality of raw material can be enhanced by using some admixture for making it like composites; it could be an area of future work.

The downstream flood regime (dissipation of energy) is another issue concerning PKWs has not yet been thoroughly studied: it can be an area of future work. Although few authors have addressed the risk of floating debris, it merits further investigation. Experiments with trunk roots have been conducted; however, more research is required if these schemes are located near forest-covered catchment areas or in inhabited areas.

These parameter subjects, such as approach flow, aeration, tailwater submergence, and the hydrodynamics and transient flow behavior accumulation effect, have been considered and still need to be understood in order to reach a conclusion and make recommendations.

8 Conclusions

PKW is a cost-effective solution for rehabilitation and new dam projects with a high level of constraints, for instance, limited space, small reservoir level, high specific flood discharge, etc. PKW is free-surface weirs with specific flows 2 to 4 times that of a standard linear weir, increasing the discharging capacity of both existing and new dams. It is used for run-of-the-river schemes to increase the storage capacity of many existing reservoirs and increase spillway capacity to reduce dam risk failures, with approximately 0.5 m^3 of reinforced concrete used for each additional increase of 1 m^3/s depending on site conditions. PKW is more suitable for new dams or hydraulic structures as a free flow spillway and convenient for spillway rehabilitation on the existing concrete and gravity dams already equipped with a free flow spill than to its alternatives. PKWs are structurally sound, cost-effective, and have a wide range of applications. Turbulence generated near the upstream end of these weirs gives these weirs self-cleaning properties. The installation of several PKW projects worldwide has demonstrated a general design subject to the conditions of their construction. Although a few topics necessitate future scope of work, PKW provides an excellent alternative for dam rehabilitation when compared to its alternatives.

Acknowledgements We express our unique sense of gratitude and respect toward Dr. P. S. Mahar, Professor, Civil Engineering Department, GBPUA, and T. Pantnagar, for providing the guidance and support in writing this paper.

Author's contributions All authors have drafted the article or revised it critically for important intellectual content. They all have read and agreed to the published version of the manuscript.

Funding This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors (no funds were taken from any Organization).

Data Availability Statement No data, models, or code was generated or used during the study.

References

1. Novak, P.; Guinot, V.; Jeffrey, A.; Reeve, D. E.: Hydraulic modelling: an introduction: principles, methods and applications (2010)
2. Subramanya, K.: Flow in open channels. Tata McGraw Hill Education Private Limited, New Delhi (2010)
3. Ouamane, A.; Lempérière, F.: Design of a new economic shape of weir. In: Proceedings of the International Symposium of Dams in

the Societies of the 21st Century, pp. 463–470. Barcelona, Spain (2006)
4. Lempérière, F.; Ouamane, A.: The piano keys weir: a new cost-effective solution for spillways. Hydropower Dams **10**(5), 144–149 (2003)
5. Laugier, F.; Vermeulen, J.; Blancher, B.: 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., Pfister, M. (eds). Labyrinth and Piano Key Weirs III—PKW 2017, pp. 37–51. CRC Press, London (2017)
6. Ho Ta Khanh, M.: 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. (eds) Labyrinth and Piano Key Weirs III—PKW 2017, pp. 3–16. CRC Press, London (2017)
7. Erpicum, S.; Laugier, F.; Boillat, J.L.; Pirotton, M.; Reverchon, B.; Schleiss, A. J.: Labyrinth and Piano Key Weirs—PKW-2011. CRC Press, London (2011)
8. Erpicum, S.; Laugier, F.; Pfister, M.; Pirotton, M.; Cicero, G.; Schleiss, A. J.: Labyrinth and Piano Key Weirs II—PKW-2013. CRC Press, London (2013)
9. Erpicum, S.; Laugier, F.; Ho Ta Khanh, M.; Pfister, M.: Labyrinth and piano key weirs III—PKW-2017. CRC Press, London (2017)
10. Crookston, B.; Erpicum, S.; Tullis, B.; Laugier, F.: Hydraulics of labyrinth and piano key weirs 100 years of prototype structures, advancements and future research needs. J. Hydrol. Eng. ASCE **145**, 12 (2019)
11. Chemaly, A. G.: Lessons learned from the planning, design, construction, operation and maintenance (O&M), monitoring and safety evaluation of dams. In: Proceedings of the Annual SANCOLD Conference 2017, pp 271–295 (2017)
12. Mabela, M.J.: The raising of Hazelmere Dam using a Piano Key Weir. In: Procedings of the Annual SANCOLD Conference 2017, pp. 181–193 (2017)
13. Karimi, M.; Attari, J.; Saneie, M.; Ghazizadeh, M.R.J.: Side weir flow characteristics: comparison of Piano Key, Labyrinth, and linear types. J. Hydrol. Eng **144**, 12 (2018)
14. Barcouda, M.; Cazaillet, O.; Cochet, P.; Jones, B. A.; Lacroix, S.; Laugier, F.; Odeyer, C.; Vingny, J. P.: Cost-effective increase in storage and safety of most dams using fuse gates or PK Weirs. In: Proceedings of the 22nd Congress of ICOLD., Barcelona, Spain (2006)
15. Ho Ta Khanh, M.: The piano keyweirs: 15 years of research and development—Prospect. Labyrinth and Piano Key weirs II—PKW2013, pp. 3–14. CRC Press, London (2013)
16. Bieri M.; Federspiel M.; Boillat J.; Houdant, B.; Delorme, F.: Spillway capacity upgrade of Gloriettes dam: environmental integration and energy dissipation. In: Proceedings of HYDRO-2009, Lyon, France (2009)
17. Laugier, F.; Lochu, A.; Gille, C.; Leite Ribeiro, M.; Boillat, J.L.: Design and construction of a labyrinth PKW spillway at St-Marc Dam. France. Intl. J. Hydropower Dams **15**(5), 100–107 (2009)
18. Vermeulen, J.; Laugier, F.; Faramond, L.; Gille, C.: Lessons learnt from design and construction of EDF first Piano Key Weirs. In: Proceedings of the International Conference on Labyrinth and Piano Key Weirs. Liège B, pp. 215–224. CRC Press, Boca Raton (2011)
19. Schleiss, A. J.: From Labyrinth to Piano Key Weirs—A historical review. In: Proceedings of the International Conference on Labyrinth and Piano Key Weirs (PKW 2011), pp. 3–15. CRC Press, Taylor & Francis Group (2011)
20. Lempérière, F.; Vigny, J. P.; Ouamane, A.: General comments on Labyrinth and Piano Key Weirs: The past and present. In: Proceedings of the International Conference on Labyrinth and Piano Key Weirs (PKW 2011), pp. 17–24. CRC Press, Taylor & Francis Group (2011)
21. Abhash, A.; Pandey, K.K.: A review of Piano Key Weir as a superior alternative for dam rehabilitation A review of Piano Key Weir as a superior alternative for dam rehabilitation. ISH J. Hydraul. Eng. **00**(00), 1–11 (2020)
22. Lempérière, F.; Vigny, J.: General comments on labyrinth and Piano Keys Weirs–The future. Labyrinth and Piano Key weirs-PKW 2011, pp. 289–294 (2011)
23. Pfister, M.; Capobianco, D.; Tullis, B.; Schleiss, A.: Debris-Blocking sensitivity of Piano Key weirs under reservoir-type approach flow. J. Hydraul. Eng. **139**(11), 1134–1141 (2013)
24. Erpicum, S.; Tullis, B.P.; Lodomez, M.; Archambeau, P.; Dewals, B.J.; Pirotton, M.: Scale effects in physical piano key weir models. J. Hydraul. Res. **54**(6), 692–698 (2016)
25. Tullis, B.P.; Crookston, B.M.; Young, N.; Asce, M.: Scale effects in free-flow nonlinear weir head–discharge relationships. **146**(2), 1–9 (2020)
26. Dabling, M.R.; Tullis, B.P.: Piano Key Weir submergence in channel applications. In: Proceedings of the International Workshop on Piano Key Weir for In-Stream Storage and Safety, New Delhi, India, pp. 13–23 (2012)
27. Erpicum, S.; Archambeau, P.; Dewals, B.; Pirotton, M.: Hydraulics of piano key weirs: a review. In: Erpicum, S.; Laugier, F.; Ho Ta Khanh, M.; and Pfister, M. (eds) Labyrinth and piano key weirs III—PKW 2017, pp. 27–36. CRC Press, London (2017)
28. Anderson, R.M.;Tullis B.P.: Influence of piano key weir geometry on discharge. In: Proceedings International Workshop on Labyrinth and Piano Key Weirs, Liège, Belgium (2011)
29. Kabiri-Samani, A.; Javaheri, A.: Discharge coefficients for free and submerged flow over Piano Key weirs. J. Hydraul. Res. **50**(1), 114–120 (2012)
30. Mario, O.: Discharge coefficient of piano key weir from experimental and numerical models. In: E. Proceedings of the 36th IAHR World Congress (2015).
31. Eslinger, K.R.; Crookston, B.M.: Energy dissipation of type a piano key weirs (2020)
32. Justrich, S.; Pfister, M.; Schleiss, A.J.: Mobile riverbed scour downstream of a piano key weir. J. Hydraul. Eng. **142**(11), 04016043 (2016)
33. Silverstri, A.; Erpicum, S.; Archambeau, P.; Dewals, B.; Pirotton, M.: Stepped spillway downstream of a Piano Key weir—Critical length of uniform flow. In: International Workshop on Hydraulic Structures, pp. 99–107. Bundesanstalt fur, Wasserbau, Karlsruhe, Germany (2013a)
34. Noui, A.; Ouamane, A.: Study of optimization of the piano key weir. In: Labyrinth and piano key weirs II pp. 175–182. CRC/Balkema, Leiden, Netherlands (2013)
35. Machiels, O.; Erpicum, S.; Archambeau, P.; Dewals, B. J.; Pirotton, M.: Influence of the Piano Key Weir height on its discharge capacity. In: Proceedings of the International Conference on Labyrinth and Piano Key Weirs Liege B, pp. 59–66. CRC Press, Boca Raton (2011a)
36. Michael, P.; Schleiss, A.J.: Estimation of A-type Piano Key weir rating curve. In: Proceedings of the 2nd International Workshop on Labyrinth and Piano Key Weirs (PKW 2013), pp. 139–147. CRC Press, Taylor & Francis Group (2013)
37. Leite Ribeiro, M.; Pfister, M.; Boillat, J. L.; Schleiss, A. J.; Laugier, F.: Piano Key Weirs as efficient spillway structure. In: 24th ICOLD Congress Kyoto (Q94, R13), pp. 1–10 (2012b)
38. Leite Ribeiro, M.; Pfister, M.; Schleiss, A.J.; Boillat, J.L.: Hydraulic design of A-type Piano Key Weirs. J. Hydraul. Res. **50**(4), 400–408 (2012)
39. Denys, F.J.M.: Piano Key Weirs spillway standard design principles and flow induced vibrations. In: Proceedings of the 3rd International Workshop on Labyrinth and Piano Key Weirs (PKW 2017), pp. 119–126. CRC Press, Taylor & Francis Group (2017)
40. Lempérière, F.: New Labyrinth weirs triple the spillways discharge (2009). http://www.hydrocoop.org. (2010)

41. Machiels, O.: Experimental study of the hydraulic behaviour of Piano Key Weirs. Ph.D. Thesis ULgetd-09252012-224610, University of Liege (Belgium) (2012)
42. Anderson, R.M.: Piano key weir head discharge relationships. All Graduate Theses and Dissertations, Paper 880, Utah State University (2011)
43. Anderson, R.M.; Tullis, B.P.: Piano key weir hydraulics and labyrinth weir comparison. J. Irrig. Drain. Eng. ASCE **139**(3), 246–253 (2013)
44. ICOLD, ICOLD Bulletin 144: Cost Savings in Dams. Appendix 2. International Commission on Large Dams (2010)
45. Pralong, J.; Vermeulen, J.; Blancher, B.; Laugier, F.; Erpicum, S.; Machiels, O.; Pirotton, M.; Boillat, J. L.; Leite Ribeiro, M.; Schleiss, A. J.: A naming convention for the Piano Key Weirs geometrical parameters. In: Proceedings of the International Conference on Labyrinth and Piano Key Weirs (PKW 2011), pp. 271–278. CRC Press, Taylor & Francis Group (2011a)
46. Phillips, M.A.; Lesleighter, E.J.: Piano Key Weir spillway: upgrade option for a major dam. In: Proceedings of the 2nd International Workshop on Labyrinth and Piano Key Weirs (PKW 2013), pp. 159–168. CRC Press, Taylor & Francis Group (2013)
47. Pfister, M,; Schleiss, A.J.: Comparison of hydraulic design equations for A-type Piano Key weirs. In: Proceeding International conference Water Storage and Hydropower Development for Africa. Addis Ababa, Ethiopia (2013a)
48. Erpicum, S.; Archambeau, P.; Pirotton, M.; Dewals, B.J.: Geometric parameters influence on Piano Key Weir hydraulic performances. In: 5th International Symposium on Hydraulic Structures and Society: Engineering Challenges and Extremes, Brisbane, Australia, pp. 25–27 (2014). ISBN 9781742721156
49. Mehboudi, A.; Attari, J. ; Hosseini, S.A.: Flow regimes over trapezoidal Piano Key Weirs. In: Erpicum, S., Laugier, F., Ho Ta Khanh, M., Pfister, M. (eds. Labyrinth and Piano Key Weirs III—PKW 2017, pp. 65–74. CRC Press, London (2017)
50. Laiadi, A.; Athmani, B.; Belaabed, F.; Ouamane, A.: The effect of the geometric shape of the alveoli on the performance of Piano Key Weirs. In: Erpicum, S., Laugier, F., Ho Ta Khanh, M., Pfister, M. (eds) Labyrinth and piano key weirs III—PKW 2017, pp. 93–100. CRC Press, London (2017)
51. Lempérière, F.; Jun, G.: Low cost increase of dams storage and flood mitigation: the piano keys weir. In: Q.53 R. 2.06 International Commission on Irrigation and Drainage Nineteenth Congress Beijing (2005)
52. Machiels, O.; Erpicum, S.; Dewals, B. J.; Archambeau, P.; Pirotton, M.: Influence of the alveoli slopes on the discharge capacity of Piano Key Weirs. In: 34th IAHR World Congress, Brisbane, Australia (2011b)
53. Machiels, O.; Pirotton, M.; Archambeau, P.; Dewals, B.; Erpicum, S.: Experimental parametric study and design of Piano Key Weirs. J. Hydraul. Res. **52**(3), 326–335 (2014)
54. Leite Ribeiro, M.; Bieri, M.; Boillat, J.L.; Schleiss, A.J.; Singhal, G.; Sharma, N.: Discharge capacity of piano key weir. J. Hydraul. Eng. ASCE **138**(2), 199–203 (2012)
55. Hien, T.C.; Son, H.T.; Khanh, M.H.T.: Results of some piano keys weir hydraulic model tests in Vietnam. In: Proceedings of the 22nd Congress of ICOLD, Barcelona, Spain (2006)
56. Le Doucen, O.; Leite Ribeiro, M.; Boillat, J. L.; Schleiss, A.; Laugier, F.: Etude parametrique de la capacite des PK-Weirs. Modeles physiques hydrauliques-outils indispensables du XXIe siècle, Lyon: SHF (2009)
57. Cicero, G.M.; Vermeulen, J.; Laugier, F.: Influence of some geometrical parameters on piano key weir discharge efficiency. In: 6th International Symposium on Hydraulic Structures and Water System Management, Portland, Oregon, USA, 27–30 June 2016, ISBN 978-1-884575-75-4 (2016)
58. Khassaf, S.I.; Aziz, L.J.; Elkatib, Z.A.: Hydraulic behavior of piano key weir type B under free flow conditions. Int. J. Sci. Technol. Res. **5**, 03 (2016)
59. Paxson, G.S.; Tullis, B.P.; Hertel, D.J.: Comparison of Piano Key Weirs with labyrinth and gated spillways: hydraulics, cost, constructability and operations. In: Proceedings of the 2nd International Workshop on Labyrinth and Piano Key Weirs (PKW 2013). CRC Press, Taylor & Francis Group, pp. 123–130 (2013)
60. Machiels, O.; Erpicum, S.; Archambeau, P.; Dewals, B.; Pirotton, M.: Parapet wall effect on Piano Key Weir efficiency. J. Irrigat. Drain. Eng. ASCE **139**(6), 506–511 (2013)
61. Falvey H.T.: Hydraulic design of labyrinth weirs. ASCE Press. (American Society of Civil Engineers) (2003)
62. Cicero, G.M.; Delisle, J.R.: Effects of the crest shape on the discharge efficiency of a type-A Piano Key weir. In: Proceedings of the 2nd International Workshop on Labyrinth and Piano Key Weirs (PKW 2013), pp. 41–48. CRC Press, Taylor & Francis Group (2013)
63. Ali, M.S.; Mansoor, T.: Performance evaluation of piano key weir. In: Hydro-2015 20th International Conference on Hydraulics, Water Resources and River Engineering. IIT Roorkee, India (2015)
64. Erpicum, S.; Machiels, O.; Archambeau, P.; Dewals, B.; Pirotton, M.: 1D numerical modeling of the flow over a Piano Key Weir. In: Proceedings of the International Conference on Labyrinth and Piano Key Weirs. Liege B, pp. 151–158. CRC Press, Boca Raton (2011a)
65. Oertel, M.; Tullis, B.; P.: Comparison of Piano Key Weir discharge coefficients from experimental and numerical models. In: Proceedings of the 3rd IAHR Europe Congress, Porto, Portugal (2014)
66. Anderson, R.M.; Tullis, B.P.: Comparison of piano key and rectangular labyrinth weir hydraulics. J. Hydraul. Eng. ASCE **138**, 358–361 (2012)
67. Li G.; Li, S.; Hu, Y.: The effect of the inlet/outlet width ratio on the discharge of piano key weirs. J. Hydraul. Res. 1686 (2019)
68. Al-Baghdadi, MohdB.: 2D Visualization of flow phenomenon over individual inlet and outlet keys of piano key weir by CFD modeling. Int. J. Energy Environ. **10**(1), 15–24 (2019)
69. Denys, F.; Basson, G.: Transient hydrodynamics of piano key weirs. In: 7th International Symposium on Hydraulic Structures. Aachen, Germany, 15–18 May 2018, ISBN 978-0-692-13277-7 (2018)
70. Olyaie, E.; Heydari, M.; Banejad, H.; Chau, K.W.: A laboratory investigation on the potential of computational intelligence approaches to estimate the discharge coefficient of piano key weir. J. Rehabilit. Civ. Eng. **6**, 1–20 (2018)
71. Kashkaki, Z.; Banejad, H.; Heydari, M.: Application of ANN in estimating discharge coefficient of circular piano key spillways. Soft Comput. Civ. Eng. **2**(3), 39–49 (2018)
72. Zounemat-Kermani, M.; Mahdavi-Meymand, A.: Hybrid metaheuristics artificial intelligence models in simulating discharge passing the piano key weirs. J. Hydrol. **569** (2019)
73. Vischer, D.L.; Hager, W.H.: Dam Hydraulics, 1st edn. Wiley, New York (1998)
74. Blancher, B.; Montarros, F.; Laugier, F.: Hydraulic comparison between piano key weirs and labyrinth spillways. In: Proceedings of the International Conference on Labyrinth and Piano Key Weirs (PKW 2011), pp. 141–150. CRC Press, Taylor & Francis Group (2011)
75. Erpicum, S.; Machiels, O.; Dewals, B.; Archambeau, P.; Pirotton, M.: Considerations about the optimum design of PKW. In: Proceeding International conference Water Storage and Hydropower Development for Africa (Africa 2013). Addis Ababa, Ethiopia (2013a)
76. Pfister, M.; Schleiss, A.J.; Tullis, B.: Effect of driftwood on hydraulic head of Piano Key weirs. In: Erpicum et al. (eds) Proceedings of the 2nd International Workshop on Labyrinth and

Piano Key Weirs-PKW 2013, Paris, France, 20–22 november 2013, pp. 255–265. CRC Press, Boca Raton (2013b)
77. Gameson, A.L.H.: Weirs and aeration of rivers. J. Inst Water Eng **11**(5), 477–490 (1957)
78. Tebbutt, T.H.Y.: Reaeration performance of stepped cascades. J. Inst. Water Eng. Sci. **31**, 285–297 (1977)
79. Macaitis, B.; Reclamation, M.W.: Side stream elevated pool aeration station design. MOO, pp. 199 (1990)
80. Laugier, F.; Vermeulen, J.; Lefebvre, V.: Overview of piano key weirs experience developed at EDF during the past few years. In: Proceedings of the 2nd International Workshop on Labyrinth and Piano Key Weirs (PKW 2013), pp. 213–226. CRC Press, Taylor & Francis Group (2013)
81. Lombaard, J.: Evaluation of the influence of aeration on the discharge capacity and flow-induced vibrations of piano key weir spillways. Masters of Engineering thesis. Copyright© 2020, Stellenbosch University (2020)
82. Pinchard, T.; Farges, J.L.; Boutet, J.M.; Lochu, A.; Laugier, F.: Spillway capacity upgrade at Malarce dam: construction of an additional piano key weir spillway. Labyrinth and Piano Key Weirs II—PKW 2013. In: Erpicum et al. (eds) pp. 243–252 (2013). London: Taylor & Francis Group
83. Vermeulen, J.; Lassus, C.; Pinchard, T.: Design of a Piano Key Weir aeration network. In: Proceedings of the 3rd International Workshop on Labyrinth and Piano Key Weirs III—PKW 2017, Qui Nhon, Vietnam, 22–24, pp. 127–134. London: Taylor & Francis Group (2017)
84. Ervine, D.; Elsawy, E.: The effect of a falling nappe on river aeration. In: Proceedings of the 16th IAHR Congress, Sao Paulo, Brazil (1975)
85. Falvey, H.T.: Air-water flow in hydraulic structures. NASA STI/Recon Technical Report N, 81. Gabriel-Martin (1980)
86. Crookston, B.M.; Tullis, B.P.: Hydraulic design and analysis of labyrinth weirs II: nappe aeration, instability, and vibration. J. Irrigat. Drain. Eng. **139**, 371–377 (2013)
87. Lodomez, M.; Pirotton, M.; Dewals, B.; Archambeau, P.; Erpicum, S.: Nappe oscillations on free-overfall structures: experimental analysis. J. Hydraul. Eng. **144**, 3 (2018)
88. Ho Ta Khanh, M.; Sy Quat. D.; Xuan Thuy D.: PK weirs under design and construction in Vietnam (2010). Labyrinth and piano key weirs-PKW 2011, pp. 225–232. CRC Press, London (2011)
89. Pfister, M.; Justrich, S.; Schleiss, A.J.: Toe scour formation at Piano Key weirs. In: Erpicum, S., Laugier, F., Ho Ta Khanh, M., Pfister, M. (eds) Labyrinth and Piano Key Weirs III—PKW 2017. CRC Press, Leiden (2017)
90. Sharma, N.; Tiwari, H.: Experimental study on vertical velocity and submergence depth near Piano Key weir. In: Erpicum et al. (eds) Labyrinth and Piano Key Weirs II—PKW 2013–2014, pp. 93–100. Taylor & Francis Group, London, ISBN (2013)
91. Tiwari, H.; Sharma, N.: Turbulence study in the vicinity of piano key weir: relevance, instrumentation, parameters and methods. Appl. Water Sci. **7**(2), 525–534 (2017)
92. Noseda, M.; Stojnic, I.; Pfister, M.; Schleiss, A.J.: Upstream Erosion and Sediment Passage at Piano Key Weirs. J. Hydraul. Eng. **145**(8), 04019029 (2019)
93. Gebhardt, M.J.; Herbst, J.; Merkel, S.; Belzner, F.: Sedimentation at labyrinth weirs—an experimental study of the self-cleaning process. J. Hydraul. Res. (2018). https://doi.org/10.1080/00221686.2018.1494053
94. Pfister, M.; Tullis, B.; Schleiss, A.J.: Closure to debris-blocking sensitivity of piano key weirs under reservoir-type approach flow, by Michael Pfister, Damiano Capobianco, Blake Tullis, and Anton J. Schleiss. J. Hydraul. Eng. 141(10) (2015)
95. Machiels, O.; Erpicum, S.; Archambeau, P.; Dewals, B.J.; Pirotton, M.: Large scale experimental study of piano key weirs. In: Proceedings of 33rd of IAHR., Vancouver, Canada (2009)
96. Saghari, A.; Saneie, M.; Hosseini, K.: Experimental study of one- and two-cycle trapezoidal piano-key side weirs in a curved channel. Water Supply Press IWA Publishing, New York (2019)

Implementation of Dual Tree Complex Wavelet Transform with Mean Energy Features to detect Alzheimer's Disease

Neha Garg
Department of Electronics and Communication
Delhi Technological University
Delhi,India
ngcoolneha@gmail.com

Mahipal Chaudhary
Department of Electronics and Communication
Delhi Technological University
Delhi,India
msc_1976@yahoo.com

2021 7th International Conference on Signal Processing and Communication (ICSC) | 978-1-6654-2739-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICSC53193.2021.9673484

***Abstract*—Alzheimer's disease is an inveterate neurodegenerative disorder that results to gradual loss of cognitive power and memory. It is a chronic disease that reduces the count of neurons in the brain which may lead to patients' death in due course of time. To reduce the threat of this disease, it is substantial to detect it at its initial stage. Alzheimer's Disease introduces the structural atrophies in the brain, causing reduction in grey matter densities, reduction of cortical surface and altering the size, shape, and volume of hippocampus. Magnetic resonance imaging is a non-intrusive neuroanatomical technique that shows inner structural differences with good resolution and contrast. In this paper, Dual Tree Complex Wavelet Transform is used for obtaining the features from Magnetic Resonance Images. This paper presents a novel approach for Alzheimer's Disease and Normal Control classification using the mean energies of real and imaginary components of complex wavelet coefficients taken individually as features. Experimental results with seven different classifiers show that mean energy obtained from real and imaginary components of complex wavelet co-efficient taken separately as features gives more promising result compared to mean energy of real component of complex wavelet coefficient taken as a whole. Through the proposed approach using Open Access Series of Imaging Studies dataset with k-Nearest Neighbor classifier 97.3% Prediction accuracy is achieved along with 98.2% sensitivity and 96.4% specificity while using the same data set with Cubic Support Vector Machine classifier 97.3% prediction accuracy is achieved with 96% sensitivity and 99% specificity. The proposed approach is simple in comparison to existing complex algorithms still it gives better results with regards to accuracy, sensitivity, and specificity.**

Keywords— Classification, Alzheimer's Disease, k-Nearest Neighbor, Cubic Support Vector Machine, Dual Tree Complex Wavelet transform

I. INTRODUCTION

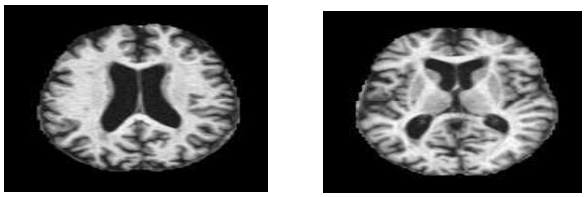
Alzheimer's disease (AD) is a vital, incorrigible, and a neurodestructive disorder, that is frequently sustained in aged persons. It results in slowly reducing their mental ability in terms of their cognitive power, social and behavioural skills [1]. In AD neuronal loss takes place, thus lot of structural changes happen in various regions of the brain including hippocampus [2] [3]. Presently, AD is not curable; however, its progression can be slowed using some preventive measures [4]. Magnetic Resonance Imaging (MRI) is an effective non-invasive neuroimaging technique that is used as a biomarker for AD. MRI provides inner structural variations in the human brain, inner organs, and various other tissues with good resolution and contrast [5]. It's very essential to extract optimum and relevant features from these images for detecting AD at its early stage. Till now more focus is on the methods like independent component analysis (ICA) [6], Fourier transform [7] and wavelet transform [8] for extracting relevant features from these images.

Chaplot et al. [9] used coefficients of second level Db4 wavelet decomposition as features for AD classification and obtained 94% accuracy with the neural network self-organizing maps (SOM). Dahshan et al. [8] used Discrete Wavelet Transform (DWT) with Principal Component Analysis (PCA) and obtained remarkable results with classifiers k-Nearest Neighbor (k-NN) and Feed forward back propagation network (FP-ANN). Zhang et al. [10] applied PCA on coefficients at third level decomposition of Haar wavelet in his work and calculated classification accuracy with Back propagation neural network. Jha et al. [11] used Dual Tree Complex Wavelet Transform (DTCWT) for extracting detail coefficients at fifth level of decomposition and applied PCA on coefficients considering them as features to reduce their dimensions. They obtained 90.06 ± 0.01 % accuracy with Feed forward Neural Network (FNN)classifier. Further on Alam et al. [12] applied PCA with Linear Discriminant Analysis (LDA) on detail coefficients extracted at fifth level of decomposition and obtained 96.68 ± 1.44% accuracy with Twin support vector machine classifier. Jha et al. [13] modified this work by using Extreme Learning Machine (ELM) for classification and obtained 95.72 % accuracy. Bhasin et al. [14] used 3D DWT with Local Binary Pattern-20 (LBP-20) for extracting features. Topannavar et al. [15] used DTCWT with Rotated complex wavelet filter (RCWF) for extracting Entropy, Energy and Standard deviation as features. Primarily the state-of-the-art methods use DWT for extracting relevant features. DWT is better than Fourier transform as it provides time and frequency domain analysis of any signal simultaneously while Fourier transform provides only frequency domain analysis. But DWT lacks in directional sensitivity, shift sensitivity, and in phase information. So, in the proposed work DTCWT is used for feature extraction which is better than DWT as it provides shift invariance and high directionality [16]. All the MRI images are pre-processed for getting good results.

In this work we have introduced two Feature groups. Feature group1 consist of features related to the mean energy

978-1-6654-2739-5/21/$31.00 ©2021 IEEE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:42:22 UTC from IEEE Xplore. Restrictions apply.



(a) (b)

Fig. 1. Sample of Dataset (Transverse view) of (a) Normal Control and (b) Alzheimer's Disease

of real component of complex wavelet coefficient obtained in all sub-bands up to 5th level of decomposition and feature group 2 includes the features related to mean energy of real and imaginary component of complex wavelet coefficient. To indicate the efficacy of our proposed method, we have used different classifiers for assessing performance parameters which includes accuracy, sensitivity, specificity [17],[18],[19]. We have plotted a bar chart to give an ease of understanding around the comparative study done for all different classifiers with both feature groups. The main contribution of the paper is summarized as

- We propose an optimum set of features which includes mean energy obtained from real and imaginary component of complex coefficient that can classify AD and Normal Controls (NC) with good accuracy.
- Feature group 2 provides the good accuracy, sensitivity, and specificity with different classifiers over Feature group 1.
- Existing method used very complex strategy which includes PCA, LDA to achieve high accuracy, sensitivity and specificity for AD and NC classification. But Proposed method uses simple methodology to extract features related to mean energy of real and imaginary component of complex coefficient of all six sub-bands at each level.

The subsequent segments of this paper talks about: Segment 2 explains the dataset [20], implementation of 2D DTCWT and proposed method while Segment 3 explains the classification performance of proposed method with different classifiers and comparison of proposed method with four state of art methods and finally, Segment 4 concludes the paper with future direction.

II. Material And Methods

A. Dataset

In this work we have used 112 2D MRI images for AD and NC respectively. These images are center slices of 3D MRI of 28 AD and NC subjects, each subject contributing 4 slices, thus forming 112 2D MRI images in each AD and NC class. 3D MRI images for AD and NC are downloaded from Open Access Series of Imaging Studies (OASIS) dataset [20]. OASIS is a data bank, which is designed for collecting Brian MRI datasets and make this data approachable by the scientific community. The MRI data can be downloaded from http://www.oasis-brains.org. A sample of AD and NC MRI image is shown in Fig. 1. OASIS dataset consists of longitudinal and cross- sectional MRIs. In proposed work, 56 cross-sectional MRI samples (including 28 ADs and 28 NCs) are utilized as these are good for automatic classification in comparison of longitudinal MRIs. Table 1 shows the statistical details of the patient included in the study. Both men and women with dextral habits are selected in the study. The statistical table provides the details about the number of subjects, mini mental state examination (MMSE), gender (M/F), age, clinical dementia rating (CDR). Cognitive impairment and dementia are monitored through a short survey known as MMSE [21]. It comprises of straightforward questions in several fields like arithmetic, comprehension, time, language utilization, place, reiterating words and basic motor skills. While the numeric scale used to measure the acuteness of signs of dementia is known as CDR [22].

TABLE I: Statistical Details Of The Subjects

Factors	Normal	AD
No of Patients	28	28
Age	84.40 (76-96)	82.11(76-96)
CDR	0	0.5
MMSE	28.72(25-30)	24.82(18-30)
Gender(M/F)	12/16	9/19

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:42:22 UTC from IEEE Xplore. Restrictions apply.

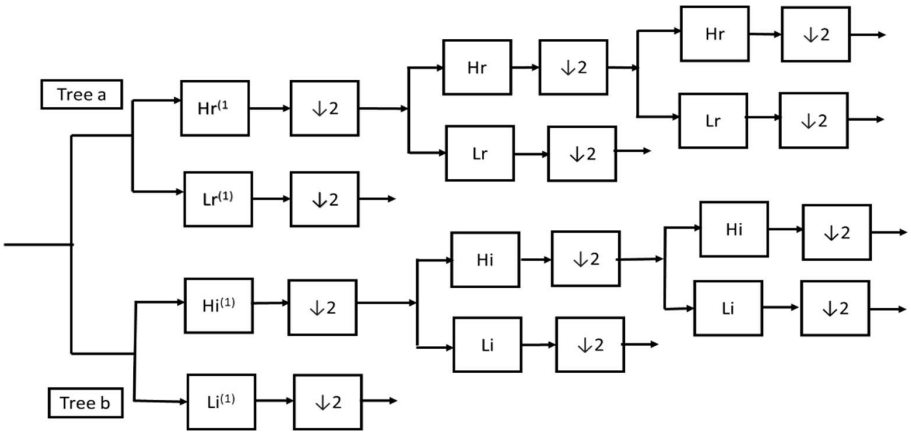


Fig.2. Implementation of one dimensional DTCWT

B. Dual Tree Complex Wavelet Transform (DTCWT)

Wavelet Transform is an effective technique for feature extraction. Proposed work used DTCWT for feature extraction. Mean energy obtained from real and imaginary component of complex wavelet coefficient from all six sub-bands at each decomposition level up to 5th level is calculated. To observe perfect number of decomposition level is always a great challenge. Large Decomposition level provides more information creating a problem of overfitting and small decomposition level provide less information. In Proposed method decomposition is done up to 5th level. We have checked classification accuracy from level 1 to level 6 of decomposition and observed maximum accuracy with fifth level of decomposition. Dual-tree complex wavelet transform (DTCWT), is highly shift invariant in nature and has high directional selectivity. It was first instigated by Kingsbury in 1998[23][24]. One dimensional DTCWT includes two real DWTs, one showing the real component and second one showing the imaginary component of the wavelet transform. Along with this, two filter banks, analysis filter banks for signal decomposition and synthesis filter banks for signal reconstruction are used in each tree [16][23]. Fig. 2 shows implementation of 1 dimensional DTCWT. Lr (1), Lr, Li (1), and Li indicate low-pass filters while H (1), Hr, Hi (1), and Hi indicate high-pass filters with 'i' and 'r' indicating imaginary and real respectively in both filters. These filters are designed in such a way that Tree b wavelet becomes Hilbert Transform of Tree a wavelet. Analysis filter in first stage of each tree is different from the analysis filters in its subsequent stages. These are required for perfect reconstruction.

A 2D DTCWT is described for two dimensional signals by two complex wavelets $\chi(u)$ and $\chi(v)$ which can be written as are $\chi(u,v) = \chi(u)(v)$, where

$\chi(u) = \chi_m(u) + j\chi_n(u)$ and $\chi(v) = \chi_m(v) + j\chi_n(v)$ (1)

where $\chi_m(*)$ and $\chi_n(*)$ represents the real and imaginary component of each complex wavelet.

Then the following expression can be obtained [16]:

$$\chi(u,v) = [\chi_m(u) + j\chi_n(u)][\chi_m(v) + j\chi_n(v)] \quad (2)$$

$$= \chi_m(u)\chi_m(v) - \chi_n(u)\chi_n(v) + j[\chi_n(u)\chi_m(v) + \chi_m(u)\chi_n(v)] \quad (3)$$

DTCWT has six oriented complex wavelets at each decomposition level, whose real component is expressed as:

$$\Psi_i(u,v) = \frac{1}{\sqrt{2}}(\chi_{1,i}(u,v) - \chi_{2,i}(u,v)) \quad (4)$$

$$\Psi_{i+3}(u,v) = \frac{1}{\sqrt{2}}(\chi_{1,i}(u,v) - \chi_{2,i}(u,v)) \quad (5)$$

Here i = 1, 2, 3 and

$$\chi_{1,1}(u,v) = \phi_m(u)\chi_m(v) \quad \chi_{2,1}(u,v) = \phi_n(u)\chi_n(v) \quad (6)$$

$$\chi_{1,2}(u,v) = \chi_m(u)\phi_m(v) \quad \chi_{2,2}(u,v) = \chi_n(u)\phi_n(v) \quad (7)$$

$$\chi_{1,3}(u,v) = \chi_m(u)\chi_m(v) \quad \chi_{2,3}(u,v) = \chi_n(u)\chi_n(v) \quad (8)$$

DTCWT has six oriented complex wavelets at each decomposition whose imaginary component is expressed as:

$$\xi_i(u,v) = \frac{1}{\sqrt{2}}(\chi_{3,i}(u,v) - \chi_{4,i}(u,v)) \quad (9)$$

$$\xi_{i+3}(u,v) = \frac{1}{\sqrt{2}}(\chi_{3,i}(u,v) - \chi_{4,i}(u,v)) \quad (10)$$

Here i = 1, 2, 3 and

$$\chi_{3,1}(u,v) = \phi_n(u)\chi_m(v) \quad \chi_{4,1}(u,v) = \phi_m(u)\chi_n(v) \quad (11)$$

$$\chi_{3,2}(u,v) = \chi_n(u)\phi_m(v) \quad \chi_{4,2}(u,v) = \chi_m(u)\phi_n(v) \quad (12)$$

$$\chi_{3,3}(u,v) = \chi_n(u)\phi_m(v) \quad \chi_{4,3}(u,v) = \chi_m(u)\chi_n(v) \quad (13)$$

In Equations (6) to (13), $\phi_m(*)$ and $\phi_n(*)$ represents the low-pass filters of tree a and tree b respectively in both dimensions, while $\chi_m(*)$ and $\chi_n(*)$ represents the high-pass filters of tree a and tree b , respectively in both dimensions. A 2D DT-CWT creates six sub-bands at each level of decomposition. These sub-bands are accompanied with complex coefficients and are aligned at angles of $\pm 15°$, $\pm 45°$, $\pm 75°$ [16].

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:42:22 UTC from IEEE Xplore. Restrictions apply.

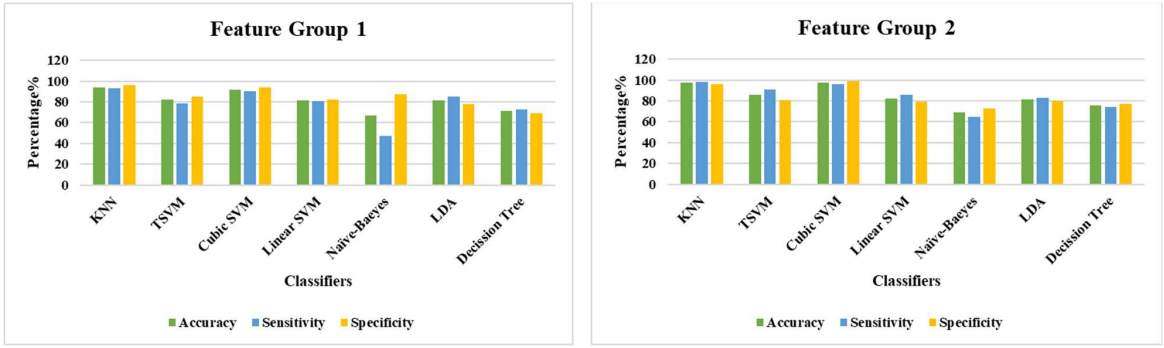


Fig. 3. Bar graph of performance comparison between different classifier for Feature group 1 and for Feature group 2

C. *Methods*

Proposed method includes the use of DTCWT for features generation. A fifth level wavelet decomposition is applied on each MRI image in the dataset and complex wavelet coefficients from all six sub-bands at each level are extracted. DTCWT has six sub-bands at each level of decomposition and the *i*th complex wavelet coefficient is indicated by c*i,d,o* where '*d*' denotes decomposition level and '*o*' stands for orientation. With the help of coefficients generated from all six sub-bands at each level Mean energy can be calculated [25] as

(a) For wavelet coefficients with real components only

$$F_{1,d,o} = \frac{1}{N}\sum_i \left(cr_{i,d,o}\right)^2 \qquad (14)$$

(b) For wavelet coefficients with imaginary components only

$$F_{2,d,o} = \frac{1}{N}\sum_i \left(ci_{i,d,o}\right)^2 \qquad (15)$$

In the above equations $cr_{i,d,o}$ and $ci_{i,d,o}$ denote the real and imaginary components of *i*th wavelet coefficient . *N* indicates the total amount of complex wavelet coefficients at '*d*' decomposition level and '*o*' orientation. Thus, the obtained features for all six sub-bands from each level are then arranged into the given feature group 1and feature group 2.

Feature Group 1. The mean energy obtained from the real components from all sub-bands at each level of decomposition.

Feature Group 1 = $\{F_{1,d,o}\}$, d = 1, 2, . . ., 5, o = 1, 2, . . ., 6. (1). Thus *Set 1* has 30 features.

Feature Group 2. The mean energy obtained from the real components and imaginary components from all sub-bands at each level of decomposition.

Feature Group 2 = $\{F_{1,d,o}, F_{2,d,o}\}$, d = 1, 2, . . ., 5, o = 1, 2, . . ., 6. (2). Thus Set *2* has 60 features.

These features from feature group1 and feature group2 are given to the different classifiers like kNN, Twin Support Vector Machine (TSVM), Cubic SVM, Linear SVM, Naïve-Bayes, Linear Discriminant Analysis (LDA) and Decision Tree for evaluating the performance of proposed approach.

III. Results and Discussion

A. *Performance Evaluation Parameter*

Accuracy [17] is calculated to evaluate the success rate of classifier. Sensitivity [17],[18],[19] indicates the success rate of classifier for detecting true positive, and specificity [17] indicates success rate of classifier for detecting true negative. These are defined in equations from (16) to (18).

$$\text{Accuracy} = (T_P + T_N) / (T_P + T_N + F_P + F_N) \qquad (16)$$

$$\text{Sensitivity} = T_P / (T_P + F_N) \qquad (17)$$

$$\text{Specificity} = T_N / (T_N + F_P) \qquad (18)$$

T_P denotes True positive i.e. person having Alzheimer's disease is rightly classified to be component of AD class. T_N indicates True negative i.e. person not having Alzheimer's disease is rightly classified to belong to NC class. F_P means False Positive i.e. person not having Alzheimer's Disease is wrongly classified to be component of AD class while F_N indicates False Negative i.e. person having Alzheimer's Disease is wrongly classified to belong to NC class.

B. *Classification Performance*

Fig. 3 compares the performance of different classifiers on two feature groups. Feature group 1 consists of the mean energy of the real component of complex wavelet coefficient and Feature group 2 consists of the mean energy of real and imaginary component of the complex wavelet coefficient taken separately. Performance is evaluated in terms of accuracy, sensitivity, and specificity with KNN, TSVM, Cubic SVM, Linear SVM, Naïve-Bayes, LDA and Decision Tree classifiers. The result of all classifiers with Feature group 2 is better than the result obtained with Feature group 1. KNN and Cubic SVM classifiers are giving good accuracy of 97.3 %, sensitivity of 98.2 %and 96% respectively, and specificity of 96.4% and 99% respectively with Feature group 2. Table II indicates the performance of different classifier

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:42:22 UTC from IEEE Xplore. Restrictions apply.

TABLE II: PERFORMANCE COMPARISON BETWEEN TWO FEATURE GROUP WITH DIFFERENT CLASSIFIER

Method	Feature Group	Features	Classifier	Accuracy [17]	Sensitivity [17] [18] [19]	Specificity [17][18]
DTCWT	Feature Group 1	Mean energy of Real component only	KNN	94.2	93	96
			TSVM	82.08	78.8	85.3
			Cubic SVM	92	90	94
			Linear SVM	81.7	81	82
			Naïve-Bayes	67	47	87
			LDA	81.3	85	78
			Decision Tree	71	73	69
	Feature Group 2	Mean energy of Real and imaginary component taken separately	KNN	97.3	98.2	96.4
			TSVM	85.9	91.2	80.73
			Cubic SVM	97.3	96	99
			Linear SVM	82.1	86	79
			Naïve-Bayes	69.2	65	73
			LDA	81.7	83	80
			Decision Tree	75.4	74	77

TABLE III: PERFORMANCE COMPARISON OF PROPOSED METHOD WITH EXISTING ALGORITHM

Method	Features	Classifier	DATASET	Accuracy [17]	Sensitivity [17][18][19]	Specificity [17] [18]
DTCWT [Proposed]	Mean energy of real component and imaginary component	KNN/Cubic SVM	OASIS	97.3/97.3	98.2/96	96.4/99
DTCWT+PCA [11]	detail coefficients at 5th level of decomposition	FNN	OASIS	90.06 ± 0.01	92.00 ± 0.04	87.78 ± 0.04
DTCWT+PCA+LDA [12]	detail coefficients at 5th level of decomposition	TSVM	OASIS	96.68 ± 1.44	97.72 ± 2.34	95.61 ± 1.67
DTCWT+PCA+LDA [13]	detail coefficients at 5th level of decomposition	ELM	OASIS	95.72±1.54	96.59±2.34	93.03±1.67
DTCWT+RCWF [15]	Energy, entropy, standard deviation	FNN	OASIS	91.72	92.01	88.00

with Feature group 1 and Feature group 2. Proposed method is tested on OASIS dataset and outperformed the existing methods [11], [12], [13], [15] shown in Table III. Existing methods [11], [12],[13] used the detail coefficients at fifth level as features and used PCA [11], [12], [13] and LDA [12],[13] for feature dimension reduction. They could achieve classification accuracy of 90.06 ± 0.01[11] , 96.68 ± 1.44 [12], 95.72±1.54[13] and 91.72[15],sensitivity of 92.00 ± 0.04[11] , 97.72 ± 2.34 [12] , 96.59±2.34[13]and 92.01[15] and specificity of 87.78 ± 0.04[11] , 95.61 ± 1.67[12], 93.03±1.67[13] and 88.00[15] with FNN[11] , TSVM[12] ,ELM [13] and FNN[15] classifier respectively. And the proposed method obtained the classification accuracy of 97.3% with both KNN and cubic SVM, sensitivity of 98.2% with KNN and 96% with cubic SVM, and specificity of 96.4% with KNN and 99% with cubic SVM. Hence our proposed method is simple. There is no need of reducing feature dimension further and thus no more complexity is added in the approach and results are coming more promising in comparison to the existing methods using DTCWT.

IV. CONCLUSIONS

In this paper, a simple technique for feature generation using DTCWT based complex wavelet coefficients is proposed for AD and NC classification. Proposed method signifies that AD can be classified effectively by taking the mean energy of complex wavelet coefficient with real and

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:42:22 UTC from IEEE Xplore. Restrictions apply.

imaginary components taken individually as features. In addition to this, our proposed method on OASIS dataset works superior in comparison to the other methods stated in this paper—providing 97.3% accuracy along with equally high sensitivity and specificity. In future we will focus on (i) testing with other wavelets also like 3DDTCWT on 3D MRI images, (ii)testing with new features also and (iii) will introduce a feature selection stage for further improvement in accuracy.

ACKNOWLEDGMENT

This research was supported by the Department of Electronics and Communication Engineering, Delhi Technological University, Delhi.

REFERENCES

[1] Brookmeyer R, Johnson E, Ziegler-Graham K, and Arrighi HM "Forecasting the global burden of Alzheimer's disease." Alzheimer's & dementia: the journal of the Alzheimer's Association vol. 3,3 (2007): pp. 186-91. doi: 10.1016/j.jalz.2007.04.381

[2] C. A. Raji, O. L. Lopez, L. H. Kuller, O. Carmichael, J. T. Becker "Age, Alzheimer disease, and brain structure" Neurology Dec 2009, 73 (22) pp.1899-1905; doi: 10.1212/wnl.0b013e3181c3f293

[3] Chupin, Marie et al. "Fully automatic hippocampus segmentation and classification in Alzheimer's disease and mild cognitive impairment applied on data from ADNI." Hippocampus vol. 19,6 (2009) pp. 579-87. doi:10.1002/hipo.20626

[4] Jack, Clifford R Jr et al. "Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade." The Lancet. Neurology vol. 9,1 (2010): pp. 119-28. doi:10.1016/S1474-4422(09)70299-6

[5] Glodzik-Sobanska, Lidia et al. "The role of quantitative structural imaging in the early diagnosis of Alzheimer's disease." Neuroimaging clinics of North America vol. 15,4 (2005): pp.803-26, x. doi:10.1016/j.nic.2005.09.004

[6] C. H. Moritz, V. M. Haughton, D. Cordes, M. Quigley, and M. E. Meyerand, "Whole-brain functional MR imaging activation from finger tapping task examined with independent component analysis," American Journal of Neuroradiology, Vol. 21, no. 9, pp. 1629–1635, 2000

[7] R. N. Bracewell, The Fourier Transform and Its Applications, McGraw-Hill, New York, Third edition, 1999.

[8] T. H. El-Dahshan and A. B. M. Salem, "Hybrid intelligent techniques for MRI brain images classification," Digital Signal Processing, vol. 20, no. 2, pp. 433–441, 2010

[9] S. Chaplot, L. M. Patnaik, and N. R. Jagannathan, "Classification of magnetic resonance brain images using wavelets as input to support vector machine and neural network," Biomedical Signal Processing and Control, vol. 1, no. 1, pp. 86–92, 2006.

[10] Y. Zhang, Z. Dong, L. Wu, and S. Wang, "A hybrid method for MRI brain image classification," Expert Systems with Applications, vol. 38, no. 8, pp. 10049–10053, 2011.

[11] Jha, J.-I. Kim, and G.-R. Kwon, "Diagnosis of alzheimers disease using dual-tree complex wavelet transform, pca, and feed-forward neural network," Journal of Healthcare Engineering, vol. 2017, (2017) https://doi.org/10.1155/2017/9060124.

[12] Alam, Saruar et al. "Twin SVM-Based Classification of Alzheimer's Disease Using Complex Dual-Tree Wavelet Principal Coefficients and LDA." Journal of healthcare engineering vol. 2017 (2017): 8750506. doi:10.1155/2017/8750506

[13] Jha, Debesh et al. "Alzheimer's Disease Detection Using Extreme Learning Machine, Complex Dual Tree Wavelet Principal Coefficients and Linear Discriminant Analysis." *Journal of Medical Imaging and Health Informatics* 8 (2018): 881-890.

[14] Bhasin, H., Agrawal, R.K. & For Alzheimer's Disease Neuroimaging Initiative. "A combination of 3-D discrete wavelet transforms and 3-D local binary pattern for classification of mild cognitive impairment." *BMC Med Inform Decis Mak* **20,** 37 (2020). https://doi.org/10.1186/s12911-020-1055-x

[15] Preeti S. Topannavar et al. "Rotational Wavelet Filters for Analysis of Brain MRI in Detection of Alzheimer's Disease" Vol. 12 No. 6 (2021), https://doi.org/10.17762/turcomat.v12i6.5789.

[16] Selesnick, I.W.—Baraniuk, R.G.: The Dual-Tree Complex Wavelet Transform. IEEE Signal Processing Magazine, Vol. 22, 2005, No. 6, pp. 123–151.

[17] Trevethan R. Sensitivity, Specificity, and Predictive Values: Foundations, Pliabilities, and Pitfalls in Research and Practice. Front Public Health. 2017 Nov 20;5:307. doi: 10.3389/fpubh.2017.00307.

[18] Abdul Ghaaliq Lalkhen, MB ChB FRCA, Anthony McCluskey, BSc MB ChB FRCA, Clinical tests: sensitivity and specificity, Continuing Education in Anaesthesia Critical Care & Pain, Volume 8, Issue 6, December 2008, Pages 221–223, https://doi.org/10.1093/bjaceaccp/mkn041

[19] Parikh R,Mathai A,Parikh S,Chandra Sekhar G,Thomas R, Understanding and using sensitivity, specificity and predictive values. Indian journal of ophthalmology. 2008 Jan-Feb

[20] http://www.oasisbrain.org

[21] Kukull, W A et al. "The Mini-Mental State Examination score and the clinical diagnosis of dementia." Journal of clinical epidemiology vol. 47,9 (1994): pp.1061-7. doi:10.1016/0895-4356(94)90122-8

[22] Morris, J C. "The Clinical Dementia Rating (CDR): current version and scoring rules." Neurology vol. 43,11 (1993): pp. 2412-4. doi:10.1212/wnl.43.11.2412-a

[23] N. Kingsbury, "Design of Q-shift complex wavelets for image processing using frequency domain energy minimization," in Proceedings of the International Conference on Image Processing (ICIP '03), vol. 1, pp. 1013–1016, Barcelona, Spain, September 2003.

[24] N. Kingsbury, "Complex wavelets for shift invariant analysis and filtering of signals," Applied and Computational Harmonic
Analysis, vol. 10, no. 3, pp. 234–353, 2001.

[25] Vermaak, H.; Nsengiyumva, P.; Luwes, N. "Using the Dual-Tree Complex Wavelet Transform for Improved Fabric Defect Detection." J. Sens. 2016, 2016, pp.1–8

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:42:22 UTC from IEEE Xplore. Restrictions apply.

Proceedings of the SMART–2021, IEEE Conference ID: 52563
10th International Conference on System Modeling & Advancement in Research Trends, 10th–11th December, 2021
Faculty of Engineering & Computing Sciences, Teerthanker Mahaveer University, Moradabad, India

2021 10th International Conference on System Modeling & Advancement in Research Trends (SMART) | 978-1-6654-3970-1/21/$31.00 ©2021 IEEE | DOI: 10.1109/SMART52563.2021.9676248

Implementation of LSTM for Prediction of Diabetes using CGM

Sunny Arora[1], Shailender Kumar[2], Pardeep Kumar[3]
[1,2]*Computer Science dept. of Delhi Technological University, Delhi, India*
[3]*Computer Science dept. of Swansea University, Swansea, United Kingdom*
E-mail: [1]Sunnyarora_2k18phdco@dtu.ac.in, [2]shailenderkumar@dce.ac.in, [3]pardeep.kumar@swansea.ac.uk

***Abstract*—Deep learning has added conveniences for the diagnosis and prediction of various diseases making an influence in healthcare facilities. Diabetes mellitus is a dominant health issue faced by many around the globe. The number of people with this disease went up from one hundred eight million to six hundred in 1980, to four sixty million in 2019. Predicting trends of blood glucose prediction using deep learning methods make the management of the disease much easier. In this work, we are predicting future trends of the disease using training data. We have used the publically available dataset Ohio T1DM dataset in this work. In this paper, we have implemented LSTM to predict future trends. Root mean square error is used as the performance evaluation measure for this work.**

Keywords: Diabetes, Type 1, Regression Models, Linear Regression, Support Vector Regression, LSTM.

I. Introduction

Diabetes mellitus is a public health challenge faced commonly by the general population of these recent times. Universally, the figure of grown-ups living with the infection has ascended from 108 million in 1980 to 422 million in 2014 [1]. In India, this number has already touched 77 million. Diabetes may lead to several complications such as vision loss, terminal renal disease, and amputations it can also lead to macrovascular/ microvascular disease [2]. The pancreas embraces the crucial accountability to uphold blood glucose levels amid around 70 mg/dL and 180 mg/ dL [3] by matching the levels of insulin and glucagon in the bloodstream. Type 1 diabetes is a more chronic disease, men in general fall trap to this disease more often than women [4]. In Nordic countries like Finland, Germany, and Norway the average twelve-monthly upsurge of 2·76% in its occurrence has been reported [5-7]; while it is still most prevalent in Finland and Sardinia

Adults have the highest prevalence of this disease in the Middle East as well as in North Africa. It occurs in different people can be attributed to their diminished pancreatic function and/or problematic insulin sensitivity. They have to overcome the challenge of preserving their blood glucose levels within the reference range via proper medicines, lifestyle changes, and exercise [8-10]. Figure 1 depicts the different types of Diabetes. Type 1 diabetes mellitus is a category of diabetes under which 5%-10% of subjects are diagnosed [11] and it is owing to the devastation of β cells of the pancreas [12,13]. Children and adolescents are the most affected population type under this category [14]. While Idiopathic type 1 diabetes is an infrequent form of type 1 diabetes and is distinctively found in people of African/ Asian descent [15]. Fulminant type 1 diabetes is a distinctive form of type 1 diabetes, [16,17]. It is categorized by ketoacidosis soon after the onset of hyperglycemia [18,19].

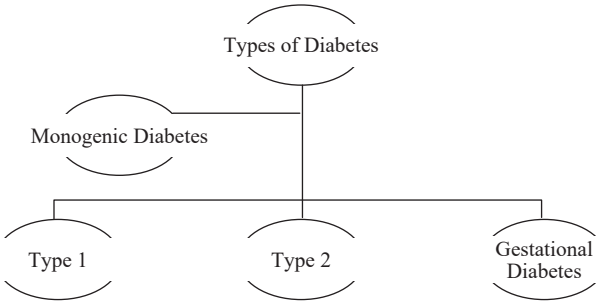


Fig.1: Types of Diabetes [20]

The proven features of this subtype of type 1 diabetes include strangely unexpected commencement of illness, a diminutive interval of indications, which may include events like bodyweight loss, polyuria, need for liquids, and commonly acidosis. According to a study, the prevalence of diabetes is going to double by 2030 [22]. The problem with this disease is the affected person may not have symptoms for a prolonged time duration which makes it more dangerous. Type 2 diabetes is more commonly found in adults and children with obesity issues. It is caused due to the deficiency in the mechanism the body controls and customs sugar as energy. In type 2 diabetes, there are principally two interconnected difficulties one of which is that the pancreas is not able to produce sufficient insulin which is the primary hormone that controls the movement of sugar through the cells and another being the response of the cells become poor towards insulin.

Over the last two decades, deep learning has supplemented the diagnosis and prediction of many diseases. Prediction of CGM values in the case of diabetes

Copyright © IEEE–2021 ISBN: 978-1-6654-3970-1

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 07:35:23 UTC from IEEE Xplore. Restrictions apply.

will aid the management of the disease for the patient. In this work, we have used LSTM (Long Short Term Memory) to predict blood glucose levels from CGM history. Publically available database Ohio T1DM dataset was incorporated in this study. The paper is divided into the following sections: Section 2 discusses the methodology followed during the work. Section 3 details the results and Section 4 holds the discussion section.

II. Methodology

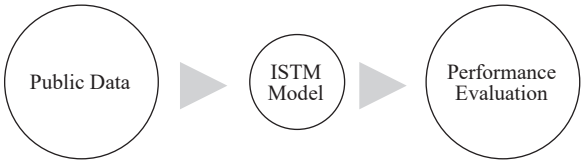


Fig. 2: Flowchart of Work

A. Public Data

The OhioT1DM Dataset is used in this work [21]. The data comprises two months of data for each of twelve persons, who have type 1 diabetes. The people in the database are denoted by arbitrarily designated ID figures. Medtronic 530G or 630G insulin pumps were worn by these papers and they were on insulin pump therapy. Sensors used through this period were Medtronic Enlite CGM sensors. They were designed to provide information life-event data via a custom smartphone application. The data was for two months, consisting of eight months, and collected at an interval of five minutes. The age range of patients was 40-60 years among which four females and two male patients. Figure 3 and Figure 4 depict the number of data points in each patient's dataset for both training and testing data.

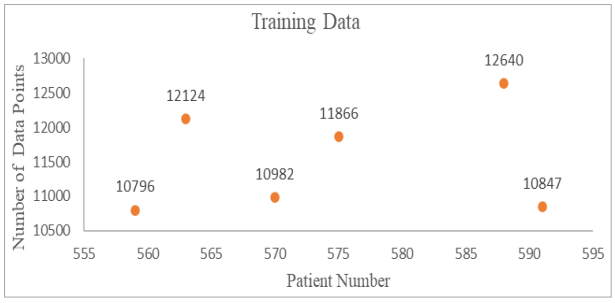


Fig. 3: Training Data Details

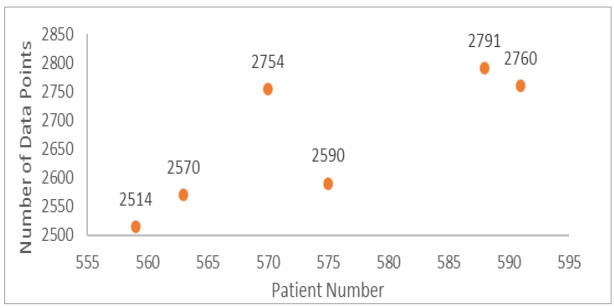


Fig. 4: Testing Data Points

B. Model: LSTM

It is an exceptional sort of network that is equipped for learning long-term dependencies in the given information. This is accomplished because the repetitive module of the model has a blend of four layers associating with one another. To avoid the gradients from the explosion, in this work we have established the gradient threshold to one. The initial learning rate used in this model is 0.005, and after a certain period, which for this work is 125 epochs, we update the learning weight by multiplying it by 0.2. Table 1 holds the details of hyperparameters used in the model.



Fig. 5: LSTM Model Layers

Table 1: Parameters used in the Model

Parameters	
Optimizer	**Adam**
Maximum no of Epocs	250
Number of Hidden Units	200
Initial Learning Rate	.005
Learn Drop Rate	125

C. Performance Evaluation Methods

In this work, the performance measure root mean square error (RMSE) is used to evaluate the model used. Equation 1 depicts the formulation of the RMSE.

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(Z_t - \hat{Z}_t)^2}, \quad (1)$$

III. Results

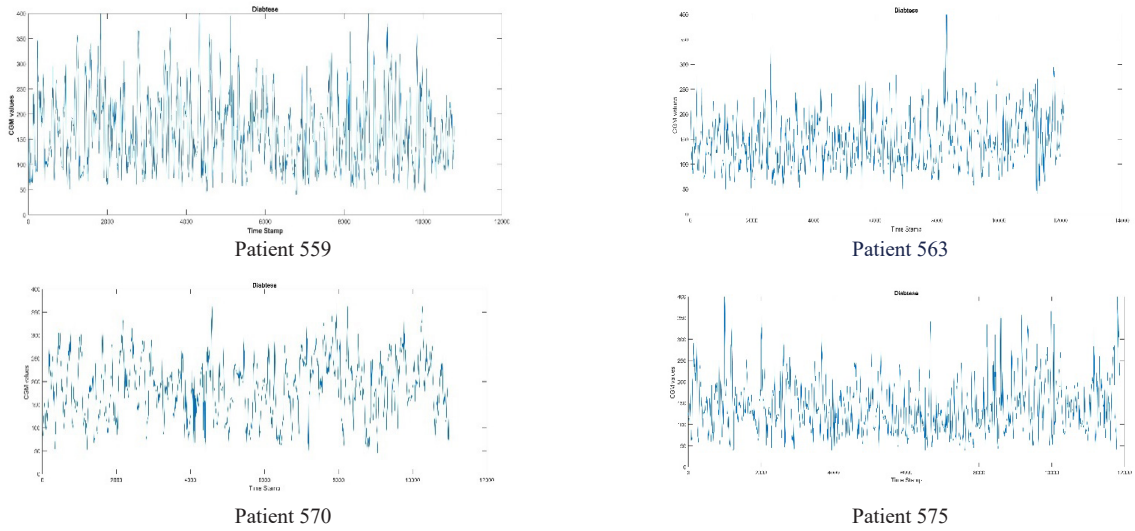
Figure 6 is a graphical representation of the training dataset for all six patients and Figure 7 shows the Root mean square error concerning the test set. Figure 6 depicts the CGM values of six different patients over eight weeks, the x-axis depicts the data received, over the gap of five minutes and the y-axis depicts the CGM value. Figure 8 depicts the RMSE achieved using LSTM over testing data and predicted values. Table 1 depicts the RMSE values achieved after the application of the model. From figure 7 it is evident that the model performed with great efficiency, on the test as well as the training data. For Patient number 565 it provided a minimum RMSE 3.9, followed by patient no. 570 and patient no. 559. It gave the average RMSE of 4.02.

Table 2: Results

Patient No.	RMSE
559	4.565
563	3.9414
570	4.2829
575	5.1865
588	4.8565
591	5.6247

Copyright © IEEE–2021 ISBN: 978-1-6654-3970-1

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 07:35:23 UTC from IEEE Xplore. Restrictions apply.



Patient 559 Patient 563

Patient 570 Patient 575

Fig. 6: CGM Value of All Patients

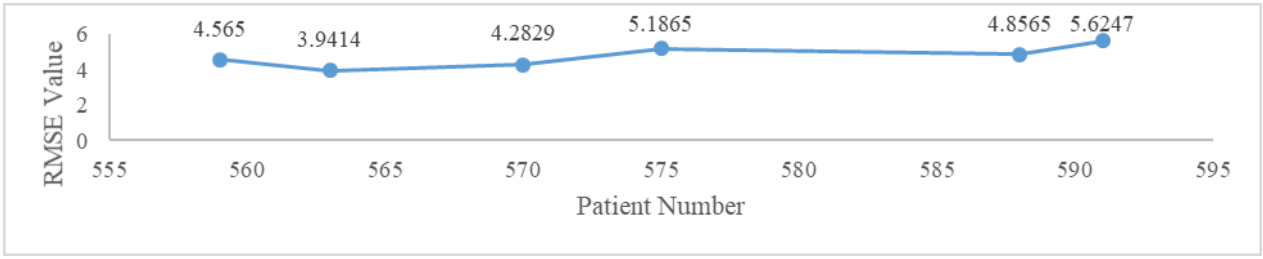
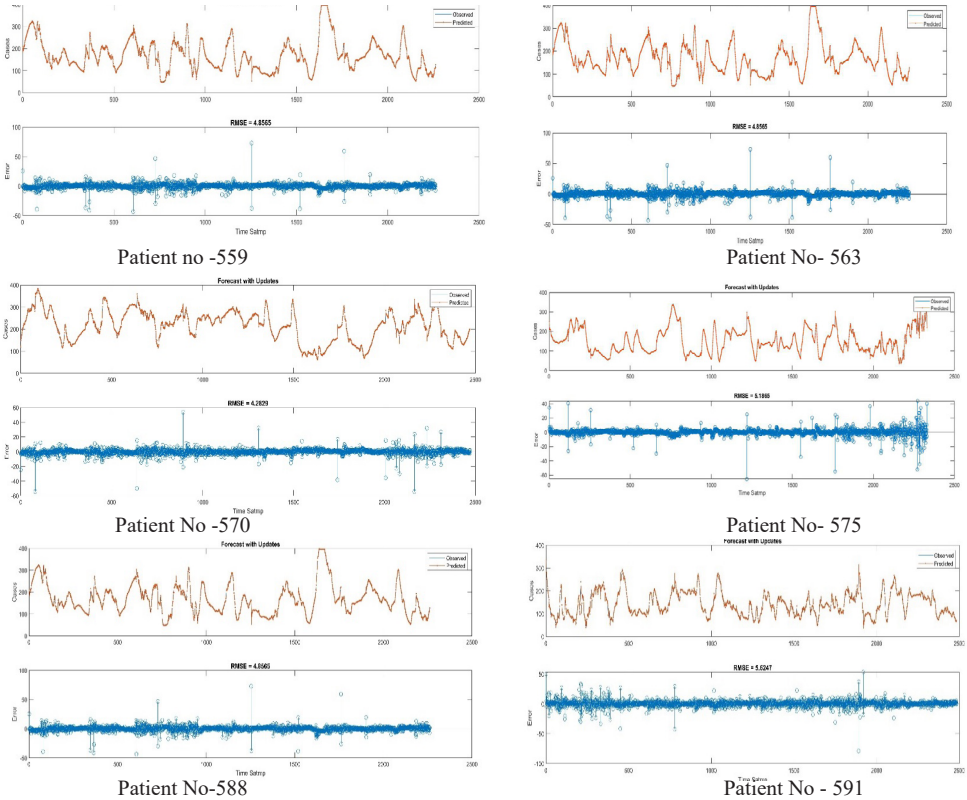


Fig. 7: RMSE Values for Different Patient Cases



Patient no -559 Patient No- 563

Patient No -570 Patient No- 575

Patient No-588 Patient No - 591

Fig. 8: Forecast Values of Patients

Copyright © IEEE–2021 ISBN: 978-1-6654-3970-1

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 07:35:23 UTC from IEEE Xplore. Restrictions apply.

IV. Discussion

In the past years, many researchers have presented their models for the public database. Table 2 has elaborated those, works for comparison with the current study. The work on glucose level prediction can be found three decades earlier by Bremer, *et al*. [32] (1992), where he used autocorrelation functions. Different models such as the support vector machine and the feed-forward neural network have been used in [27, 33-35]. Deep learning approaches are also popular among authors, different authors have used models such as RNN in [36,37], LSTM [24,25], and CNN in [38]. Evaluating table 3, we can see that our model has improved the RMSE value, and performed better than the state of art models.

Table 3: Comparison with Recent Research Work

Author	Reference	Year	Approach	Result (Average RMSE)
Bertachi, *et al.*	[26]	2018	feed-forward NN	19.04
Xie, *et al.*	[27]	2018	SVM	19.32
Xie, *et al.*	[28]	2018	ARX linear regression	19.53
Martinsson, *et al*.	[29]	2018	LSTM RNN	20.06
Midroni, *et al.*	[30]	2018	XGBoost	20.37
Contreras *et al.*	[31]	2018	Gramm. evolution	21.18
Zhu, *et al.* 2018	[32]	2018	Wave Net convolutional NN	21.72
Martinsson, *et al.*	[24]	2019	LSTM	18.67
Proposed		**2021**	**LSTM**	**4.02**

V. Conclusions

This work used the LSTM network for the prediction of CGM levels in Type 1 diabetes patients. As discussed in the results, we have compared our performance evaluation parameter RMSE with those obtained by different works and found that we have achieved better results from state-of-the-art methods available.

VI. Acknowledgment

The author is thankful for the public database provided by C. Marling and R. Bunescu[21].

VII. Acronyms

RMSE: Root mean square error

References

[1] World Health Organization. Global Report on Diabetes. Geneva, Switzerland: World Health Organization; 2016.
[2] Klein R. Hyperglycemia and microvascular and macrovascular disease in diabetes. Diabetes Care 1995 Feb;18(2):258-268.
[3] Leahy J, Clark N, Cefalu W. Medical Management of Diabetes Mellitus. New York, USA: Dekker; 2000.
[4] Ostman J, Lonnberg G, Arnqvist HJ, *et al*. Gender differences and temporal variation in the incidence of type 1 diabetes: results of 8012 cases in the nationwide Diabetes Incidence Study in Sweden 1983–2002. J Intern Med. 2008;263:386–94.
[5] Patterson CC, Dahlquist GG, Gyürüs E, Green A, Soltész G, EURODIAB Study Group Incidence trends for childhood type 1 diabetes in Europe during 1989–2003 and predicted new cases 2005–20: a multicentre prospective registration study. Lancet. 2009;373:2027–33.
[6] Thunander M, Petersson C, Jonzon K, *et al*. Incidence of type 1 and type 2 diabetes in adults and children in Kronoberg, Sweden. Diabetes Res ClinPract. 2008;82:247–55.
[7] Ehehalt S, Dietz K, Willasch AM, Neu A. Prediction model for the incidence and prevalence of type 1 diabetes in childhood and adolescence: evidence for a cohort-dependent increase within the next two decades in Germany. Pediatr Diabetes. 2012;13:15–20.
[8] American Diabetes Association. Diagnosis and classification of diabetes mellitus. Diabetes Care. 2012;35(suppl 1):64–71.
[9] Rodbard D. Continuous glucose monitoring: a review of successes, challenges, and opportunities. Diabetes TechnolTher 2016 Feb;18(Suppl 2):S3-13
[10] Onkamo P, Vaananen S, Karvonen M, Tuomilehto J. Worldwide increase in incidence of type I diabetes—the analysis of the data on published incidence trends. Diabetologia 1999; 42: 1395–403.
[11] Maahs DM, West NA, Lawrence JM, Mayer-Davis EJ. Epidemiology of type 1 diabetes. EndocrinolMetabClin North Am. 2010;39:481–497.
[12] Daneman D. Type 1 diabetes. Lancet. 2006;367:847–858.
[13] Devendra D, Liu E, Eisenbarth GS. Type 1 diabetes: recent developments. BMJ. 2004;328:750–754.
[14] Dabelea D, Mayer-Davis EJ, Saydah S, Imperatore G, Linder B, Divers J, Bell R, Badaru A, Talton JW, Crume T, *et al*. Prevalence of type 1 and type 2 diabetes among children and adolescents from 2001 to 2009. JAMA. 2014;311:1778–1786.
[15] Abiru N, Kawasaki E, Eguch K. Current knowledge of Japanese type 1 diabetic syndrome. Diabetes Metab Res Rev. 2002;18:357–366.
[16] Imagawa A, Hanafusa T, Miyagawa J, Matsuzawa Y. A proposal of three distinct subtypes of type 1 diabetes mellitus based on clinical and pathological evidence. Ann Med. 2000;32:539–543.
[17] 47. Imagawa A, Hanafusa T, Miyagawa J, Matsuzawa Y. A novel subtype of type 1 diabetes mellitus characterized by a rapid onset and an absence of diabetes-related antibodies. Osaka IDDM Study Group. N Engl J Med. 2000;342:301–307.
[18] Imagawa A, Hanafusa T. Fulminant type 1 diabetes--an important subtype in East Asia. Diabetes Metab Res Rev. 2011;27:959–964.
[19] Shibasaki S, Imagawa A, Hanafusa T. Fulminant type 1 diabetes mellitus: a new class of type 1 diabetes. AdvExp Med Biol. 2012;771:20–23.
[20] Centers for Disease Control and Prevention. National diabetes statistics report, 2017. Centers for Disease Control and Prevention website. www.cdc.gov/diabetes/pdfs/data/statistics/national-diabetes-statistics-report.pdf External link (PDF, 1.3 MB). Updated July, 18 2017. Accessed August 1, 2017.
[21] C. Marling and R. Bunescu. The OhioT1DM dataset for blood glucose level prediction: Update 2020. In Proceedings of the 5th International Workshop on Knowledge Discovery in Healthcare Data, Santiago de Compostela, Spain, August 30, 2020, http://ceur-ws.org/Vol-2675/paper11.pdf
[22] TANZOLA, M. (2010) 'Type 2 Diabetes Screening Found Cost Effective', Internal Medicine News, 43(8), p. 21. doi: 10.1016/s1097-8690(10)70423-6.
[23] Imagawa, A. *et al*. (2003) 'Fulminant type 1 diabetes: A nationwide survey in Japan', Diabetes Care, 26(8), pp. 2345–2352. doi: 10.2337/diacare.26.8.2345.
[24] Martinsson J, Schliep A, Eliasson B *et al*. Blood glucose prediction with variance estimation using recurrent neural networks. J Healthc Inform Res 2019, p. 1-18. DOI: 10.1007/s41666-019- 00059-y. 34. Aiello E
[25] Chen J, Li K, Herrero P *et al*. Dilated recurrent neural network for short-time prediction of glucose concentration. Proc 3rd Int Workshop on Knowl Discov in Healthcare Data 2018, p. 69-73. URL: http://ceur-ws.org/Vol-2148/paper11.pdf [accessed 2020-03-09]

Copyright © IEEE–2021 ISBN: 978-1-6654-3970-1

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 07:35:23 UTC from IEEE Xplore. Restrictions apply.

[26] Bertachi A, Biagi L, Contreras I *et al*. Prediction of blood glucose levels and nocturnal hypoglycemia using physiological models and artificial neural networks. Proc 3rd Int Workshop on Knowl Discov in Healthcare Data 2018, p. 85-90. URL: http://ceur-ws.org/Vol-2148/paper14.pdf [accessed 2020-03-09].

[27] Xie J, Wang Q. Benchmark machine learning approaches with classical time series approaches on the blood glucose level prediction challenge. Proc 3rd Int Workshop on Knowl Discov in Healthcare Data 2018, p. 97-102. URL: http://ceur-ws.org/Vol-2148/paper16.pdf [accessed 2020-03-09].

[28] Martinsson J, Schliep A, Eliasson B *et al*. Automatic blood glucose prediction with confidence using recurrent neural networks. Proc 3rd Int Workshop on Knowl Discov in Healthcare Data 2018, p. 64-68. URL: http://ceur-ws.org/Vol-2148/paper10.pdf [accessed 2020-03-09].

[29] . Midroni C, Leimbigler PJ, Baruah G *et al*. Predicting glycemia in type 1 diabetes patients: experiments with XGBoost. Proc 3rd Int Workshop on Knowl Discov in Healthcare Data 2018, p. 79-84. URL: http://ceur-ws.org/Vol-2148/paper13.pdf [accessed 2020-03-09].

[30] Contreras I, Bertachi A, Biagi L *et al*. Using grammatical evolution to generate short-term blood glucose prediction models. Proc 3rd Int Workshop on Knowl Discov in Healthcare Data 2018, p. 91- 96. URL: http://ceur-ws.org/Vol-2148/paper15.pdf [accessed 2020-03-09].

[31] Zhu T, li K, Herrero P *et al*. A deep learning algorithm for personalized blood glucose prediction. Proc 3rd Int Workshop on Knowl Discov in Healthcare Data 2018, p. 74-78. URL: http://ceurws.org/Vol-2148/paper12.pdf [accessed 2020-03-09].

[32] Bremer T, Gough DA (1999) Is blood glucose predictable from previous values? A solicitation for data. Diabetes 48(3):445–451

[33] Wiley MT (2011) Machine learning for diabetes decision support, pp 55–72

[34] Perez-Gand ´ ´ıa C, Facchinetti A, Sparacino G, Cobelli C, Gomez E, Rigla M, de Leiva A, Hernando ´ M (2010) Artificial neural network algorithm for online glucose prediction from continuous glucose monitoring. Diab Technol Therapeut 12(1):81–88

[35] Pappada SM, Cameron BD, Rosman PM, Bourey RE, Papadimos TJ, Olorunto W, Borst MJ (2011) Neural network-based real-time prediction of glucose in patients with insulin-dependent diabetes. Diab Technol Therapeut 13(2):135–141

[36] Mougiakakou SG, Prountzou A, Iliopoulou D, Nikita KS, Vazeou A, Bartsocas CS (2006) Neural network based glucose-insulin metabolism models for children with type 1 diabetes. In: Engineering in medicine and biology society, 2006. EMBS'06. 28th annual international conference of the IEEE. IEEE, pp 3545–3548

[37] Chen J, Li K, Herrero P, Zhu T, Georgiou P (2018) Dilated recurrent neural network for short-time prediction of glucose concentration. In: 3rd International workshop on knowledge discovery in healthcare data, KDH@ ICML/IJCAI 2018, 13 July 2018, pp 69–73

[38] Zhu T, Li K, Herrero P, Chen J, Georgiou P (2018) A deep learning algorithm for personalized blood glucose prediction. In: 3rd International workshop on knowledge discovery in healthcare data, KDH@ ICML/IJCAI 2018, 13 July 2018, pp 64–78

Copyright © IEEE–2021 ISBN: 978-1-6654-3970-1

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 07:35:23 UTC from IEEE Xplore. Restrictions apply.

Magnification Dependent Histopathology Breast Cancer Image Binary Classification Using Different Levels Of EfficientNet

Saman Saadi Zadeh
Electronics and Communication Department
Delhi Technological University
Delhi, India
Saman.Saadizadeh@gmail.com

2021 6th International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS) | 978-1-7281-6714-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICIIBMS52876.2021.9651644

***Abstract*—Breast cancer kills millions of women annually around the world; early diagnosis of this perilous disease can result in efficient treatment. Computer Aided Detection systems by using advanced screening tools can automatically help the physician all over the world to detect the cancer. In our approach we are dealing with histopathological Images and state-of-art deep learning network named EfficientNet, it has been applied to BreakHis images to classify them into benign or malignant. For breast cancer classification in histopathological images using deep learning we applied eight different stages of EfficientNet that has different amounts of parameters on four diverse resolutions from our image dataset. This pre-trained Network with ImageNet weights represents high accuracy in training as well as in testing part. The best result among 32 findings in our proposed work gives the accuracy of 94.41% for training, 99.28% while predicting and 99% Sensitivity, Recall and F1-score.**

Keywords—EfficientNet, BreakHis,Cancer Classification, Histhopathological Images, CAD system

I. INTRODUCTION

Cancer kills over millions of patients every year around the world. Among variant types of cancer, breast cancer for women and lung cancer for men are the most common disease for patients. Findings in American Cancer Society, depicts a decrease in mortality rate among breast cancer patients, which seemingly is the results of early-stage cancer diagnosis possibility by using brand new tools of screening, more women self-consciousness and enhanced cure methods which make the treatments more successful [1], [2]. Tumors also are the cells that abnormally grown. There are two main types of tumors: Benign and Malignant. In Benign tumors tissues mostly wont treated as cancer as the cells have undergone a negligible change and they are not invasive. Malignant tumors in the other hand have the ability to grow and release in other organs. Histopathologist visually inspect the histology images to check the discipline of cell morphology and tissue spreading, accordingly, they determinate the cancerous tissue and decide about the level of malignancy [3].The histopathological images have a significant point that they are taken while an open biopsy surgery therefor they can give accurate information about the condition of breast tissue. Cancer CAD or Cancer Computer-Aided Diagnosis is an advanced tool based on computer to diagnose the cancer and help the efficiency of treatment process by detecting cancer accurately and reducing the death rates. CAD basically searches in the medical images and classify them into categories and stages. In recent advancement Artificial Intelligence (AI) has shown promising and great results to detect cancers. Computer-Aided Diagnosis (CAD) utilizes novel machine learning and other AI methods like Deep Learning (DL) can be a great help in the medical field, apparently it avoids many drawbacks of traditional methods of cancer detection such as human errors that may occur.

The aim of this study is to develop a CAD system for breast cancer classification that is accurate and authentic. To achieve this goal we applied a state-of-art deep learning model for the automate process.

The remain part of this paper consists of: In section II, related work done on histopathological images for the breast cancer classification using CAD systems have been described. In section III, we concentrate on our proposed method for histopathological image classification. In section IV, the results of the proposed methods on histopathological images using EfficientNet are presented, compared and pointed out the best results. In section V, conclusion and the forthcoming studies have written briefly.

II. LITRATURE REVIEWS

Cancer Computer-Aided detection is developing rapidly by emerging new advancements of machine learning and deep learning networks. CAD can ease the burden of recognizing disease and also it can reduce the false negative and false positives that may occur by the naked eyes of physicians. Although the results of new techniques have gained high accuracies but due to the importance of health issues and the patients lives there should be many more researches to find the best alternate for the previous networks [4]. Before the emergence of Deep Neural Networks, the breast images were classified with classical approaches, they already had many significant results but there were some drawbacks like

978-1-7281-6714-5/21/$31.00 ©2021 IEEE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 09:32:44 UTC from IEEE Xplore. Restrictions apply.

overfitting, the computational time process, low training speed, and also the parameters were set only manually by the experiences [5].

In computer vision (CV) and other applications like natural Language processing (NLP), deep learning obtains excellent performance. CNN or Convolutional Neural Network basically is an algorithm that performs efficiently to extract important features and learn from them in the image. Nowadays the CNN are modifying rapidly to gain a better performance, in [6] Sohail A., et al. proposes a new architecture based on deep CNN that is multiphase diagnosis framework and detect the mitotic nuclei in breast cancer by the means of histopathological images. In their work a label filter is used to refine the weak labels that has semi-semantic data and use them in Convolutional Neural Network training process. This deep model is based on instance segmentation and detection; also in their work they detect possibly weak mitotic regions on tissue fragments. Then the more likely areas are monitored based on the spot parts of cells, therefore it is classified by developing a Convolutional Neural Network named "MitosRes-CNN". The developed architecture for this method got the F-1 score (0.75) value and precision of (0.71), AUC or Area Under Curve (0.78) and recall (0.76) on TUPAC16 dataset. In "MitosRes" results from previous layer comes as output, and then fed as an input for the next layer, also there are pooling and fully connected layers which considers as the main layers. Generally in deep learning more data gives the better results but for the era of information, training a huge amount of data from scratch is a remarkable issue due to the lack of time and high expenses. In [7] Mahesh G. et al. uses BreakHis dataset for their proposed work and design an architecture named "ResHis" which is based on Residual CNN and includes 152 Layers to extract the features and classify the images into binary classification. Authors in their proposed work find the better result after data augmentation method, by stain normalization, generating image fragments and affine transformation. Finally the results achieve the accuracy of 92.52% and F1- score (93.45). In [8] the Yao, H. et. al. , uses BACH for their proposed work, BACH dataset has four types of labeled breast histopathologcal images, In situ, Invasive, Normal, Benign. The authors for their work introduce a parallel structure based on DNN that utilizes Convolutional Neural Network "DenseNet" and Recurrent Neural Network "LSTM", they also use Switch normalization instead of batch normalization and in regularization method, they replace the general dropout by targeted dropout, during testing process a data augmentation technique is performed to enhance the test set that consist of rotation, clipping, scaling and other techniques. In their model in [8] they obtain 97.5% accuracy. In [9] the Carlos A. F. and his team used pre-trained ImageNet for transfer learning and inception-Resnet-v2 for the CNN network on BACH dataset. In order to extract features they also used the fine-tuning the frozen layers, the authors got the accuracy of 0.76 in blind testing set after data augmentation for multi class classification. R. Yan et. al. in [10] proposed a hybrid Convolutional neural network and recurrent neural network for classification of histology breast cancer images. Initially they do preprocessing and data enhancement by H&E staining then the image is portioned into 12 patches and used GoogleNet inception-V3 to extract the features. For classification they used bidirectional Long Short Term Memory to classify the images into multi-classes. The average accuracy gained from this method was 91.3%. In [11] Wang J et. al. utilize an state-of-art network "Boosted EfficientNet" to detect the lymph node metastases in breast cancer histopathologic images from Rectified Patch Camelyon RPCam dataset. They mentioned most of the image resolution of lymph node metastasis in the RPCam dataset is not fitted as input of CNN network due to its small scale, therefore they used RCC or Random Center Cropping and a Boosted CNN network to address this issue. However for a better result they used an attention mechanism as well, in which they reduced the redundant information comes from background of image, finally they find the 97.96% accuracy and the 99.68% AUC with this network. In [12] Chanaleä M. et al. proposed a new architecture by using transfer learning and EfficientNet network for the classification the histopathologic images from ICIAR2018 dataset into Normal, Benign, In-situ carcinoma and Invasive carcinoma. To compare the results in their research they use both Reinhard and Macankon stain normalization stain methods. The training have done on eight types of EfficientNet and find the best result on EfficientNet B2, after using ImageNet [13] to pretrain and fine-finetuning the dataset they obtain the accuracy of training process in 98.33%. Researches depict that the EfficientNet have shown great results in works and worth to employ more for the breast cancer classification, as the CNN architecture needs more data to train better and provide a better result we are going to use the BreakHis Dataset for our proposed work which has the 7909 images.

III. METHODOLOGY

In this part we introduce the method utilized in our proposed work. For dataset we choose histopathological images of breast cancer from BreakHis database and applied DNN network on it to train and we also use this model for test set to predict the class of blind images, whether it is benign or malignant.

A. Proposed Method Using Histopathological Images

The main consistence of our proposed method that has been utilized for the classification of histopathological images into benign or malignant described here. Due to the importance of breast cancer treatment, breast cancer detection using CAD is a remarkable debate for researchers and there are already a considerable amount of researches on histopathological images, but still for a better result in health science to find the best accuracy, researches must continue on by using state-of-art architectures.

As it is shown in Fig.1. First step towards every DNN research is to acquire the dataset, here the BreakHis dataset [14], which is publically available, have been used, the details of these

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 09:32:44 UTC from IEEE Xplore. Restrictions apply.

histopathological images have been described ahead. Then we split the data into 80% and 20% ratio respectively for training and testing process. Next we apply the preprocessing for each input image to make the images suitable for training and a better result. Then the data augmentation is applied on training set, further the ImageNet weights for the EfficientNet has been utilized to pre-train the model and at the end, the model used learnt parameters for evaluating and classify the test set.

1) BreakHis Image Acquisitions

BreakHis database compiled from the result of an investigation by P&D Lab, Brazil. Breast tissues are taken as a specimen by the procedure of surgical (open) biopsy (SOB). Samples are stained by Hematoxylin and Eosin (H&E) and conducted by a standard paraffin process in which specimen infiltration and embedment are done in paraffin.

The BreakHis images consist of TrueColor (8bits per color channel) and three-channel color (Red,Green,Blue) coding arrangement. This database has total amount of 7909 images with 700×460 pixel resolution. The allotment of images is categorized below in TABLE I.

TABLE I. AMOUNT OF LABELED IMAGES IN DATASET

magnification	malignant	benign	total
40x	652	1370	**1995**
100x	644	1437	**2081**
200x	623	1390	**2013**
400x	588	1232	**1820**
Total	**2480**	**5429**	**7909**

2) Pre-Processing

For classification using deep learning, preprocessing is a vital part of workflow. To make the suitable input images for CNN architecture, some changes are expected, here we are resizing images as the resolution of original images in BreakHis dataset is 700×460 so it is generally bigger than what is needed for the EfficientNet networks and should have to size normalize it. To perform the resizing we crop and rescale the data.

3) Data Augmentation

CNN networks need a large amount of data to obtain a better result, also data imbalance is prone to overfitting, an issue whereby the networks result in high accuracy for training process but during testing the performance is poor, in other word not good as compared to training accuracy, in order to cope with this issue we are applying data augmentation method in this work. We have extended our dataset with zooming, rotation and flipping in both vertical and horizontal direction.

TABLE II. DATA AUGMENTATION PROCESS

Transformation	Value
Zooming Range	2
Rotation Range	90 degree
Horizontal flipping	True
Vertical flipping	True

4) Transfer Learning

Transfer Learning or TL is a process in which we reuse the pre-trained CNN model and obtained information from another semi-related task, for our problem. TL is worthy as it needs less computational time and also it will reduce the cost overall. Medical imaging is a time-consuming and complex task, TL can help to obtain a better accuracy and boost the proficiency in a quicker time in comparison to training our model from scratch.

By fine-tuning we will fix the first layers that has trained from another model and contains the primary information which is mostly common in all images (e.g. edges), and make the last layers to retrain the layers again which is more task specific and consists of important and more unique features from our dataset.

Here we are using ImageNet for our pre-trained model to initiate the primitive weights then we use the weights for our EfficientNet CNN architecture.

ImageNet is a dataset of images with twenty two thousand of different labeled classes with fourteen million images. Here we are using weights of ImageNet online and download it through Keras library for our code in python.

B. Training Methodology

In [15] the authors by study and precising the convolution network scaling, pointed out the importance of balancing network between width, depth and resolution that was still not taken into consideration in investigations. EfficientNet is a scaled up baseline convolution neural network, that shows many effective and accurate results and surpasses the other state-of-art architecture results. EfficientNet apart from its effectiveness, uses less floating point per second (FLOS) and parameters on ImageNet.

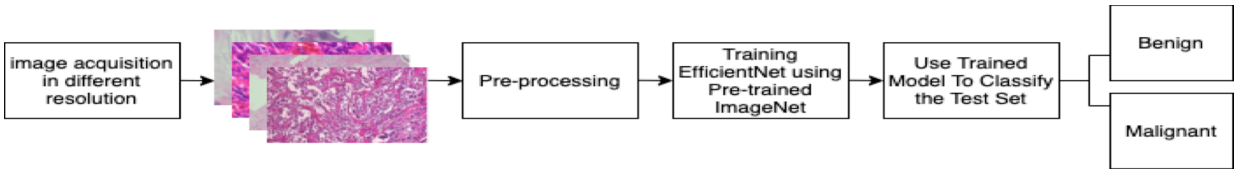


Fig. 1.Shows the framework for our proposed method in histopathological images.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 09:32:44 UTC from IEEE Xplore. Restrictions apply.

ImageNet dataset, lead the machine learning to escalate the performance, still there were some limitations such as computational workload [16]. EfiicientNet obtains 84.4 % accuracy in the ImageNet dataset and showed that this new architecture gives a higher value with only 66 million parameters to train and less computational time among other state-of-art architectures.

EfiicientNet is made from 8 different levels (B0 to B7) and with the increment in versions the parameters and the complexity of model increases; it also uses Swish activation function instead of Rectified Linear Unit (ReLU)[17].

The main goal of DNN approaches is to develop new architectures with high performance and less computation load, EfficientNet reaches to this goal by scaling depth, width and resolution wise. Initially in order to developing the compound scaling method we should find the coefficient that we can apply it on the baseline network to scale up depth, width and resolution dimensions. Architecture in EfficientNet is made from the bottleneck MBConv blocks, that has gained from the MobileNetV2 architecture [18], however in EfficientNet due to the higher floating point operation per second (FLOPS) budget, MBConv used awhile more. Blocks in MBConv include a layer that primarily enlarge and then compact the channel, in order to make a straight connection between bottlenecks which has less channel connections in comparison to expansion layers.

Due to the in-depth separation convolution functionality the EfficientNet model decreased the computational time with the k^2 factor in comparison to the traditional approaches that K represents the kernel size, and height and width is the 2D convolution window[18].Compound coefficient denoted by ψ, utilized in below equation to show the relationship among depth (D), width (W), and resolution (R) in balance, it is specified by the user, and audit the avail resources to scale the model.

$$D = \alpha^{\psi} \qquad W = \beta^{\psi} \quad R = \gamma^{\psi} \qquad \alpha \geq 1, \beta \geq 1, \gamma \geq 1 \quad (3.1)$$

α, β, γ are constants and evaluated by investigation in network, they simply evaluate that to what extend these additional resources are dedicated to the width, depth and resolution of network in orderly. By giving the different values of ψ we can obtain B0 to B7. We also use an attention mechanism to boost the performance, the learning rate in every Neural Network layer is the amount of steps to investigate in the layers with, here we are using the Learn Control method, by which after 5 epochs, if the validation accuracy did not change the LR will reduce automatically to obtain a better result. More information is provided in TABLE III.

TABLE III. LEARNING CONTROL PARAMETERS

Reduce LR	Value
Quantity to be monitored	Validation Accuracy
Factor	0.2
No. of epochs with no improvements	5
Minimum Learning Rate	Le-7

C. *Evaluation Metrics*

Evaluation Metrics are very strong equations to show the evaluation of our model in binary classification, tables below show the evaluation metrics and their equations used for our proposed work.

Evaluation Metrics	Equation
Rate for Accuracy	TP+TN/(TP+FP+TN+FN)
Precision or the positive value predicted	TP/ (TP+FP)
Recall or Sensitivity is the rate for Positive prediction	TP/ (TP+FN)
F1 Score	(2xSensitivityxPrecision)/ (Sensitivity+Precision)

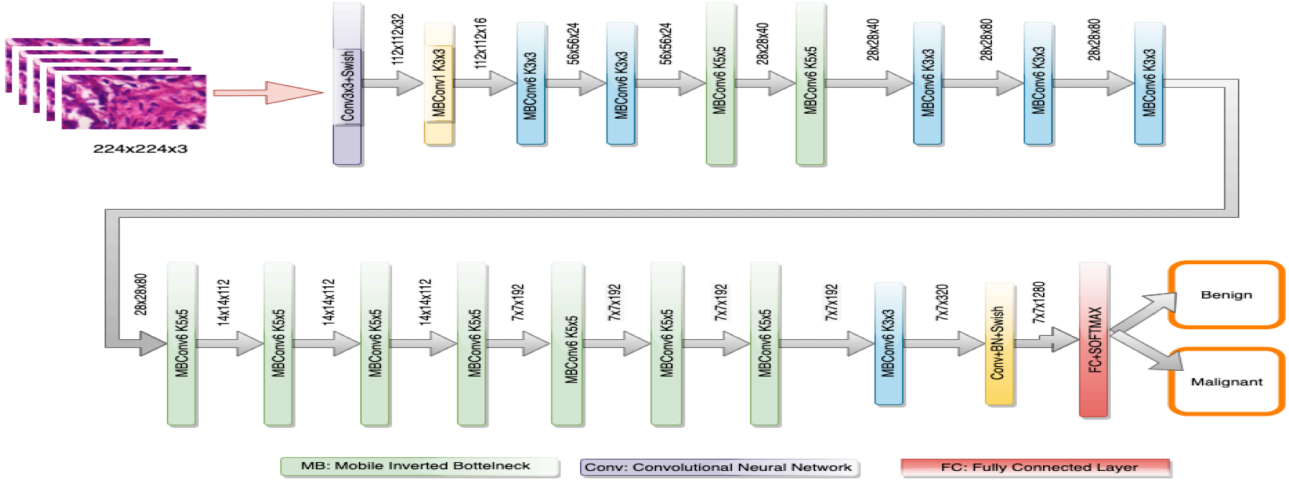


Fig. 2. Representation of EfficientNetB0 for the proposed work

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 09:32:44 UTC from IEEE Xplore. Restrictions apply.

IV. RESULTS AND DISCUSSION

In this section we represent the results for classification with EfficientNet using Histopathological images. Our dataset contains 7907 images.We split the BreakHis dataset into 80% for training set and 20% for testing set part. We use data augmentation on the training set to reduce the misbalancing in dataset, furthermore, we also applied dropout method in every pre-trained layer to reduce the overfitting as well. Adam or Adaptive Moment Estimation is used in our proposed work as an optimizer and the learning rate is le-4 (0.0001), We evaluate our model performance with the evaluation metrics like classification accuracy, precision, Recall, and F1 score.

Here we are investigating the eight levels of EfficientNet (From EfficientNetB0 to EfficientNetB7) Network for images in four resolutions: 40X, 100X, 200X 400X.

The highest Accuracy gained from this proposed work is refer to the 200X magnification with EfficientNetB6 network, which is 99.28% (fig. 3). Surprisingly in this work it is showing that by increasing the parameters in different levels of efficientNet the accuracy for the same epoch is not getting higher and each of the networks has its own characteristic for the classification. In every table the best result has been bolded.

The result for our proposed method surpasses recent findings on other histopathology methods by achieving higher accuracy. Table 4.5 shows the comparison among state-of-art proposed methods on histopathological images with our method.

V. COCLUSION AND FEUTURE WORK

A. Conclusion

In this work we have proposed a deep learning method for Breast cancer image classification, we concentrate on histopathological images form BreakHis dataset which contains 7909 images in four different resolution in 40x, 100x, 200x and 400x, then we applied state-of-art deep learning network to automatically train and predict the blind test set images, to perform this classification we utilized eight different levels of EfficientNet and also we used ImageNet weights for transfer leaning instead of training from scratch, then we used Learn Control on test set to reduce the Learning Rate in case the predicted accuracy does not increase in five continues epochs. We compare the different levels of network with each other in different resolutions (TABLE IV.) and also compare the best results obtained from EfficientNet with other recent works on histopathological images (TABLE V.), the results shows that our proposed method surpass the previous recent methods with the higher predicted accuracy of 99.28%

TABLE IV. RESULTS OF FOUR DIFFERENT RESOLUTION (BEST RESULT HAS HIGHLIGHTED)

	EfficientNet	Training Accuracy%	Testing Accuracy%	Training Loss	Testing Loss	Precision Average%	Recall Average%	F1-Score Average%
40X Resolution	**B0**	93.65	**99.2**	0.17	0.03	99	99	99
	B1	96.39	97.44	0.08	0.09	97	97	97
	B2	94.52	98.37	0.14	0.07	98	99	98
	B3	94.11	98.37	0.15	0.05	98	98	98
	B4	96.56	97.9	0.10	0.04	98	98	98
	B5	95.05	98.37	0.14	0.03	98	98	98
	B6	98.83	94.41	0.15	0.07	99	99	99
	B7	95.75	98.83	0.12	0.04	99	99	99
100X Resolution	B0	94.77	96.88	0.15	0.09	97	95	96
	B1	94.47	97.6	0.18	0.07	98	97	97
	B2	93.93	**98.8**	0.16	0.03	98	99	99
	B3	93.75	98.56	0.15	0.04	98	98	98
	B4	94.53	96.64	0.16	0.08	96	97	96
	B5	95.61	98.32	0.13	0.04	98	98	98
	B6	95.37	96.64	0.13	0.07	96	95	96
	B7	94.65	97.6	0.13	0.07	97	98	97
200X Resolution	B0	95.09	96.77	0.13	0.08	97	96	97
	B1	93.91	98.26	0.16	0.06	97	99	98
	B2	94.47	99.01	0.14	0.05	99	99	99
	B3	94.47	97.02	0.14	0.09	96	97	97
	B4	95.4	99.01	0.12	0.03	99	99	99
	B5	95.22	99.26	0.13	0.03	99	99	99
	B6	94.41	**99.28**	0.15	0.04	99	99	99
	B7	95.59	98.01	0.12	0.06	97	98	98
400X Resolution	**B0**	94.87	**98.60**	0.14	0.05	97	96	95
	B1	92.03	97.25	0.19	0.07	96	98	97
	B2	91.96	96.15	0.21	0.1	96	95	96
	B3	93.61	97.25	0.1642	0.09	97	97	97
	B4	94.51	94.23	0.16	0.16	95	92	93
	B5	93.48	97.53	0.16	0.13	98	98	98
	B6	93.61	94.23	0.16	0.13	95	92	93
	B7	94.92	94.51	0.15	0.22	95	95	94

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 09:32:44 UTC from IEEE Xplore. Restrictions apply.

in 200x image resolution and 99% of Precision , Recall and F1-Score.

B. Future Work

Breast Cancer arouse high attention around the world and in the time of AI that shows many breakthrough results in health issues we are eagerly want to extend our work to use the more state-of-art deep learning result to escalate the accuracy and obtain more precise results since in the health issues such as breast cancer, we need 100% accurate system, also we want to apply the other datasets with higher and less amount of data on our pre-trained model and same proposed work to see the effect of different dataset.

TABLE V. COMPARISON BETWEEN STATE-OF-ART METHODS WITH OUR PROPOSED WORK

Deep Learning Model	*Reference*	*Dataset*	*Staining*	*Results*
MitosRes-CNN	[6]	TUPAC16	H&E	AUC(0.78)
ResHis	[7]	BreakHis	H&E	ACC (92.52%)
DenseNet+ LSTM	[8]	BACH	H&E	ACC (97.5)
Inception-Resnet-v2	[9]	BACH	H&E	ACC (0.76) in test set
Boosted EfficientNet	[11]	RPCam	H&E	ACC(97.96%)
EfficientNet	[12]	BACH	H&E	ACC(98.33%) in Training set
EfficientNet (Our Proposed Work)	-	**BreakHis**	**H&E**	**ACC(99.28%)**

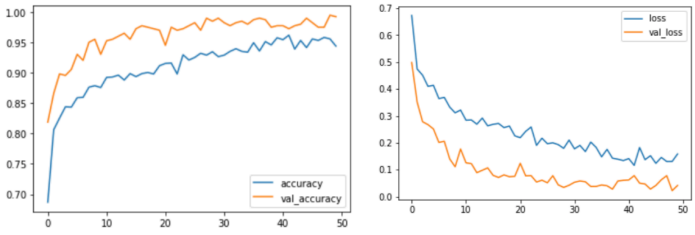


Fig. 3. (Left) Shows the training and predicted accuracy graph using EfficientNetB6 for 200X magnification images.
(Right) Shows the training and predicted loss graph using EfficientNetB6 for 200X magnification images.

References

[1] "breast cancer fact and figure 2019-2020," American Cancer Society, Atlanta, Georgia., 2019. [Online]. Available: https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/breast-cancer-facts-and-figures/breast-cancer-facts-and-figures-2019-2020.pdf

[2] B. E. Sirovich and J. Sox, "Breast cancer screening," Surgical Clinics of North America, vol. 79, no. 5, pp. 961–990, 1999, doi: 10.1016/S0039-6109(05)70056-6.

[3] He L, Long LR, Antani S, Thoma GR. Histology image analysis for carcinoma detection and grading. Comput Methods Pro- grams Biomed. 2012;107:538-556.

[4] S. Z. Ramadan, "Methods Used in Computer-Aided Diagnosis for Breast Cancer Detection Using Mammograms: A Review," Journal of Healthcare Engineering, vol. 2020, pp. 1–21, 2020.

[5] X. Zhou, C. Li, M. M. Rahaman, Y. Yao, S. Ai, C. Sun, Q. Wang, Y. Zhang, M. Li, X. Li, T. Jiang, D. Xue, S. Qi, and Y. Teng, "A Comprehensive Review for Breast Histopathology Image Analysis Using Classical and Deep Neural Networks," IEEE Access, vol. 8, pp. 90931–90956, 2020

[6] A. Sohail, A. Khan, N. Wahab, A. Zameer, and S. Khan, "A multi-phase deep CNN based mitosis detection framework for breast cancer histopathological images," Scientific Reports, vol. 11, no. 1, 2021 https://doi.org/10.1038/s41598-021-85652-1

[7] Gour, M, Jain, S, Sunil Kumar, T. Residual learning based CNN for breast cancer histopathological image classification. Int J Imaging Syst Technol. 2020; 30: 621– 635. https://doi.org/10.1002/ima.22403

[8] Yao, H.; Zhang, X.; Zhou, X.; Liu, S. Parallel Structure Deep Neural Network Using CNN and RNN with an Attention Mechanism for Breast Cancer Histology ImageClassification. Cancers 2019, 11,1901.https://doi.org/10.3390/cancers11121901

[9] C. A. Ferreira, T. Melo, P. Sousa, M. I. Meyer, E. Shakibapour, P. Costa, and A. Campilho, "Classification of Breast Cancer Histology Images Through Transfer Learning Using a Pre-trained Inception Resnet V2," Lecture Notes in Computer Science Image Analysis and Recognition, pp.763–770,2018. https://doi.org/10.1007/978-3-319-93000-8_86

[10] R. Yan, F. Ren, Z. Wang, L. Wang, T. Zhang, Y. Liu, X. Rao, C. Zheng, and F. Zhang, "Breast cancer histopathological image classification using a hybrid deep neural network," Methods, vol. 173, pp. 52–60, 2020.https://doi.org/10.1016/j.ymeth.2019.06.014

[11] Wang J, Liu Q, Xie H, Yang Z, Zhou H. Boosted EfficientNet: Detection of Lymph Node Metastases in Breast Cancer Using Convolutional Neural Networks. Cancers. 2021; 13(4):661. https://doi.org/10.3390/cancers13040661

[12] Eosin-Stained Breast Cancer Histology Microscopy Images Using Transfer Learning with EfficientNets", Computational Intelligence and Neuroscience, vol. 2021, Article ID 5580914, 17 pages, 2021. https://doi.org/10.1155/2021/5580914

[13] J. Deng, W. Dong, R. Socher, L. Li, Kai Li and Li Fei-Fei, "ImageNet: A large-scale hierarchical image database," 2009 IEEE Conference on Computer Vision and Pattern Recognition, 2009, pp. 248-255, doi: 10.1109/CVPR.2009.5206848.

[14] (2015). BreakHis: Breast Cancer Histopathological Database BreakHis. [Online]. Available: http://web.inf.ufpr.br/vri/databases/breast-cancerhistopathological-database-breakhis/

[15] Tan, Mingxing, and Quoc Le. "EfficientNet: Rethinking model scaling for convolutional neural networks." International Conference on Machine Learning. PMLR, 2019.

[16] Ü. Atila, M. Uçar, K. Akyol, and E. Uçar, "Plant leaf disease classification using EfficientNet deep learning model," Ecological Informatics, vol. 61, p. 101182, 2021. https://doi.org/10.1016/j.ecoinf.2020.10118.

[17] W. Hao, M. Han, H. Yang, F. Hao, and F. Li, "A novel Chinese herbal medicine classification approach based on EfficientNet," Systems Science & Control Engineering, vol. 9, no. 1, pp. 304–313, 2021. DOI: 10.1080/21642583.2021.1901159

[18] Mark, S., Andrew, H., Menglong, Z., Andrey, Z., Liang-Chieh, C., 2018. Mobilenetv2: inverted residuals and linear bottlenecks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4510–4520.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 09:32:44 UTC from IEEE Xplore. Restrictions apply.

International Journal of Clinical Oncology
https://doi.org/10.1007/s10147-021-02108-2

ORIGINAL ARTICLE

Mapping of global, regional and national incidence, mortality and mortality-to-incidence ratio of lung cancer in 2020 and 2050

Rajesh Sharma[1]

Received: 11 October 2021 / Accepted: 15 December 2021
© The Author(s) under exclusive licence to Japan Society of Clinical Oncology 2022

Abstract

Aim Lung cancer is the leading cause of cancer deaths worldwide. This study examines the current and future burden of lung cancer at global, regional, and national levels.

Methods The estimates of lung cancer incident cases, deaths, and their age-standardized rates are drawn from GLOBOCAN 2020 for 21 regions and 185 countries. Mortality-to-incidence ratio (MIR) is considered as a proxy indicator of 5-year survival rates. Lung cancer burden in 2050 is projected using age-specific incidence and death rates in 2020.

Results In 2020, there were 2.21 million new cases and 1.8 million deaths due to lung cancer worldwide with age-standardized incidence rate (ASIR) of 22.4/100,000 (male: 31.5; female: 14.6) and age-standardized mortality rates (ASMR) of 18.0/100,000, (male: 25.9; female: 11.2/100,000). Global MIR of lung cancer was 0.82 (males 0.83; females: 0.79), varying from 0.59 (Japan) to 1.0 (Belize). Hungary had the highest age-standardized rates (ASIR: 50.1/100,000; ASMR: 42.4/100,000) and Nigeria (ASIR: 0.88; ASMR: 0.86) had the lowest age-standardized rates in 2020. Both ASIR and ASMR were positively correlated with country-level tobacco smoking prevalence and human development index (HDI), whereas MIR exhibited a negative correlation with HDI. As per our projections, there will be 3.8 million incident cases and 3.2 million deaths globally due to lung cancer in 2050.

Conclusion With close to 2 million cases and deaths already in 2020, lung cancer has already become a global public health threat. Even with current risk levels and age-specific rates, lung cancer annual cases are expected to reach 3.8 million in 2050. Until smoking prevalence is reduced and ambient air pollution levels are checked, particularly in low/medium HDI countries, the lung cancer epidemic will continue unfolding.

Keywords Lung cancer · Incidence · Mortality · Epidemiology · GLOBOCAN

Introduction

Lung cancer is the second leading cancer in terms of new cases and the leading cancer in terms of cancer deaths worldwide [1]. From a rare disease about a 100 years ago [2], lung cancer has gained prominence as one of the leading cancers by the twenty-first century [1]. The growth in lung cancer burden has largely been attributed to the rise in cigarette smoking, which is expected to have peaked in industrialized countries (e.g., United States) by the later half of the twentieth century, after the landmark study by Doll and Hill [3] and United States surgeon general's report linking smoking as the causative factor of lung cancer [4]. Still, the tobacco smoking epidemic is unfolding in regions of Asia and Africa, due to which, the lung cancer burden is increasing in several countries of these regions [5]. Although lung cancer survival rates are one of the lowest among frequent cancer groups, there are wide disparities in 5-year survival rates across countries, from less than 10% in countries such as Brazil, Bulgaria, India, and Thailand to 32.9% in Japan [6].

About 80–85% of lung cancer is non-small cell lung cancer, which is further sub-divided into adenocarcinoma, squamous cell carcinoma, and large cell carcinoma [7] and 10–15% of lung cancer is small cell lung cancer and the rest is lung carcinoid tumor and other lung cancers [8]. About 8% of lung cancer is inherited or occur due to genetic predisposition [9, 10]. Tobacco smoking is by far the single

✉ Rajesh Sharma
rajesh.sharma@dtu.ac.in

[1] University School of Management and Entrepreneurship, East Delhi Campus, Delhi Technological University, Room No. 305, Vivek Vihar Phase II, Delhi 110095, India

Published online: 12 January 2022

most significant risk factor of lung carcinoma [11], followed by air pollution (particularly particulate matter $PM_{2.5}$) [12], passive smoking [13, 14], and occupational exposure to carcinogens such as radon, and asbestos [15]. The estimated economic cost of lung cancer to world economy was about $180 billion in 2008 [16]. In the United States, the average treatment cost of lung cancer exceeds $45,000 as compared to approximately $3,000 in non-cancer patients [17]. As lung cancer is the leading cause of deaths and entails a huge economic burden globally, it is pertinent to understand the lung cancer burden such that interventions can be designed and implemented appropriately.

In this study, we provide a comprehensive and comparable examination of lung cancer epidemiology at global, regional, and national levels in 2020 using recently released GLOBOCAN 2020 estimates produced by International Agency for Research on Cancer [1, 18]. The lung cancer burden is evaluated using age, sex, and country-wise estimates of number of incident cases and deaths, age-standardized rates (per 100,000 population) and mortality-to-incidence ratio (MIR) in 185 countries. The human development index (HDI) is considered as the indicator of socio-economic status of countries. As tobacco smoking is the biggest risk factor of lung cancer, we also investigate the bivariate association between tobacco smoking prevalence and age-standardized rates of lung cancer. Lastly, we project the lung cancer incident cases and deaths to 2050 using the age-specific estimates of incidence and death rates in 2020 and age-specific population projections.

This comprehensive report on lung cancer epidemiology in the light of HDI and tobacco smoking prevalence is expected to aid our understanding of the lung cancer burden and help in policy formulation towards fighting lung cancer burden at the global, regional, and national levels.

Data and Methods

The age, sex, and country-specific estimates of lung cancer incidence and mortality in 185 countries for 2020 are procured from GLOBOCAN 2020 database [19]. The lung cancer burden described and examined in this study pertains to codes C33–34 (including trachea and bronchus) of the international classification of disease tenth version (ICD 10). The MIR for each country is calculated as the ratio between all-age death counts and all-age incident cases of lung cancer. The absolute counts (not the age-standardized rates) are provided with 95% uncertainty intervals. The age-specific counts and age-specific rates (per 100,000 populations) are available for eighteen 5-year non-overlapping age groups from below −5 to 85-plus. The age-standardized rates are expressed per 100,000 population and are calculated using Segi's method [1, 18].

The estimates of lung cancer are available for 21 regions defined by the United Nations: Australia/New Zealand, Caribbean, Central America, Central, and Eastern Europe, Eastern Africa, Eastern Asia, Melanesia, Micronesia, Middle Africa, Northern Africa, Northern America, Northern Europe, Polynesia, South America, South-Central Asia, South-Eastern Asia, Southern Africa, Southern Europe, Western Africa, Western Asia, and Western Europe. Using age-specific rates of incidence and mortality in 21 regions in 2020 and age-specific population projections from the United Nations world population prospects (WPP) report, we project the lung cancer burden in 2050. The basic assumption underlying these 2050 projections is that the age-specific rates will remain unchanged by 2050 (i.e., the risk factors and the state of health care stays the same in each region). The equation for arriving at 2050 projection of incident cases is described below.

$$IC_{2050} = \sum_{i=1}^{n} IR_{2020i} \times P_{2050i}$$

In the above equation, IC_{2050} denotes total number of incident cases in 2050, i denotes the age group, IR_{2020i} implies age group specific incidence rate in 2020 in the i^{th} age group and P_{2050i} denotes the projected population count in the i^{th} age group in 2050 as per WPP report by the United Nations. The estimates of number of deaths due to lung cancer in 2050 are also produced in the same manner. The HDI for 2019 represented a country's socio-economic status or development [20]. We also investigate the association between lung cancer age-standardized rates and tobacco smoking prevalence in countries. For this purpose, we employ Pearson's correlation coefficient to describe the bivariate association. The data on age-standardized tobacco smoking prevalence (%) are procured from Global Health Observatory database of World Health Organization (WHO), which provided data for 149 countries and 6 WHO regions [Africa (32 countries), Americas (22), Eastern Mediterranean (15), Europe (48), South-East Asia (9), Western Pacific (23)] [21].

All the data analysis and data visualizations in this study are done using statistical software R 4.1.1, Stata 13.1, and Python 3.8.8.

Results

Global burden of lung cancer in 2020

Globally, there were 2.21 million [2.18–2.24 million] new cases and 1.8 million [1.77–1.83 million] deaths due to lung cancer in 2020. Males accounted for approximately two-thirds of the global burden (1.44 million [1.41–1.46 million] new cases and 1.19 million [1.16–1.21 million] deaths). The

global age-standardized incidence rate (ASIR) of lung cancer was 22.4/100,000, with males ASIR (31.5/100,000) being more than double that of females (14.6/100,000). Global age-standardized mortality rates (ASMR) was 18.0/100,000, varying from 11.2/100,000 in females to 25.9/100,000 in males. The MIR at the global level was 0.82 with not much difference between males (0.83) and females (0.79).

Lung cancer burden by region in 2020

East Asia was the leading region in terms of incident cases and deaths with an estimated 1.01 million [1.00–1.02 million] incident cases and 841,174 [831,852–850,601] deaths in 2020, followed by Northern America (cases: 253,537 [252,452–254,627]; deaths: 159,641 [158,736–160,551]). Lung cancer was ranked among the top three cancer groups out of 34 cancer groups in terms of new cases in 11/21 regions and among top three in 15/21 regions in terms of cancer deaths (Table 1). The MIR of lung cancer varied from 0.63 in Northern America to 0.96 in Micronesia, with MIR being 0.80 or above in the majority of regions except Australia/New Zealand (0.69), Northern Europe (0.73) and Western Europe (0.78). Polynesia had the highest ASIR (37.3/100,000) and Micronesia had the highest ASMR (34.9/100,000). Majority of African regions had low age-standardized rates, with Western Africa recording the lowest ASIR (2.2/100,000) and ASMR (2.1/100,000).

The male–female disparity in age-standardized rates was again visible across regions (Supplementary Fig. S1), with male age-standardized rates being higher than females in almost all regions. The male–female differences in age-standardized rates were starker in regions of Asia (Western Asia, South-East Asia) and Africa (Northern Africa, Southern Africa) than in Northern America, Northern Europe, Western Europe, and Australia/New Zealand (Supplementary Fig. S1).

Lung cancer burden by country

In absolute terms, China, United States, and Japan were the three leading countries with 815,563 [808,051–823,145], 227,875 [223,940–231,879], and 138,532 [134,618–142,560] incident cases in 2020, respectively; these three countries were also the leading countries in terms of death counts estimated at 714,699 [706,501–722,992], 138,225 [136,168–140,313], and 82,369 [80,835–83,932], respectively. In terms of age-standardized rates, Hungary was the leading country (ASIR:50.1/100,000 and ASMR:42.4/100,000), followed by Serbia (ASIR:47.3/100,000 and ASMR:40.0/100,000). Nigeria (ASIR: 0.88 and ASMR: 0.86) and Mozambique (ASIR: 1.4 and ASMR: 1.4) had the lowest age-standardized rates among 185 countries in 2020 (Fig. 1).

Age-specific incidence and mortality of lung cancer

The lung cancer burden by age-groups illustrates that age groups below 45 years accounted for a very small percentage of lung cancer burden (2.4% of incident cases and 1.9% of deaths). Both the metrics (incidence and deaths) peaked in the 65–69 years age group and declined thereafter (Fig. 2). The age-specific incidence and death rates rose with age such that the oldest age groups had the highest age-specific rates. To gauge the age-specific patterns of lung cancer burden in different regions, we aggregated the age-specific incident cases and deaths in four age groups: below 45 years, 45–59 years, 60–69 years, and 70–85 plus and present the percentage of lung cancer burden in these age groups in Supplementary Fig. S2. In regions of Africa, a sizable fraction of incident cases occurred in age groups below 45 years (e.g., Western Africa 8.9%, Eastern Africa 10.4%, Middle Africa 11.1%), whereas only 1–2% of lung cancer cases occurred in age groups below 45 years in Europe (Central and Eastern Europe 1.6%, Western Europe 1.0%, Southern Europe 1.0%). In regions such as Western Africa and South-Eastern Asia, around 60% of cases and deaths occurred in age groups 60 years or older, whereas in Western Europe and South America, around 80% of new cases and deaths occurred in age groups 60 years or older. In Western Europe, Southern Europe, Northern Europe, and Australia/New Zealand, more than 50% of incident cases and deaths occurred in age groups 70 years or older, while in regions of Africa (Western Africa, Middle Africa, and Eastern Africa), approximately 30% of lung cancer burden occurred in age groups 70 years or older (Supplementary Fig. S2).

Bivariate relationship between HDI and lung cancer burden

Figure 3 displays the bivariate associations between country-level age-standardized rates and MIR with the socio-economic development status of countries measured by HDI. It shows that both ASIR and ASMR were positively associated with HDI with pairwise correlation coefficient of 0.75 and 0.70, respectively; on the other hand, MIR was negatively correlated with country-level HDI ($r = -0.67$). This implies that, on average, countries with higher HDI had higher age-standardized rates (Fig. 3a, b); however, more developed countries (countries with higher HDI) had lower MIR (i.e., higher 5-year survival rates) than lesser developed ones. For instance, high HDI countries such as Japan (0.59), United States (0.61), and Australia

Table 1 Lung cancer burden by region in 2020

Population	Incidence	ASIR	Deaths	ASMR	MIR
Australia and New Zealand	15,587 [15305–15874] [[5]]	25.2	10,791 [10550–11038] [[1]]	16.2	0.69
Caribbean	11,058 [10485–11662] [[4]]	17.6	10,079 [9604–10578] [[6]]	15.8	0.91
Central America	9934 [9615–10264] [[9]]	5.2	9236 [8979–9501] [[6]]	4.8	0.93
Central and Eastern Europe	151,632 [149886–153398] [[3]]	26.9	130,596 [129152–132056] [[1]]	22.7	0.86
Eastern Africa	7419 [5268–10448] [[11]]	3.5	6758 [4577–9978] [[11]]	3.2	0.91
Eastern Asia	1,012,021 [1003100–1021020] [[1]]	34.4	841,174 [831852–850601] [[1]]	28.1	0.83
Melanesia	918 [768–1097] [[6]	13.1	798 [736–865] [[4]]	11.5	0.87
Micronesia	202 [152–269] [[1]]	36.4	194 [146–258] [[1]]	34.9	0.96
Middle Africa	2037 [1382–3003] [[12]]	2.5	1897 [1287–2795] [[10]]	2.4	0.93
Northern Africa	23,179 [20865–25750] [[3]]	11.1	20,728 [18443–23296] [[3]]	9.9	0.89
Northern America	253,537 [252452–254627] [[2]]	32.6	159,641 [158736–160551] [[1]]	19.3	0.63
Northern Europe	75,051 [74261–75849] [[2]]	29.7	54,421 [53732–55119] [[1]]	20.1	0.73
Polynesia	268 [212–339] [[1]]	37.3	229 [159–330] [[1]]	31.8	0.85
South America	76,609 [72588–80853] [[4]]	13.6	67,312 [66320–68319] [[1]]	11.8	0.88
South-Central Asia	121,369 [116397–126554] [[4]]	6.6	109,356 [104882–114021] [[2]]	5.9	0.9
South-Eastern Asia	123,309 [118258–128576] [[2]]	17.2	109,520 [104946–114294] [[1]]	15.3	0.89
Southern Africa	9178 [8893–9473] [[4]]	16.9	7939 [7707–8178] [[1]]	14.6	0.87
Southern Europe	104,391 [101664–107191] [[3]]	28.7	85,635 [84449–86838] [[1]]	21.9	0.82
Western Africa	4175 [2688–6485] [[11]]	2.2	3849 [3370–4396] [[10]]	2.1	0.92

Table 1 (continued)

Population	Incidence	ASIR	Deaths	ASMR	MIR
Western Asia	58,437 [54808–62307] [[2]]	24.2	52,467 [51245–53718] [[1]]	21.9	0.9
Western Europe	146,460 [144672–148270] [[2]]	32.7	113,524 [112268–114794] [[1]]	23.8	0.78

The MIR is calculated as the ratio between all-age death counts and all-age cases. The numbers in the parenthesis represent 95% uncertainty interval. The number in double parenthesis denotes the rank of lung cancer among 34 cancer groups in that region. Data source: GLOBOCAN 2020 (International Agency for Research on Cancer)

Incidence incident cases, *ASIR* age-standardized incidence rate, *Deaths* number of lung cancer deaths, *ASMR* age-standardized mortality rate; *MIR* mortality-to-incidence ratio

(0.67) had much lower MIR than countries such as Belize, Comoros, and Vanuatu, which had MIR of 1.0 (Fig. 3c).

Smoking and lung cancer

Worldwide, tobacco smoking prevalence was 21.9% in 2016, varying from 4% in Ghana to 47.4% in Kiribati (Fig. 4) with tobacco smoking prevalence exceeding 25% in 57/149 countries. The tobacco smoking prevalence was high in European countries with 5 of the top-10 countries among 149 countries belonged to Europe. Notably, 11/33 countries in Africa had a smoking prevalence less than 10%, whereas all countries in Europe had a smoking prevalence greater than 10%. Next, we explore the bivariate relationship between age-standardized rates of lung cancer and tobacco smoking prevalence (Supplementary Fig. S3). Tobacco smoking prevalence was positively associated with ASIR ($r = 0.61$) and ASMR ($r = 0.64$) (Supplementary Fig. S3a, b). It implies that countries with higher smoking prevalence rates also had higher age-standardized rates of lung cancer.

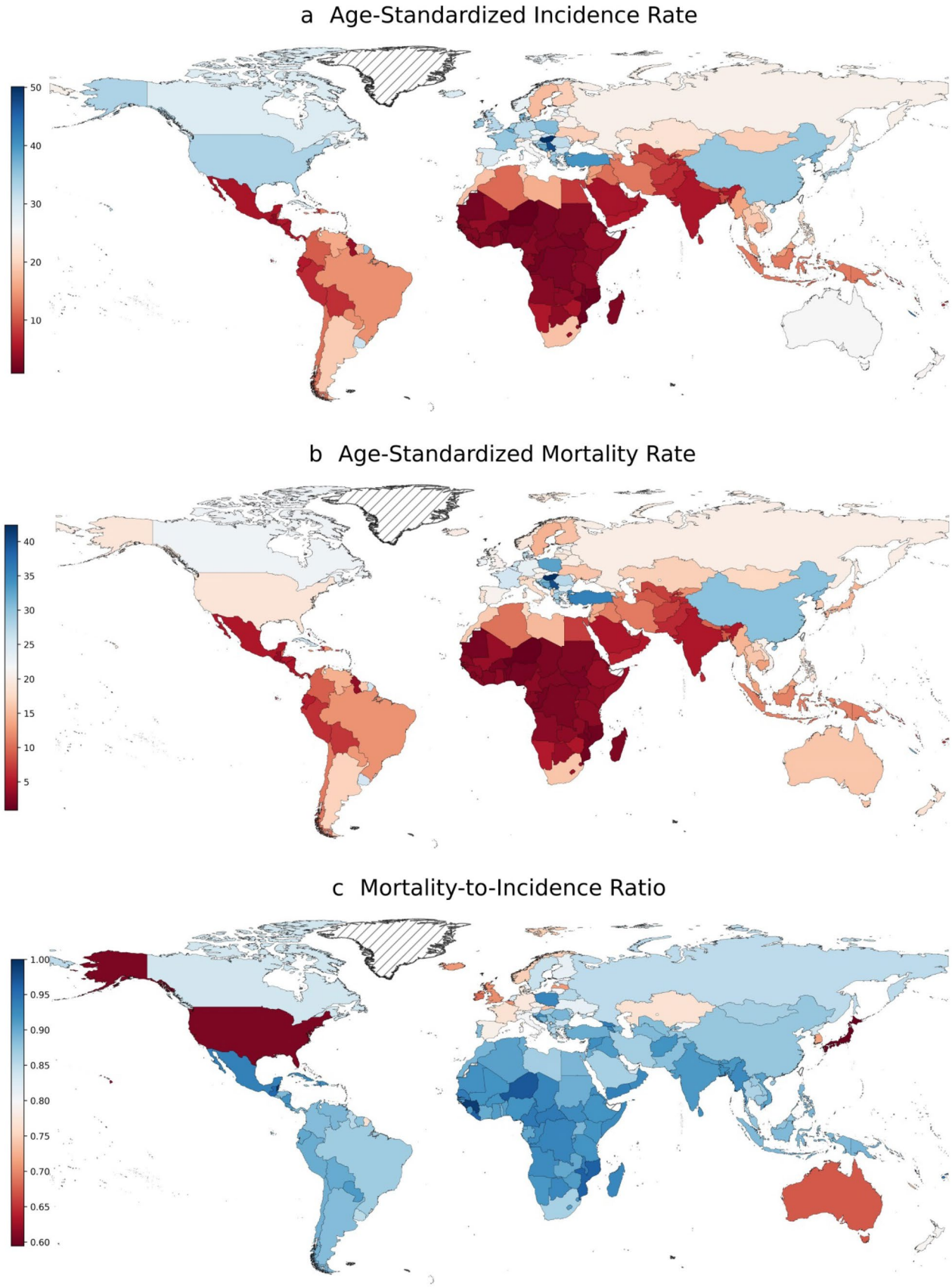
Lung cancer burden in 2050

In 2050, the global burden of lung cancer is projected to reach 3.8 million new cases and 3.2 million deaths each year. In 2050, lung cancer cases and deaths will be more than 100,000 in 10/21 regions led by Eastern Asia, which is projected to record 1.7 million incident cases and 1.5 million deaths (Fig. 5). The burden of lung cancer in regions of Asia (South-Eastern Asia, South-Central Asia, and Western Asia) and Africa (Northern Africa, Southern Africa, and Western Africa) is expected to double or more between 2020 and 2050, whereas regions in Europe (Northern Europe, Southern Europe, Western Europe, Central and Eastern Europe) are expected to witness the smallest increases in the lung cancer burden (Fig. 5). As a result of these changes, the lung cancer burden in several Asian regions will surpass that in a few regions of Europe. For instance, South-Eastern Asia (123,309) and South-Central Asia (121,369) had lesser number of incident cases than Central and Eastern Europe (151,632) and Western Europe (146,460) in 2020; however, by 2050, these two regions in Asia will surpass the number of incident cases in Central and Eastern Europe and Western Europe. In 2050, lung cancer is projected to claim similar number of lives in South-Eastern Asia (248,326) and South-Central Asia (238,020) as that in Northern America (253,058), although the projected number of incident cases in Northern America (378,587) is much higher than projected new cases in South-Eastern Asia (271,416) and South-Central Asia (264,309) (Fig. 5).

Discussion

With close to 2 million incident cases and deaths along with an MIR of 0.82 in 2020, lung cancer is one of the most lethal neoplasms and was ranked first among 34 cancer groups in 13/21 regions in terms of cancer-related deaths in 2020. The age-standardized rates were the highest in regions of Europe and the Americas, where smoking prevalence is still higher than that in Africa and Asia. Due to lower smoking prevalence and lower life expectancy in sub-Saharan Africa (SSA), the age-standardized rates, as well as the absolute burden of lung cancer, is low; as a result, lung cancer, which is currently ranked first in terms of cancer deaths in regions of Europe and America, is ranked much behind other neoplasms in SSA. Country-wise, China, United States, and Japan are the leading countries, accounting for 53.6% of incident cases and 52.1% of lung cancer deaths.

The age-standardized rates of lung cancer displays positive correlation with HDI as well as tobacco smoking prevalence, which implies that more developed countries in Europe, America, Oceania have higher smoking prevalence (also clear from Fig. 4 on smoking prevalence) as well as higher age-standardized rates of lung cancer. In Europe, the rates of lung cancer are decreasing among males in the majority of countries, whereas the lung cancer rates are rising or plateauing among females in several countries and


a Age-Standardized Incidence Rate
50
40
30
20
10
b Age-Standardized Mortality Rate
40
35
30
25
20
15
10
5
c Mortality-to-Incidence Ratio
1.00
0.95
0.90
0.85
0.80
0.75
0.70
0.65
0.60

◂**Fig. 1** Geographical pattern of age-standardized rates of lung cancer in 2020. The mortality-to-incidence ratio is calculated as the ratio between all-age death counts and all-age cases. The age-standardized rates are expressed in terms of counts per 100,000 population. Data source: GLOBOCAN 2020 (International Agency for Research on Cancer)

these patterns most likely reflect the differing patterns of the smoking epidemic among males and females [22].

Although MIR is negatively associated with HDI, yet the 5-year survival rates of lung cancer are one of the lowest (i.e., high MIR) among all cancers, even in the high HDI countries. For instance, the 5-year survival rate of lung cancer is 19% in the United States, one of the lowest among the most frequent cancer groups [23]. Low survival rates of lung cancer also reflect a higher percentage of cases detected in the metastatic phase, which has much lower survival rate (5%) than localized lung cancer (57%) [23]. Low-dose computed tomography (LDCT) has been recently suggested as a screening solution for lung cancer. The LDCT has been shown to reduce lung cancer mortality in both National Lung Screening Trial and Multicentric Italian Lung Detection trials [24, 25]. For select current/former heavy smokers, the American Cancer Society and US Preventive Services Task Force have recommended LDCT screening [23]. However, there are questions on LDCT screening in terms of high false-positive rates [24, 26–28], overdiagnosis [24], large percentage of tumors detected in advanced stages [26] and risk of radiation-induced neoplasms [29]. Moreover, the LDCT screening becomes more challenging in low-resource countries as healthcare workers' expertise in adjusting radiation dose is low and secondly, LDCT screening is expected to cost $81,000 per one quality-adjusted life-year gained [30].

Currently, countries in Africa have very low incidence and mortality rates barring a few countries in Northern and Southern Africa. These patterns also follow the low smoking prevalence rates in Africa, barring a few countries in Northern Africa (e.g., Algeria 15.6%) and Southern Africa (e.g., Lesotho 27.2%; South Africa 20.7%; Botswana: 20.1%). However, in persons living with HIV/AIDS, smoking rates are 2–4 times higher than the general population in Africa [31]. The lung cancer burden in Africa is expected to double or more by 2050 based on demographic changes alone. As smoking initiation before the age of 18 years significantly predicts smoking in adult age (and hence the elevated lung

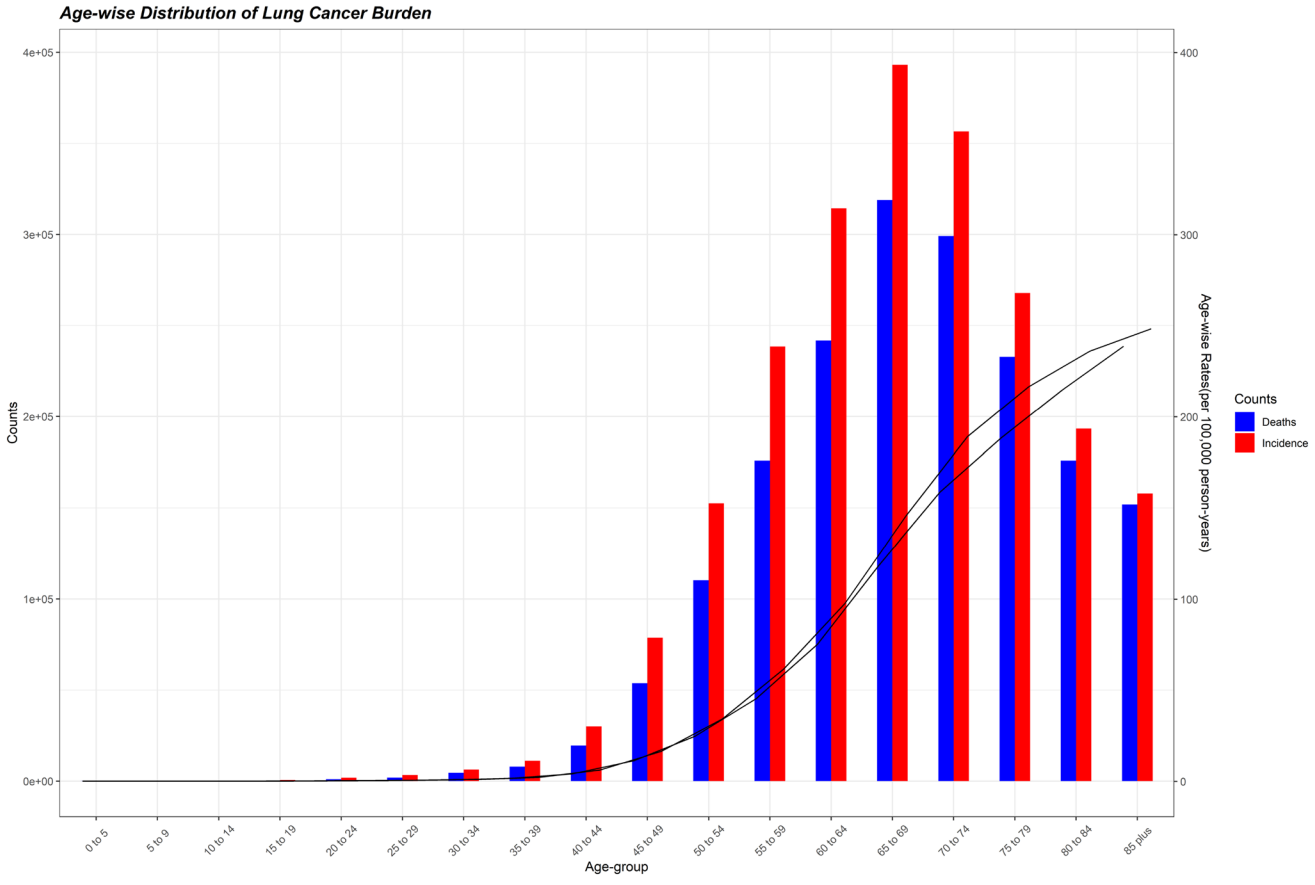


Fig. 2 Age-specific burden of lung cancer in 2020. Deaths: age-specific death count; incidence: age-specific incident cases. Data source: GLOBOCAN 2020 (International Agency for Research on Cancer)

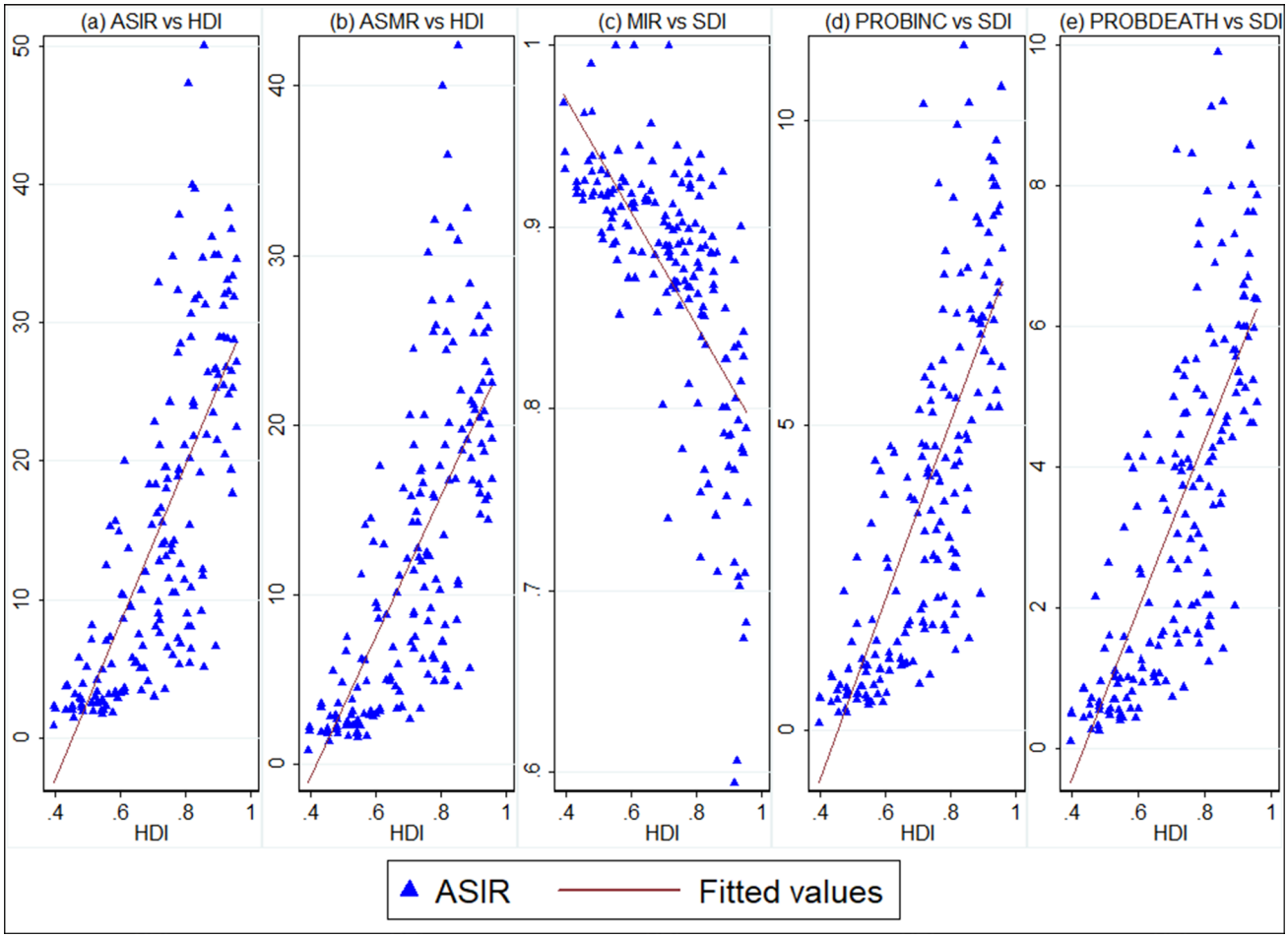


Fig. 3 Bivariate relationship between age-standardized rates, MIR of lung cancer and HDI. *ASIR* age-standardized incidence rate; *ASMR* age-standardized mortality rate; *MIR* mortality-to-incidence ratio; *HDI* human development index. MIR was calculated as the ratio between all-age death counts and all-age incident cases. Data source of cancer estimates: GLOBOCAN 2020 (International Agency for Research on Cancer). The data pertaining to HDI are procured from United Nations Development Program

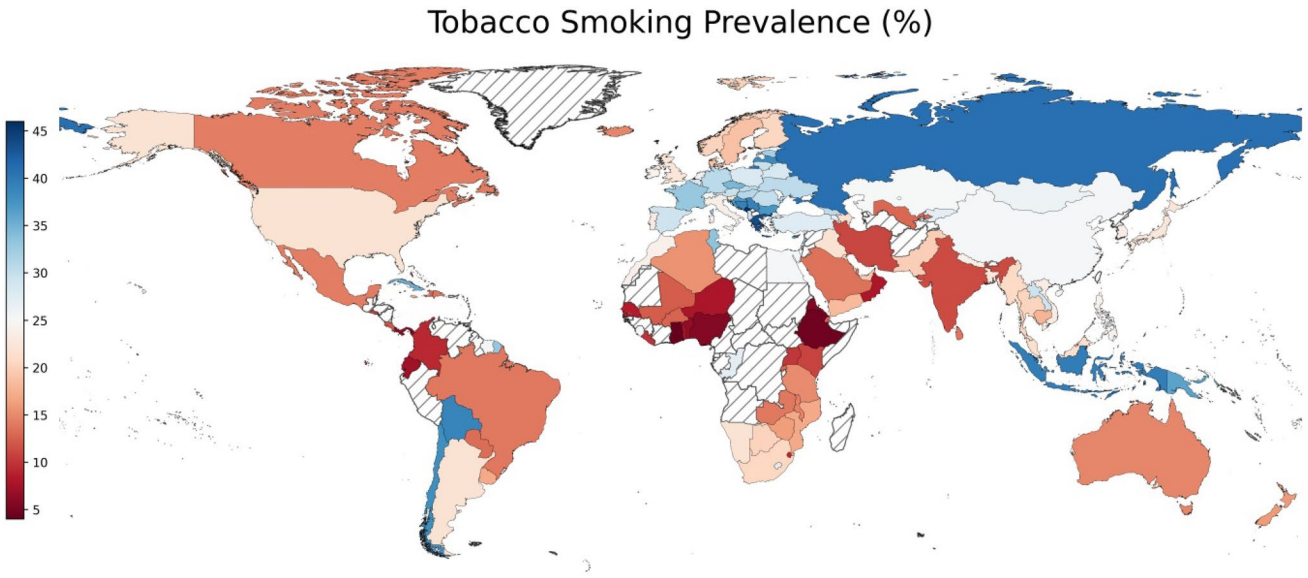


Fig. 4 Geographical pattern of tobacco smoking prevalence. Tobacco smoking prevalence (%) pertains to the year 2016. Data source: World Health Organization (WHO)

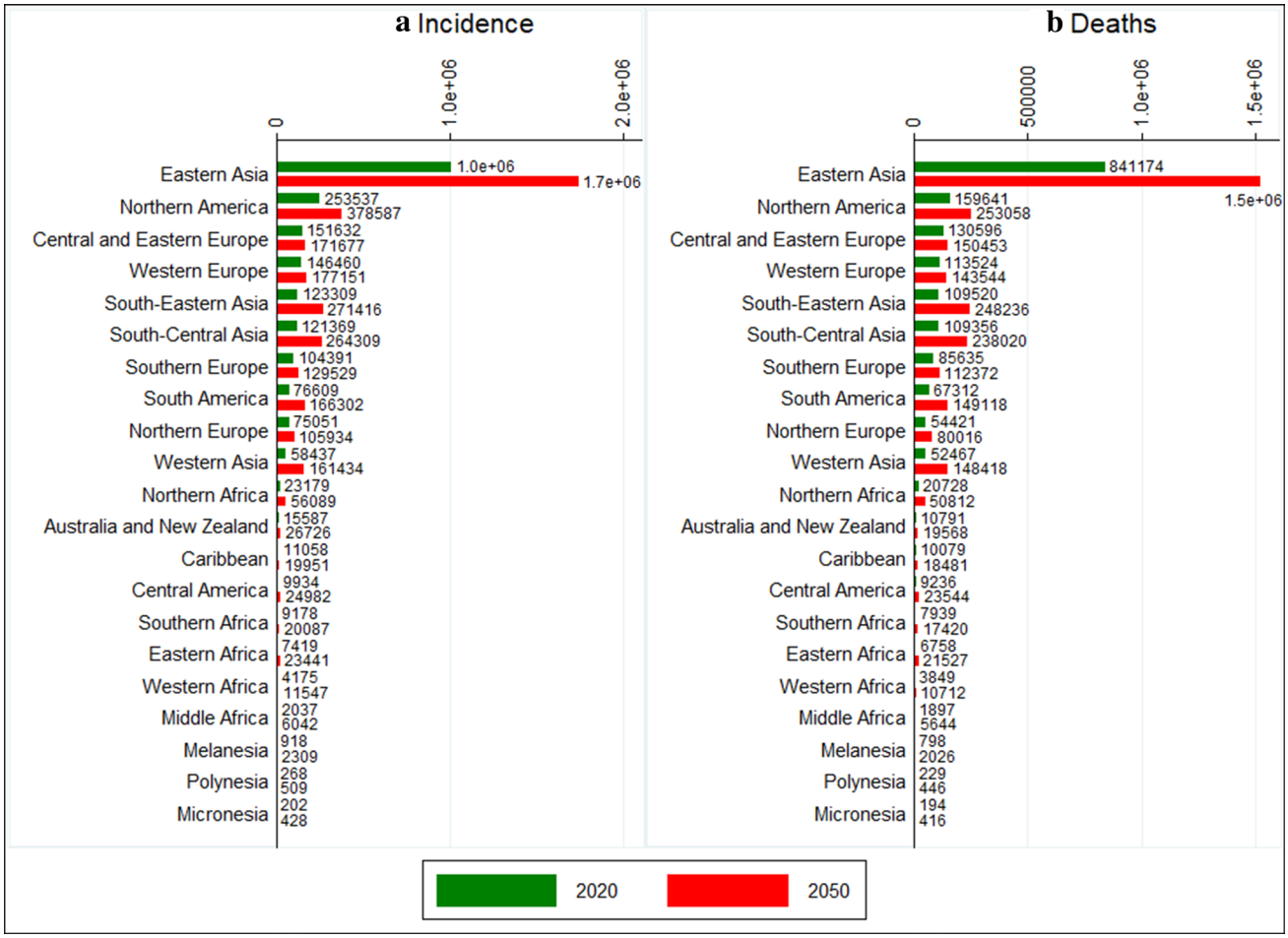


Fig. 5 Region wise projected incidence and deaths of lung cancer in 2050. Incidence: all-age incident cases; deaths: all-age deaths. Data source: authors' calculation based on age-specific rates of lung cancer as per GLOBOCAN 2020 estimates

cancer risk), a study found smoking prevalence as high as 9.6% in Cote d'Ivoire among 12/13 year old [32]. Moreover, air pollution levels are also rising in Africa. As per the State of Global Air 2020 report [33], 5/10 countries with the worst annual average $PM_{2.5}$ concentrations belonged to Africa (Niger: 3rd, Nigeria: 5th; Egypt: 6th; Mauritiania: 7th; Cameroon: 8th). If these trends of early smoking initiation and ambient air pollution levels continue along with rising life expectancy, the lung cancer burden in Africa is expected to be even higher in 2050 than our projections.

Even though Asia is in the initial stages of the smoking epidemic [34–37], the continent already accounted for 60% of incident cases and 62% of all lung cancer deaths in 2020. If current rates persist in the future, there will be 2.4 million new cases and 2.2 million deaths in Asia due to lung cancer in 2050. Previous research has also found that while lung cancer burden has been decreasing in the industrialized countires after a peak in the 2000s, it is gradually increasing in China [38]. A recent study has found that smoking prevalence has increased among males in China and India and plateaued in Japan, South Korea, Singapore, and Taiwan [37]. Moreover, the mean age of smoking initiation has decreased in successive birth cohorts, and the mean number of cigarettes smoked per day has increased [37]. Therefore, anti-smoking campaigns and public awareness campaigns about health risks of smoking are urgently required to stop and reverse the smoking epidemic in Asia.

The absolute counts as well as age-standardized rates of lung cancer are consistently higher in males than females across regions. Higher proportion of lung cancer burden in males is attributed to differences in smoking prevalence, with 80% of males and 50% of females with lung cancer are/were smokers [7]. However, male–female differences in lung cancer burden cannot be fully explained by differences in smoking prevalence. As many as one-third of lung cancer patients in East Asia are non-smokers, most of whom present with lung adenocarcinoma [39]. Even in industrialized countries, high lung cancer incidence rate in women born around 1960 cannot be fully explained by increasing smoking incidence rates in women [40]. Second-hand smoke at home and

workplace and cooking oil fumes are implicated as a few of the risk factors apart from other factors (genetic susceptibility, occupational exposure, hormonal status, and pre-existing lung disease) for rising lung cancer in never-smokers in East Asia [39]. Apart from tobacco smoking (personal or passive), air pollution is another significant risk factor of lung cancer [41], particularly among non-smokers [42, 43]. A meta-analysis of the relationship between $PM_{2.5}$ and lung cancer burden (incidence and deaths) found a higher risk of both lung cancer incidence and death due to higher levels of $PM_{2.5}$ levels, irrespective of smoking status and gender [44].

Limitations

As with all studies, this study also has some limitations. First, the lung cancer estimates of several countries are derived either from regional cancer registries or from neighboring countries due to the absence of data from high-quality population-based cancer registries by the GLOBOCAN [1, 18]. Therefore, in the countries lacking PBCRs, the actual estimates are likely to be different from the estimates reported in this study. Second, the 2020 estimates by GLOBOCAN do not reflect the effect of the COVID-19 pandemic. Therefore, actual cancer-related incident cases and deaths might be different as health services to non-COVID patients have been severely disrupted across the globe. Third, the 2050 projections reflect the demographic changes in population composition but do not reflect changes in risk factors (e.g., smoking, air pollution, among others) and improvements in cancer registration and health care delivery, which may occur in developing countries of Asia and Africa. As a result, due to worsening smoking and air pollution levels in developing countries of Asia and Africa, the lung cancer burden is expected to be much higher than our projections.

Conclusion

Lung cancer incident cases and deaths are projected to reach 3.8 million and 3.2 million, by 2050 based on demographic changes alone. Although age-standardized rates are the highest in high HDI countries of Europe and America, possibly on account of high smoking prevalence; the absolute burden, however, is the largest in developing countries led by China. With worsening smoking prevalence and $PM_{2.5}$ levels in countries such as China, India, and Indonesia; Asia is expected to remain the epicenter of the lung cancer epidemic. As Africa is experiencing increases in smoking incidence, rising life expectancy, and pollution levels, the lung cancer burden is expected to be much higher in Africa by 2050 than our projections. The experience of high-income countries has taught that unless smoking patterns are reversed, the lung cancer cases and deaths are expected to grow unabated in developing countries. Additionally, countering the burden of lung cancer also requires curtailment of other risk factors such as air pollution and exposure to carcinogens.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10147-021-02108-2.

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

Availability of data and material All the data are available on https://gco.iarc.fr/today/online-analysis-table.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval and consent to participate The research was conducted using data available in the public domain and did not include any human participants or animals. Therefore, no ethical approvals were required.

Consent for publication The authors of the study have consent and responsibility for submission to the journal.

Patient consent statement Not Applicable.

Permission to reproduce material from other sources All the data used in this study are available in public domain. The geographical maps in the manuscript are drawn using GADM in Python and does not reflect author's views on geographical boundaries of the represented countries.

References

1. Sung H, Ferlay J, Siegel RL et al (2021) Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 71(3):209–249
2. Adler I (1912) Primary malignant growths of the lungs and bronchi: A pathological and clinical study. Longmans, Green.
3. Doll R, Hill AB (1954) The mortality of doctors in relation to their smoking habits. BMJ 1(4877):1451
4. Terry LL (1964) Washington, DC: U.S. Department of Health, Education, and Welfare; 1964. Smoking and Health: Report of the Advisory Committee to the Surgeon General of the Public Health Service. In: U-23 U.S. Department of Health, Education, and Welfare. Public Health Service; Publication No. 1103
5. Deng Y, Zhao P, Zhou L et al (2020) Epidemiological trends of tracheal, bronchus, and lung cancer at the global, regional, and national levels: a population-based study. J Hematol Oncol 13:98
6. Allemani C, Matsuda T, Carlo V et al (2018) Global surveillance of trends in cancer survival 2000–14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. Lancet 391(10125):1023–1075

7. Gridelli C, Rossi A, Carbone D et al (2015) Non-small-cell lung cancer. Nat Rev Dis Primers 1:15009
8. American Cancer Society (2021). https://www.cancer.org/cancer/lung-cancer/about/what-is.html. Accessed 15 June 2021
9. Vikis HG, Wang D, Lu Y et al (2008) Familial aggregation of common sequence variants on 15q24–251 in lung cancer. J Natl Cancer Inst 100:1326–1330
10. Jonsson S, Thorsteinsdottir U, Gudbjartsson DF et al (2004) Familial risk of lung carcinoma in the icelandic population. JAMA 292:2977–2983
11. Sasco AJ, Secretan MB, Straif K (2004) Tobacco smoking and cancer: a brief review of recent epidemiological evidence. Lung Cancer 45:S3-9
12. Turner MC, Andersen ZJ, Baccarelli A et al (2020) Outdoor air pollution and cancer: An overview of the current evidence and public health recommendations. CA Cancer J Clin 70(6):460–479
13. International Agency for Research on Cancer (2002) Tobacco smoke and involuntary smoking. IARC, Lyon
14. Office on Smoking and Health (US) (2006) The health consequences of involuntary exposure to tobacco smoke: a report of the surgeon general [internet]. Centers for Disease Control and Prevention, Atlanta
15. de Groot PM, Wu CC, Carter BW et al (2018) The epidemiology of lung cancer. Transl Lung Cancer Res 7(3):220
16. American Cancer Society (ACS) (2021) The Global Economic Cost of Cancer. http://phrma-docs.phrma.org/sites/default/files/pdf/08-17-2010_economic_impact_study.pdf. Accessed 10 Oct 2021
17. Kutikova L, Bowman L, Chang S et al (2005) The economic burden of lung cancer and the associated costs of treatment failure in the United States. Lung Cancer 50(2):143–154
18. Ferlay J, Colombet M, Soerjomataram I et al (2020) Cancer statistics for the year 2020: an overview. Int J Cancer. https://doi.org/10.1002/ijc.33588
19. Ferlay J, Ervik M, Lam F et al (2020) Global cancer observatory: cancer today. International Agency for Research on Cancer, Lyon
20. United Nations Development Program (2020) Human development database. http://hdr.undp.org/en/data#. Accessed 7 Apr 2021
21. World Health Organisation (2021). https://apps.who.int/gho/data/node.sdg.3-a-viz?lang=en. Accessed 16 Jun 2021
22. Bray FI, Weiderpass E (2010) Lung cancer mortality trends in 36 European countries: secular trends and birth cohort patterns by sex and region 1970–2007. Int J Cancer 126(6):1454–1466
23. Siegel RL, Miller KD, Jemal A (2020) Cancer statistics, 2020. CA A Cancer J Clin 70:7–30. https://doi.org/10.3322/caac.21590
24. National Lung Screening Trial Research Team, Aberle DR, Adams AM et al (2011) Reduced lung-cancer mortality with low-dose computed tomographic screening. N Engl J Med 365:395–409
25. Pastorino U, Silva M, Sestini S et al (2019) Prolonged lung cancer screening reduced 10-year mortality in the MILD trial: new confirmation of lung cancer screening efficacy. Ann Oncol 30:1162–1169
26. Ruano-Ravina A, Heleno B, Fernández-Villar A (2015) Lung cancer screening with low-dose CT (LDCT), or when a public health intervention is beyond the patient's benefit. J Epidemiol Commun Health 69:99–100
27. Kramer BS, Berg CD, Aberle DR et al (2011) Lung cancer screening with low-dose helical CT: results from the National Lung Screening Trial (NLST). J Med Screen 18:109–111
28. Patz EF et al (2014) Overdiagnosis in low-dose computed tomography screening for lung cancer. JAMA Intern Med 174:269–274
29. Moyer VA (2014) Screening for lung cancer: US Preventive Services Task Force recommendation statement. Ann Intern Med 160:330–338
30. Black WC, Gareen IF, Soneji SS et al (2014) Cost-effectiveness of CT screening in the National Lung Screening Trial. N Engl J Med 371:1793–1802
31. Peprah E, Armstrong-Hough M, Cook SH et al (2021) An emerging syndemic of smoking and cardiopulmonary diseases in people living with HIV in Africa. Int J Environ Res Public Health 18(6):3111
32. Veeranki SP, John RM, Ibrahim A et al (2017) Age of smoking initiation among adolescents in Africa. Int J Public Health 62(1):63–72
33. Health Effects Institute (2020) State of Global Air 2020. Data source: Global Burden of Disease Study 2019. IHME, Washington
34. Clow B, Pederson A, Haworth-Brockman M et al (2009) Rising to the challenge: sex- and gender-based analysis for health planning, policy and research in Canada. Atlantic Centre of Excellence for Women's Health, Halifax
35. Lopez AD, Collishaw NE, Piha T (1994) A descriptive model of the cigarette epidemic in developed countries. Tob Control 3(3):242–247. https://doi.org/10.1136/tc.3.3.242
36. Zheng W, McLerran DF, Rolland BA et al (2014) Burden of total and cause-specific mortality related to tobacco smoking among adults aged ≥45 years in Asia: a pooled analysis of 21 cohorts. PLoS Med 11(4):e1001631
37. Yang JJ, Yu D, Wen W et al (2019) Tobacco smoking and mortality in Asia: a pooled meta-analysis. JAMA Netw Open 2(3):e191474
38. Yang D, Liu Y, Bai C et al (2020) Epidemiology of lung cancer and lung cancer screening programs in China and the United States. Cancer Lett 468:82–87
39. Zhou F, Zhou C (2018) Lung cancer in never smokers—the East Asian experience. Transl Lung Cancer Res 7(4):450
40. Jemal A, Miller KD, Ma J et al (2018) Higher lung cancer incidence in young women than young men in the United States. N Engl J Med 378:1999–2009
41. Loomis D, Grosse Y, Lauby-Secretan B et al (2013) The carcinogenicity of outdoor air pollution. Lancet Oncol 14(13):1262
42. Beelen R, Hoek G, Brandt PA et al (2008) Long-term exposure to traffic-related air pollution and lung cancer risk. Epidemiol 19(5):702–710
43. Raaschou-Nielsen O, Bak H, Sørensen M et al (2010) Air pollution from traffic and risk for lung cancer in three Danish cohorts. Cancer Epidemiol Biomarkers Prev 19(5):1284–1291
44. Huang F, Pan B, Wu J et al (2017) Relationship between exposure to PM2.5 and lung cancer incidence and mortality: a meta-analysis. Oncotarget 8:43322–43331

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Measuring Human Auditory Attention with EEG

Chirag Ahuja
Computer Science
Delhi Technological University
Delhi, India
cahuja1992@gmail.com

Divyashikha Setia
Computer Science
Delhi Technological University
Delhi, India
divyashikha@dtu.ac.in

***Abstract*—Human attention has been a complex topic of study for decades to understand cognitive ability. The electroencephalogram (EEG) signals act as an essential tool for understanding human cognitive ability and remain an active part of the study in clinical psychology. This paper attempts to improve upon the current research on selective auditory attention while considering the task of speech-to-text annotation along with the EEG signals. It proposes a new model to measure auditory attention using the 14 channel EEG signals from the Physiology of Auditory Attention (PhyAat) dataset. Using EEGNet and other classical machine learning techniques, a thorough study shows the improvement in the attention prediction task for test MAE of 29.65 to 22.47 on a single subject of the PhyAat dataset. This paper also demonstrates the effectiveness of EEG for auditory attention measurement across different subjects while achieving a Mean Absolute Error (MAE) of 31.47 overall and 28.83 for Subject 1 using the same model on the same test splits defined by the PhyAat dataset.**

***Index Terms*—Spectral Analysis, EEGNet, Shallow ConvNet, Deep ConvNet, Auditory Attention, Speech to Text, LightGBM, TSFresh, Laplacian Score**

I. Introduction

Auditory signals are one of humankind's most acute senses, demanding the simultaneous use of several cognitive capacities. To follow the context of a story, one must employ prior information and linguistic comprehension to create it. The growing presence of AI technology in the workplace has also increased the use of speech-to-text annotation. To do this, the annotator must listen to an audio clip and ignore semantic mistakes and noise while retaining the language. According Sohlberg and Mateer [16], the human attention is categorized into arousal, focused, sustained, selective, alternating, and divided. Auditory attention requires selective attention to filter out the noise, focused attention to comprehend and memorize the language, and sustained attention to listening to the audio.

This paper aims to analyze selective human attention using the speech-to-text task, with the use of both classical machine lerning and deep learning methodologies on Phyaat dataset. It helps us understand the relationship between human attention while listening to an audio clip, for example, a simulation of human-to-human conversation. Speech is the most used and easily interpretative mode of communication, depending on the language used. Although while listening to other people, the human brain considers many internal and external factors, which makes it the area of interest to understand the governing factors. With a focus on studying auditory attention, the most crucial physiological marker for brain signals is EEG. Hence, the paper focuses on the analysis of EEG signals for measuring Auditory Attentions.

Several datasets are available with this aim, such as Mousavi et al. [13] and Segalowitz and Barnes [14] discuss an oddball auditory paradigm where the response of participants to oddball sounds inserted in streams of sinusoidal tones. Auditory oddball paradigm dataset Segalowitz and Barnes [14] is based on the BCI three-stimulus paradigm. This oddball is identical to the normal oddball, except it has a different target: two target stimuli, one frequent stimulus. This research sought to give communication with binary options (yes/no), independent of vision, and has applications in the identification of Amyotrophic lateral sclerosis (ALS) or Locked-In Syndrome (LIS). Thus the author assessed three versions of the task in which the target stimuli changed in loudness, pitch, or direction.

Human stimuli on the speech to text task, with different semanticity of transcription for measuring the attention, is available as part of PhyAat: Physiology of Auditory Attention to Speech dataset proposed by Bajaj et al. [2]. The experiment involves transcribing one auditory stimulus for each of the 144 trials of the PhyAat. The dataset consists of three listening, writing, and resting in the trial, which tries to study the link between audio signals that humans hear and the human brain's processing. This collection seems to be the most appropriate for our research. For subject 1, the study also provides an early investigation of attention prediction using EEG. SVR, coupled with spectral characteristics, achieved a mean absolute error of 29.65 on the test split. However, the authors do not consider the efficacy of other characteristics in temporal, spectral, and wavelet domains or the resilience of algorithms over numerous subjects. The work did not consider the cutting-edge deep learning models, which can also capture interactions across different channels.

The **contributions** in this paper are summarized as follows:

- Extension over the original work of Bajaj et al. [2], while moving beyond spectral features. Using different time-domain statistics, frequency domain features (Fourier and Wavelet). Then, select the most important features using the Laplacian Score.
- Demonstration of the predictive power of the EEG Signals for Attention measurement across different subjects, rather than a single subject.

| //$31.00 ©10.1109/COMSNETS53615.2022.9668363

Authorized licensed use limited to: DELHI TECHNOLOGICAL UNIV. Downloaded on January 25,2022 at 08:40:31 UTC from IEEE Xplore. Restrictions apply.

- Comparison between three different CNN architectures for multichannel EEG, including Deep ConvNets, Shallow ConvNets, and EEGNet. Although the use of CNN-based models for EEG data is common and there is a multitude of work in this area, no results were reported with CNN-based models for the PhyAat dataset.
- Improvement of MAE by 7.18 for the reported Subject in the original work [2].

The proposed modelling methodology achieves a test MAE of 22.47 while training and testing for Subject 1 and a test MAE of 31.47 while training and testing across all 25 subjects. Also, while training on 25 Subjects and testing on Subject 1 yields the MAE of 28.83.

The outline of the rest of this papers is as follows: Section II details the statistical analysis of the data set used. Section III introduces the experimental process. Section IV presents the results of the experiments conducted in section III. Section V is the conclusion and prospect of the problem.

II. Dataset

The initial PhyAat dataset [2] has 19 feature time-series datasets, of which 14 are EEG channels, two are GSR dimensions, three are PPG measurements, and one is BPM (BPM). The dataset includes 14 channels of EEG readings collected for the 10/20 EEG placement system and captured at 128 Hz. Twenty-five university students participated in this research, including 21 men and four women, one left-handed and 24 right-handed. The model uses just the first 14 EEG channels for auditory attention. The subjects were native English speakers who did not have a known auditory processing issue in all cases. The participants were all between the ages of 16 and 34. Every subject receives a random assortment of 144 stimuli that do not get selected from the existing pool. Among the 144 people used, the semantic group consists 72, and the non-semantic group consists other 72. Each group gets split into six distinct noise levels, with 12 subjects per level. In addition, the 12 stimulus durations varied within each noise level. The specific stimuli ranged from four-word (L1) through eight-word (L2) and twelve words (L3) phrases (L3). The correlation between attentiveness and background noise is significant (r = 0.6) but not between semanticity or stimulus duration. The work presents all relationships to be statistically significant with $p < 0.0001$ using analysis of variance (ANOVA). Attention was affected by the auditory environment, as shown in the p-values (p = 0.001). As stimuli are directly proportional to audio length, therefore, The number of samples for a signal sampled at 128 Hz that the third quantile is around 500 samples. To make all the sequences of sample length, the model chose examples based on Q3 + 1.5 * IQR, which comes out to be 824 samples, and did post truncation and pre-padding of zeros.

III. Methodology and It's Explainability

Methods based on physiological signals for attention measurement are popular to understand human stimuli on different actions. Machine learning gets actively applied across multiple fields involving time-series data. Classical machine learning pipeline involves preprocessing, feature engineering, and modeling, whereas Deep Learning methods require minimal preprocessing. Fig. 1 shows the difference between the classical modeling pipeline and the deep learning pipeline. The separate feature extraction and selection step play a critical role, whereas the Deep Learning pipeline learns to extract the features. This section focuses on the effectiveness of different steps along the classical modeling pipeline and the deep learning architectures both, shallow and deep.

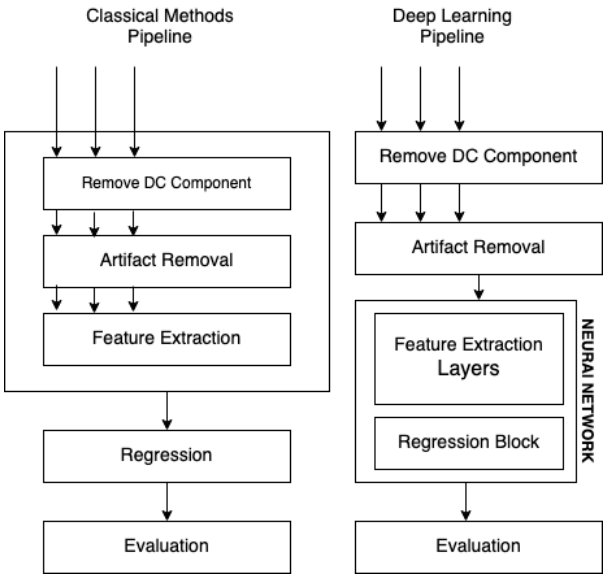


Fig. 1. High level overview of the complete processing in Traditional Methods vs. Deep Learning Methods

A. Pre-processing :

As part of standard practice in any signal processing activity, removal of Direct Current (DC) components is usually the first step to remove the trivial parts of the signal. Therefore the model uses a Fifth Order Butterworth filter. Unlike other filters, the Butterworth filter [9] has a maximally flat response, which means there is no passband ripple. As long as $j\omega = 0$, the filter is "flat maximally magnitude," and the first 2N - 1 derivatives of the transfer function are equal to zero, the filter is "flat maximally magnitude." This filter is completely defined mathematically by two parameters; they are cut off frequency and number of poles. Specific to EEG, artifacts are introduced because of eye blinks, muscle movements while trying, and pressing mouse buttons. Therefore artifact removal is a necessary preprocessing step for EEG Signals. Although the signal is multivariate, the framework processes each channel separately without considering the coherence of signals amongst each other. There are different steps in the complete artifact removal process as described in Fig. 2. The entire pipeline comprises different stages of decomposition that take care of artifact removal. Artifact removal is similar to

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:40:31 UTC from IEEE Xplore. Restrictions apply.

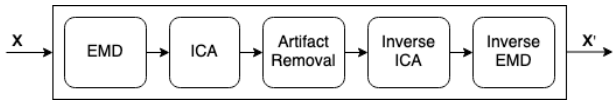


Fig. 2. Artifact Removal Pipeline, takes input signal with artifact as (x), passes through the pipeline and returns the cleaned signal x'

TABLE I
COMPARISON OF FEATURE SELECTION METHODS WITH SVR AS REGRESSOR

FeatureSelector	MAE(Validation)
Correlation	48.87
Mutual Information	48.43
RandomForest	45.76
Boruta	44.89
LaplacianSore	42.76

the cocktail party problem. Therefore the framework evaluates different techniques for Blind Source Separation (BSS). The paper compares Independent Component Analysis (ICA) and its other implementations, such as FastICA [1], InfoMax [12], Extended-Infomax [12]. However, initial experiments show the effectiveness of ICA on Intrinsic Mode Functions (IMF) better than the raw signal. Therefore, artifact removal does ICA on every IMF component obtained using Empirical mode decomposition (EMD). Using SVR with default parameters, the framework tune and select amongst the different artifact removal techniques. Each method has its corresponding set of Hyper Parameters like the Number of Independent Components (IC) and Window Size to process. The Hyper Parameter Optimizer (HPO) tunes independent component frequency and amplitude, i.e., the threshold of kurtosis and correlation coefficient. After detecting the artifacts, the cubic spline interpolates the artifacts. Finally, upon cleaning all the channels independently, the cleaned signals go through the modeling pipeline of classical models and deep learning models.

B. Classical Modelling Pipeline

1) Feature Extraction and Selection: Traditional machine learning methods can not directly work with the raw signals, making feature extraction a hard requirement. The framework extracts different features by considering the whole signal or as rolling window statistics with finite windows size. It is done for a complete trial rather than a segment of a trial. One set of features that the original PhyAat uses belongs to spectral bands, alpha (8–14 Hz), beta (14–30 Hz), low gamma (30–47 Hz), high gamma (47–64 Hz), delta (0.5–4 Hz) and theta (4–8 Hz). Apart from these features, the paper also evaluates time-domain statistical, frequency domain, and time-frequency domain, and some other features that are part of the TSFresh package [4]. As TSFresh yields hundreds of features, it is important to reduce the feature space. This process is called feature extraction and feature selection.

There are multiple ways of selecting features, such as Statistical Measures like Correlation and Mutual Information. Supervised learning approaches like Lasso, RandomForest, or Boruta Kursa et al.[10]. Unsupervised approach like Principal Component Analysis (PCA), Graph Spectral Clustering, Laplacian Score He et al. [7]. Unsupervised Learning techniques model the latent information much nicely as compared to Statistical and Supervised Learning techniques and are evident from Table I. Hence the model uses the Laplacian Score measure for the feature importance. Algorithm 1 represents the implementation of the Laplacian score for feature selection, where a lower score signifies higher feature importance.

Algorithm 1: Select important features using Laplacian Score

```
Result: Laplacian Score for each feature
Initialization: Construct the affinity matrix W using
  the k-Nearest Neighbours (kNN) approach, with
  weights as cosine distance between the node i and j;
scores = [] ;
for rth feature do
    fr = X(:,r);
    D = diag(W*ones);
    ones = [1,...,1];
    L = D - W;
    frHat = fr - (fr'*D*ones)*ones/(ones'*D*ones);
    score[r] = (frHat'*L*frHat)/(frHat'*D*frHat);
end
sortedFeatures = features[argsort(scores)];
```

The framework searches for the optimal number of features (k), as a hyperparameter optimization problem only, with search space ranging in percentage form (10% to 80%).

2) Classical Models: There is a vast range of data modeling algorithms for regression tasks, with corresponding hyperparameters tuned. Therefore, the whole framework follows a pipeline design pattern to add transformations, feature extraction, and selection, followed by models like (Linear, Random Forest, LightGBM, and Support Vector Regressors into a pipeline [8] [5]. Which uses Bayesian optimization with Gaussian Process as a surrogate function using the Facebook ax framework [6]. Which uses cross-validation error as shown in Fig. 3 as the objective function.

C. Deep Learning Models

Deep learning models have proven effectiveness in predictive modeling across different domains. Therefore, it can go beyond the classical modeling pipeline, with no feature engineering even with the physiological dataset. This paper explores both deep and shallow architectures along with the famous EEGNet Architecture proposed by Lawhern et al. [11]. For understanding the effectiveness of EEGNet that contains only 1,890 parameters due to depthwise separable convolution layers, this paper compares it with a shallow convolution network with 30,928 parameters and a deep convolution network with 158,232 parameters.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:40:31 UTC from IEEE Xplore. Restrictions apply.

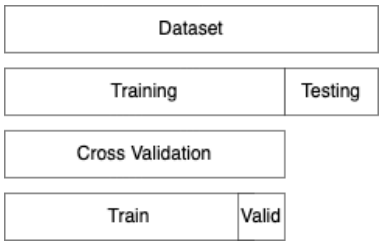


Fig. 3. Validation Strategy for Hyperparameter tunning and final evaluation on the test data

D. Evaluation

The machine learning model must work well on unseen data, which also generalizes well. The framework evaluates the model on scenarios that might come in the future. Therefore, it takes into account the following things:

- Designing the cross-validation sample when tuning the model and
- Designing the holdout sample when assessing the model's performance on unseen data.

Fig. 3 shows that the evaluation strategy comprises of two levels. The first level is the train-test split, where the test split generates benchmarking results. Hyperparameter tuning uses the second level split within the train.

E. Optimal Pipeline Search

The framework uses function approximation set up to find the optimal pipeline over the hyperparameter space (transformations, feature extraction & selection, model tuning, and selection) space; it can learn to fit the function using some surrogate model Gaussian Process. However, Sequential Module Based optimization (SMBO) using Gaussian Surrogate is a good way of reaching the optimal pipeline but is slow and overfits quickly. Therefore, went ahead with another surrogate in this Bayesian Optimization setup, i.e., Tree Parzen Optimizer (TPE), which is available as a part of HyperOpt Bergstra et al. [3]. Where the objective function is being computed over Mean Absolute Error (MAE) as $MAE(y, \hat{y}) = \frac{1}{m}\sum_{i=1}^{n} |y_i - \hat{y_i}|$. Although anything amongst mean square, mean absolute, or Huber error are valid errors to be used, MAE serves as a metric for comparing results with the original work of PhyAat.

IV. Results

This paper presents the experiments on PhyAat Task 1, with Subject 1 test MAE for attention score prediction as to the baseline. It includes all the subjects together for the training, i.e., 25 subjects x 143 trials. To avoid overfitting and tune the hyperparameters, the framework uses the cross-validation strategy shown in Fig. 3. Similar to PhyAat, post-cleaning, the framework extracts and selects features from every channel before fitting the regressor. Table II shows the results across two different feature extraction and selection strategies. Spectral feature refers to the six rhythmic used in the original work of PhyAat. Spectral features, time statistics, and CWT covers a wide variety of frequency, time, time-frequency domain features, followed by feature selection method compared in I. It tunes the cross-validation MAE and then compares it with the baseline testing MAE for selecting amongst different options available at every processing step.

TABLE II
Comparison of Traditional Approaches with EEGNet for Subject 1

Model	Features	Feature Selection	Test
SVR	Spectral Features	None	28.76
SVR	Spectral + Time Statistics + CWT	Laplacian Score	30.
RandomForest	Spectral Features	None	42.76
RandomForest	Spectral + Time Statistics + CWT	Laplacian Score	41.62
LightGBM	Spectral Features	None	28.35
LightGBM	Spectral + Time Statistics + CWT	Laplacian Score	28.65

TABLE III
Comparison of Deep Learning Architecture with and without artifact removal

Model	Subject(s)	Preprocessing	Validation MAE	Test MAE
DeepConvNet	All	Artifact Removal	25.67	35.83
DeepConvNet	All	Without Artifact Removal	28.98	37.58
ShallowConvNet	All	Artifact Removal	29.64	38.12
ShallowConvNet	All	Without Artifact Removal	31.89	40.76
EEGNet	All	Artifact Removal	20.73	31.47
EEGNet	All	Without Artifact Removal	24.74	34.23
DeepConvNet	Subject1	Artifact Removal	26.67	32.76
DeepConvNet	Subject1	Without Artifact Removal	27.65	33.54
ShallowConvNet	Subject1	Artifact Removal	32.64	42.73
ShallowConvNet	Subject1	Without Artifact Removal	34.89	46.84
EEGNet	Subject1	Artifact Removal	18.73	22.47
EEGNet	Subject1	Without Artifact Removal	21.76	26.98

Table II shows that classical models do not reduce the MAE score even after adding more expressive features from time and wavelet domain. However, theoretically, these features are more explicit. Deep Learning techniques result in better testing MAE because of a learning, rich set of feature space across multiple channels rather than concatenating features of every channel. Although EEGNet has fewer parameters to be fit, it still proves to be better than the shallow and deep convolution architectures with more parameters. It trains the model with MAE as a loss function and replacing the softmax layer with a dense layer. Facebook Ax finds the best fit model while minimizing for cross-validation MAE, using Bayesian Optimization and percentile strategy for early stopping. Table III shows that EEGNet has best test MAE.

V. Conclusion and Future Work

This paper looks into the human stimuli on speech-to-text tasks, with different semanticity of transcription for measuring

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:40:31 UTC from IEEE Xplore. Restrictions apply.

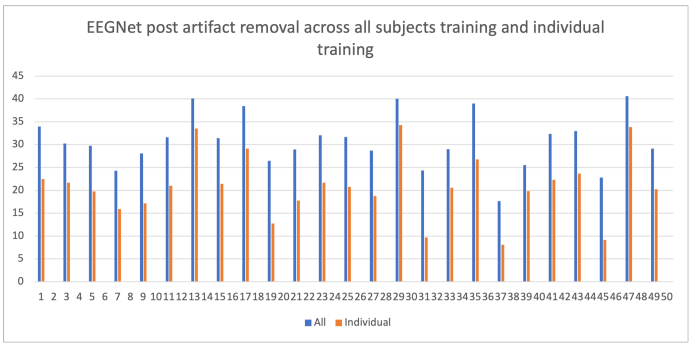


Fig. 4. EEGNet across all subjects training vs individual training

attention, using the PhyAat dataset. This paper demonstrates that EEG signals effectively quantify auditory attention by utilizing a dataset acquired via a speech-to-text annotation job. With LightGBM Regressor, strong feature engineering paired with hyperparameter tuning strategies such as Bayesian Optimization results in an MAE of 35.76 across all subjects. Deep learning models work better than classical approaches due to the limitations of capturing more expressive information and establishing relationships across many channels. EEGNet outperforms both deep and shallow networks across all individuals, with an overall MAE of 31.47. Additionally, the attention measuring job on audio signals generates signals with variable durations, requiring preprocessing to produce consistent dimensions. A simple technique such as post truncation and pre-padding, when utilized with EEGNet, functions optimally. The two modelling paradigms—classical signal processing with machine learning and other end-to-end deep learning provide a foundation for studying the effectiveness of EEG for human attention measurements. However, due to inter- and intra-subject variability, it is not applicable in the real world. Additionally, artifact removal is critical for both deep Learning and classical models, and therefore remains an area of continuous progress. We realize that utilizing different data augmentation methods would strengthen the model's resilience in the face of unknown circumstances. Explainable AI algorithms such as DeepLift [15] serves as a tool to explain the global and local significance of various channels and their features. We realize that utilizing different data augmentation methods would strengthen the model's resilience in the face of unknown circumstances.

REFERENCES

[1] P. Ablin, J.-F. Cardoso, and A. Gramfort. Faster independent component analysis by preconditioning with hessian approximations. *IEEE Transactions on Signal Processing*, 66(15):4040–4049, 2018.

[2] N. Bajaj, J. R. Carrión, and F. Bellotti. Phyaat: Physiology of auditory attention to speech dataset. *arXiv preprint arXiv:2005.11577*, 2020.

[3] J. Bergstra, D. Yamins, D. D. Cox, et al. Hyperopt: A python library for optimizing the hyperparameters of machine learning algorithms. In *Proceedings of the 12th Python in science conference*, volume 13, page 20. Citeseer, 2013.

[4] M. Christ, N. Braun, J. Neuffer, and A. W. Kempa-Liehr. Time series feature extraction on basis of scalable hypothesis tests (tsfresh–a python package). *Neurocomputing*, 307:72–77, 2018.

[5] H. Drucker, C. J. Burges, L. Kaufman, A. Smola, V. Vapnik, et al. Support vector regression machines. *Advances in neural information processing systems*, 9:155–161, 1997.

[6] Facebook. facebook/ax: Adaptive experimentation platform. URL `https://github.com/facebook/Ax`.

[7] X. He, D. Cai, and P. Niyogi. Laplacian score for feature selection. *Advances in neural information processing systems*, 18, 2005.

[8] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu. Lightgbm: A highly efficient gradient boosting decision tree. *Advances in neural information processing systems*, 30:3146–3154, 2017.

[9] M. Kumngern, N. Aupithak, F. Khateb, and T. Kulej. 0.5 v fifth-order butterworth low-pass filter using multiple-input ota for ecg applications. *Sensors*, 20(24):7343, 2020.

[10] M. B. Kursa, W. R. Rudnicki, et al. Feature selection with the boruta package. *J Stat Softw*, 36(11):1–13, 2010.

[11] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance. Eegnet: a compact convolutional neural network for eeg-based brain–computer interfaces. *Journal of neural engineering*, 15 (5):056013, 2018.

[12] T.-W. Lee, M. Girolami, and T. J. Sejnowski. Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources. *Neural computation*, 11(2):417–441, 1999.

[13] Z. Mousavi, M. M. Kiani, and H. Aghajan. Brain signatures of surprise in eeg and meg data. *bioRxiv*, 2020.

[14] S. J. Segalowitz and K. L. Barnes. The reliability of erp components in the auditory oddball paradigm. *Psychophysiology*, 30(5):451–459, 1993.

[15] A. Shrikumar, P. Greenside, and A. Kundaje. Learning important features through propagating activation differences. In *International Conference on Machine Learning*, pages 3145–3153. PMLR, 2017.

[16] M. M. Sohlberg and C. A. Mateer. *Cognitive rehabilitation: An integrative neuropsychological approach*. Guilford Press, 2001.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:40:31 UTC from IEEE Xplore. Restrictions apply.

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/357791872

Multi-functional Fractional-order Shadow Filter using OTA

Conference Paper · November 2021

CITATIONS
0

READS
35

3 authors:



Garima Varshney
Delhi Technological University
10 PUBLICATIONS **7** CITATIONS
SEE PROFILE



Neeta Pandey
Delhi Technological University
291 PUBLICATIONS **1,737** CITATIONS
SEE PROFILE



Rajeshwari Pandey
Delhi Technological University
102 PUBLICATIONS **802** CITATIONS
SEE PROFILE

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

Project AEÜ - International Journal of Electronics and Communications 2021 Best Paper Award View project

Project fractional order circuits View project

All content following this page was uploaded by Garima Varshney on 13 January 2022.

The user has requested enhancement of the downloaded file.

Multi-functional Fractional-order Shadow Filter using OTA

Garima Varshney
ECE deptt.
Delhi Technological University
Delhi, India
gvarshn1@gmail.com

Neeta Pandey
ECE deptt.
Delhi Technological University
Delhi, India
neetapandey@dce.ac.in

Rajeshwari Pandey
ECE deptt.
Delhi Technological University
Delhi, India
rpandey@dce.ac.in

***Abstract*—The realization of a new multi-functional shadow fractional-order filter (FOF) based upon operational transconductance amplifier (OTA) is proposed in this paper. The shadow FOF consists of a basic 2α FOF and an external amplifier with gain A. The filter parameters such as pole quality factor Q and pole frequency f_0 can be regulated with the help of external amplifier's gain, without changing the active or passive components of basic FOF. The proposed shadow FOF offers fractional-order high-pass filter (FHPF), fractional-order low-pass filter (FLPF) and fractional-order band-pass filter (FBPF). The three cases of stability with different feedback signals are discussed and simulation results are presented. The simulations are accomplished using SPICE with 180nm CMOS technology model parameters. Corner analysis and Monte-Carlo analysis are also executed to verify the robustness of the proposed shadow FOF.**

***Index Terms*—fractional-order filter, fractional-order shadow filter, pole frequency tuning, pole quality factor, OTA.**

I. Introduction

At present, the applications of fractional-order (FO) calculus are spreading at a high pace, due to its inherit ability of offering more precise interpretation of real objects [1]. Enormous possibilities of research lies in FO systems at interdisciplinary level because its applications are found in different branches of science, such as: electronics [2], biomedical [3], thermodynamics [4], electrical engineering [5] and many more.

Filters are dominant circuits in analog signal processing domain and the demand for active filters is increasing day by day in various fields of communication [6]. FOFs are wider form of conventional integer-order filters. FOF provides added control over transition band's slope, as slope is equal to -20*(order of FOF) dB/dec and the order of FOF depends upon α ($0<\alpha<1$), where α is the fractional part of the order of FOF [7].

In general, tuning of FOF parameters, such as pole frequency f_0 and pole quality factor Q, is necessary for several applications and can be achieved by tuning the constituent elements. The constituent elements can be either passive or active, in case of active elements, the electronic tuning of FOF parameters is possible through varying transconductance/current gain [8]. In the framework of this work, the FOFs with electronic tunability are compared on the basis of used active block(s), filter response(s), order of filter and parameter(s) tuned. The summarized comparison is listed in Table I and below are the findings:

- Current Feedback Operational Amplifier (CFOA) [12], Operational Transconductance Amplifier (OTA) [8] - [11], [13] & [16], Multiple Output-Current Follower (MO-CF) [14], Adjustable Current Amplifier (ACA) [9], [14], Fully Differential-Current Follower (FD-CF) [15] based electronically tunable FOFs are reported in respective references.
- Lean literature is accessible for the regulation of Q in FO domain.
- The filter's parameters can be electronically tuned by deviating the current gain [9], [12] & [14] - [15], transconductance gain of the constituent active elements [8], [10] - [11], [13] & [15] or the gain of external amplifier [16].

Recently, the shadowing concept to tune FOF's parameters, without changing the active and passive components of the basic 2α FOF, using an external amplifier of gain A in the feedback loop has been reported in [16] for FO domain. The concept of shadowing in FO domain is used in this work to vary the parameters of FOF. A multi-functional basic 2α FOF is taken, which consist of three outputs such as FHPF, FBPF and FLPF, the f_0 and Q of FOF are varied using an external amplifier of gain A in the feedback loop. Three cases of stability, each for different case of feedback signal has been discussed and verified through SPICE simulation using 180nm MOS technology parameters.

The remaining paper is arranged as follows: Section 2 gives basic idea of shadow FOF and Section 3 deals with the proposed circuit. Section 4 talks about simulation results in detail followed by the concluding remarks in Section 5.

II. Basics of shadow FOF

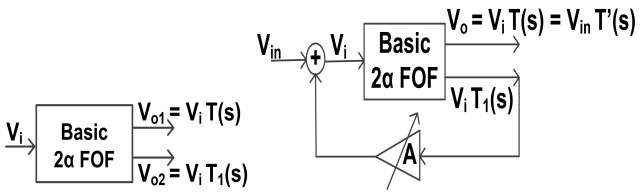
The basic scheme of shadow FOF is discussed in [16] and briefly outlined here for quick reference. Consider a basic 2α FOF is having single input multiple output (SIMO), appeared in Fig. 1(a), the transfer functions for two outputs are depicted as $T(s)$ and $T_1(s)$. Any one or combination of the outputs of basic 2α FOF can be given as the input to the external amplifier in the feedback loop to constitute a shadow FOF. The resulting shadow FOF's block diagram is appearing in Fig. 1(b). This feedback amends the transfer function $T(s) = V_0/V_i$ to $T'(s) = V_0/V_{in}$.

Let $T(s)$ and $T_1(s)$ are given by (1), by regular analysis of

978-1-6654-2690-9/21/ $31.00 © 2021 IEEE

TABLE I: COMPARISON OF ELECTRONICALLY TUNABLE FOFs

Ref.	Active block	Response	Filter's order	Electronic tunabillity	Type of tuning
[8]	OTA	FHPF, FAPF, FLPF	α	Yes	f_0
[9]	OTA, ACA	FHPF, FBPF, FLPF, FBSF	$(1+\alpha)$	Yes	f_0, Q
[10]	OTA	FBPF, FLPF	2α	Yes	f_0, f_{rp}
[11]	OTA	FHPF, FBPF, FLPF, FAPF, FBSF	$(\alpha+\beta)$	Yes	f_0, order
[12]	CFOA	FHPF, FBPF, FLPF, FBSF	$(1+\alpha)$	Yes	Order
[13]	OTA	FBPF	2α	Yes	Order
[14]	MO-CF, ACA	FLPF	$(1+\alpha)$	Yes	f_0, order
[15]	FD-CF, OTA	FLPF	$(1+\alpha)$	Yes	f_0, order
[16]	OTA	FBPF, FLPF	2α	Yes	f_0, Q
Proposed	OTA	FLPF, FBPF, FHPF	2α	Yes	f_0, Q



(a) (b)

Fig. 1: Block diagram of shadow FOF [16] (a) Basic 2α cell (b) Basic 2α cell with external amplifier in feedback loop

Fig. 1(b) the transfer function of shadow FOF (i.e., $T'(s)$) can be articulated as (2):

$$T(s) = \frac{Fs^{2\alpha} + Es^{\alpha} + G}{s^{2\alpha} + 2ps^{\alpha} + q};\quad T_1(s) = \frac{Hs^{2\alpha} + Bs^{\alpha} + L}{s^{2\alpha} + 2ps^{\alpha} + q} \tag{1}$$

$$T'(s) = \frac{T(s)}{1 - AT_1(s)} \tag{2}$$

By routine algebra, $T'(s)$ can be calculated as (3). The characteristic equation of $T'(s)$ is different from $T(s)$ and depends upon the feedback signal $T_1(s)$ and external amplifier's gain A. Therefore, shadow FOF's pole frequency and pole quality factor can be controlled by using different types of feedback signals $T_1(s)$ and with the help of external amplifier's gain A, thus facilitating electronic tunability to the filter characteristics without effecting its core.

$$T'(s) = \frac{N'(s)}{D'(s)} = \frac{(F's^{2\alpha} + E's^{\alpha} + G')}{s^{2\alpha} + 2Xs^{\alpha} + Y} \tag{3}$$

where

- $(F', E', G') = (F, E, G)/(1 - AH)$
- $X = \frac{2p - AB}{2(1-AH)}$
- $Y = \frac{q - AL}{1 - AH}$

For in-depth discussion on the variation of quality factor and pole frequency with respect to A for various cases of stability and different feedback signals, [16] can be referred. The three cases of stability [7], i.e. case-1 $(X^2 \geq Y, X > 0, Y > 0)$ for FHPF feedback signal, case-2 $(X^2 < Y, X > 0, Y > 0)$ for FLPF feedback signal and case-3 $(X^2 < Y, X < 0, Y > 0)$ for FBPF feedback signal are discussed in this work. For case-1 of FHPF feedback the value of quality factor is fixed while pole frequency is variable with A, for case-2 of FLPF feedback both pole frequency and quality factor are variable with A whereas, for case-3 of FBPF feedback the quality factor is variable with A while possessing fixed pole frequency.

III. Proposed Shadow FOF with HP, LP and BP responses

The proposed circuit of multi-functional shadow FOF is shown in Fig. 2, consist of OTA 1-4 (construct a basic 2α FOF) and an external amplifier with gain A (constructed using an adder circuit). The basic 2α FOF contains three outputs, such as: FHPF, FBPF and FLPF, the transfer function of the basic 2α order FOF is given in (4).

$$\frac{V_{HP;LP;BP}}{V_i} = \frac{s^{2\alpha}; s^{\alpha}\frac{g_{m1}}{C_{\alpha1}}; \frac{-g_{m1}g_{m2}}{C_{\alpha1}C_{\alpha2}}}{s^{2\alpha} + s^{\alpha}\frac{g_{m1}}{C_{\alpha1}} + \frac{g_{m1}g_{m2}}{C_{\alpha1}C_{\alpha2}}} \tag{4}$$

Any one of the outputs can derive the external amplifier, thus as per the type of feedback signal, there will be three types of transfer functions for the output of shadow FOF, as given in (5), (6) and (7) for FHPF, FBPF and FLPF signal as feedback signals respectively.

$$T_{(HP;LP;BP)_HP} = \frac{s^{2\alpha}; s^{\alpha}\frac{g_{m1}}{C_{\alpha1}}; \frac{-g_{m1}g_{m2}}{C_{\alpha1}C_{\alpha2}}}{(1-A)s^{2\alpha} + s^{\alpha}\frac{g_{m1}}{C_{\alpha1}} + \frac{g_{m1}g_{m2}}{C_{\alpha1}C_{\alpha2}}} \tag{5}$$

$$T_{(HP;LP;BP)_BP} = \frac{s^{2\alpha}; s^{\alpha}\frac{g_{m1}}{C_{\alpha1}}; \frac{-g_{m1}g_{m2}}{C_{\alpha1}C_{\alpha2}}}{s^{2\alpha} + s^{\alpha}\frac{(1-A)g_{m1}}{C_{\alpha1}} + \frac{g_{m1}g_{m2}}{C_{\alpha1}C_{\alpha2}}} \tag{6}$$

$$T_{(HP;LP;BP)_LP} = \frac{s^{2\alpha}; s^{\alpha}\frac{g_{m1}}{C_{\alpha1}}; \frac{-g_{m1}g_{m2}}{C_{\alpha1}C_{\alpha2}}}{s^{2\alpha} + s^{\alpha}\frac{g_{m1}}{C_{\alpha1}} + \frac{(1-A)g_{m1}g_{m2}}{C_{\alpha1}C_{\alpha2}}} \tag{7}$$

There can be nine combinations for feedback signals and the

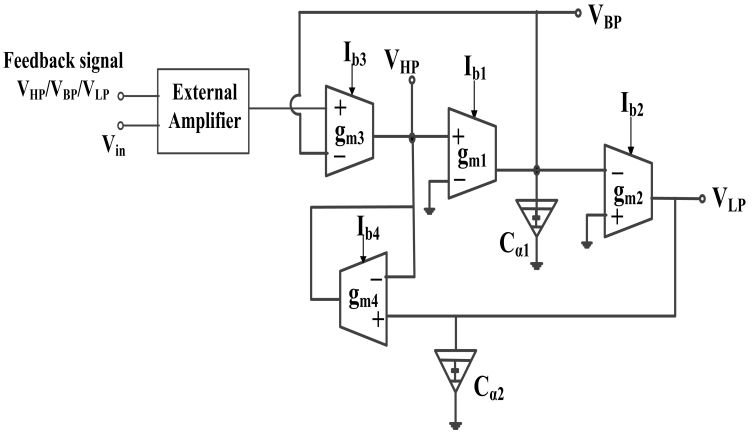


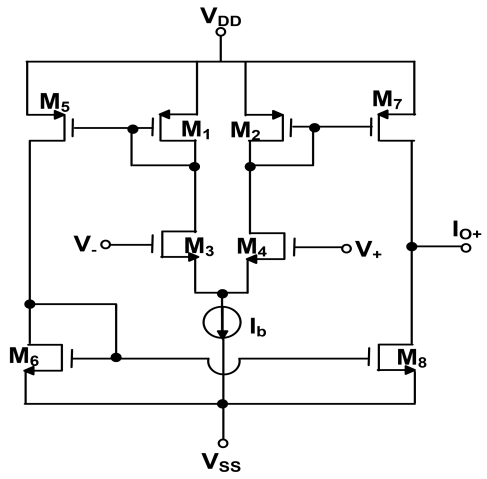
Fig. 2: Circuit diagram of proposed shadow FOF

cases of stability, such as: case-1, 2 & 3 for FHPF feedback signal; case-1, 2 & 3 for FBPF feedback signal and case-1, 2 & 3 for FLPF feedback signal. For the sake of brevity, three cases are chosen to present in this work, case-1 for FHPF feedback signal, case-2 for FLPF feedback signal and case-3 for FBPF feedback. By using the method outlined for stability constraints in [16], the values of g_{m1}, g_{m2} and A are calculated. The constraint on A & g_ms and the effect on f_0 and Q for various cases are given in Table II.

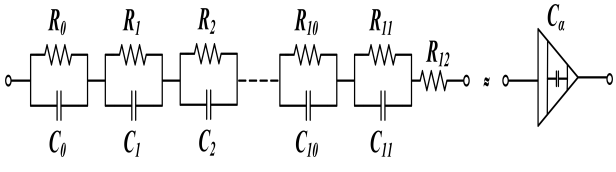
IV. SPICE Simulation Results

To illustrate the functionality of the proposed structure, SPICE is used with 180nm CMOS technology model parameters. The OTA provided in [17] is used, for quick reference Fig. 3(a) is showing CMOS structure of the OTA, the W/L ratios of the transistors M_{3-4} are taken as 5.76/0.72, $M_{1,2,5,7}$ are 2.16/0.72 and $M_{6,8}$ are 1.44/0.72. The supply voltages are used as $\pm$1.5 V. The proposed shadow FOF uses FOC, which is approximated via RC ladder network method outlined in [18]. The constructed RC network is appeared in Fig. 3(b) and passive components' values for $\alpha = 0.5$ with $C_\alpha = 3.75\mu\mho/s^\alpha$ can be referred from [18]. The approximated FOC is valid for 1Hz to 1MHz with central frequency of 1kHz. The values of $C_{\alpha 1}$ and $C_{\alpha 2}$, for all the three cases, are taken same as $3.75\mu\mho/s^\alpha$ with $\alpha = 0.5$.

To realize case-1 for FHPF feedback signal, the values of



(a)



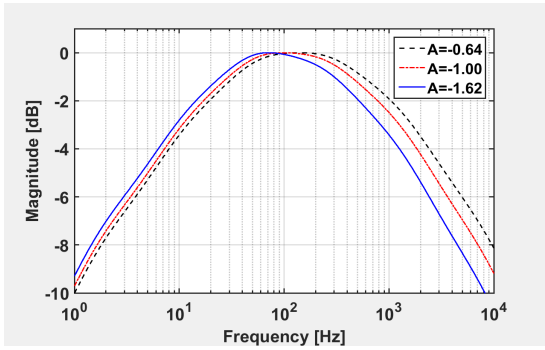
(b)

Fig. 3: Sructures used for simulation (a) MOS structure of OTA [17] (b) FOC constructed via RC network [18]

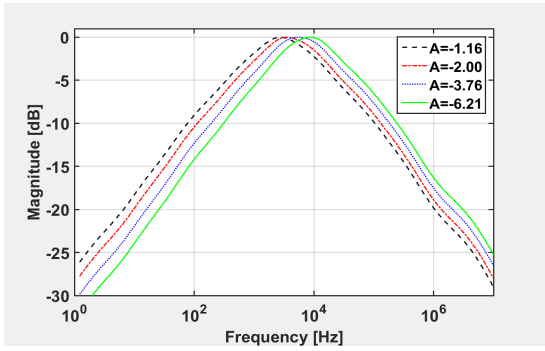
g_{m1} and g_{m2} are taken as $533.89\mu A/V$ and $40.07\mu A/V$ respectively. The gain A of external amplifier is set as -0.64, -1, and -1.62. The magnitude response of T_{BP_HP} is shown in Fig. 4(a). The value of Q is fixed at 0.085 for all the three values of A.

Further, to realize case-2 of FLPF feedback signal and case-3 of FBPF feedback signal, the values of g_{m1} and g_{m2} are taken as same and equal to $376.50\mu A/V$. The gain A of the external amplifier is set as -1.16, -2, -3.76, and -6.21 for the case of FLPF feedback signal while for FBPF feedback signal it is set as 1.62, 2, 2.12, and 2.45. The magnitude responses of T_{BP_LP} and T_{BP_BP} are shown in Figs. 4(b) & (c).

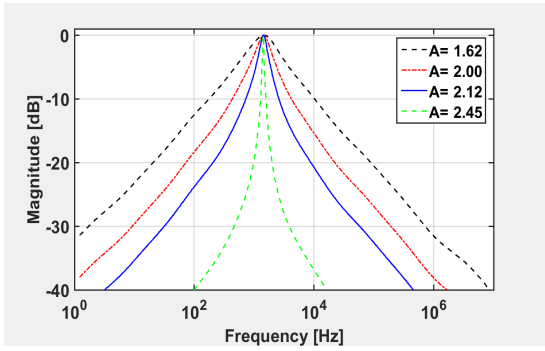
The summary of simulation results for all the above cases are enlisted in Table III. It is observed from the Table III and above discussion that for case-1 of FHPF feedback signal the pole frequency is variable with fixed Q while for case-2 of FLPF feedback signal the pole frequency and Q both are variable. For case-3 of FBPF feedback signal the pole frequency is fixed with variable Q which corroborates with the theory proposed in [16].



(a)



(b)



(c)

Fig. 4: FBPF Magnitude response for various A and feedback signal as (a) FHPF (case-1) (b) FLPF (case-2) (c) FBPF (case-3)

TABLE II: CONSTRAINTS ON g_m & A AND EFFECT ON f_0 & Q FOR DIFFERENT CASES OF STABILITY

Feedback signal	**Case**	**Constraint on g_m**	**Constraint on A**	**Effect on f_0 and Q**
FHPF	Case-1	$g_{m1} \geq 4(1-A)g_{m2}$	$A < 1$	Variable f_0, fixed Q
FLPF	Case-2	$g_{m1} < 4(1-A)g_{m2}$	$A < 1$	Variable f_0, variable Q
FBPF	Case-3	$(1-A)^2 g_{m1} < 4g_{m2}$	$A > 1$	Fixed f_0, variable Q

TABLE III: SUMMARIZED SIMULATION RESULTS

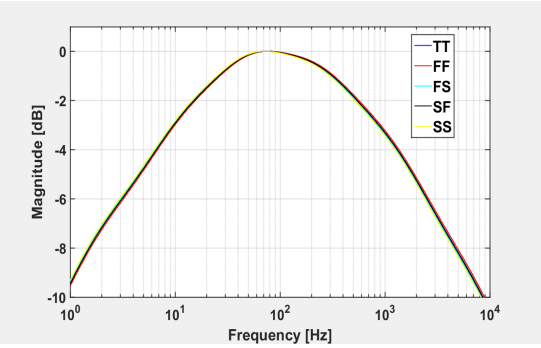
FHPF feedback signal (case-1)			**FLPF feedback signal (case-2)**			**FBPF feedback signal (case-3)**		
A	Q	f_0	A	Q	f_0	A	Q	f_0
-0.64	0.085	116.4Hz	-1.16	0.22	2.75kHz	1.62	0.36	1.50kHz
-1.00	0.085	89.1Hz	-2.00	0.23	3.45kHz	2.00	0.69	1.50kHz
-1.62	0.085	71.3Hz	-3.76	0.25	5.45kHz	2.12	1.39	1.50kHz
-	-	-	-6.21	0.29	8.80kHz	2.45	2.18	1.50kHz

Moreover to check the sensitiveness of the output of the proposed shadow FOF for the random variation in process corners during fabrication, the proposed circuit is examined at the five process corners i.e. fast fast (FF), slow slow (SS), slow fast (SF), fast slow (FS) and typical typical(TT). The magnitude response of FBPF for all the three cases i.e. FHPF feedback case-1 ($A = -1.62$), FLPF feedback case-2 ($A = -2$) and FBPF feedback case-3 ($A = 2$) are presented in Figs. 5(a), (b) and (c) respectively for all the five corners, those clearly ensure the robustness of the proposed shadow FOF.

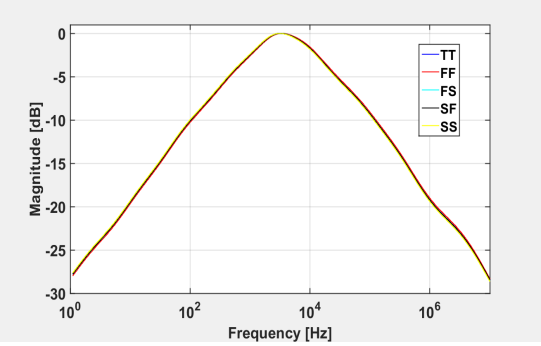
To get additional perception on the sensitivity of the proposed shadow FOF with respect to C_α, Monte-Carlo analysis is performed with uniform Gaussian distribution of passive components used to realize C_α with $\pm 10\%$ resistance and capacitance tolerances, on 500 random samples. The distribution of center frequency for 500 samples is shown in Figs. 6(a), (b) and (c) for FHPF feedback case-1 ($A = -1.62$), FLPF feedback case-2 ($A = -2$) and FBPF feedback case-3 ($A = 2$) respectively. The whole data range lies in between 4 times the standard deviation and the mean for all the cases matched with the corresponding center frequency within $\pm 10\%$ error, this shows the validity of proposed shadow FOF's for the random variation in $C'_\alpha s$ value.

V. Conclusion

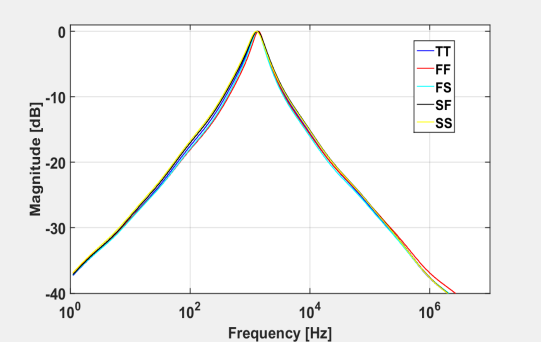
In this work, a multi-functional shadow FOF is presented and verified using SPICE with 180nm MOS technology parameters. The FOF parameters such as f_0 and Q are tuned with the help of external amplifier's gain, without changing the active or passive components of basic FOF. The magnitude response of FBPF for three cases of stability with different feedback signals are discussed and simulation results are presented. Case-1 of FHPF feedback ensures fixed Q and variable f_0 while case-2 of FLPF feedback provides Q and f_0 both variable with external amplifier's gain A. The case-3 of FBPF feedback guarantees fixed f_0 and variable Q with respect to A. The robustness of the circuit has also been validated using Monte-Carlo analysis and corner analysis.



(a)

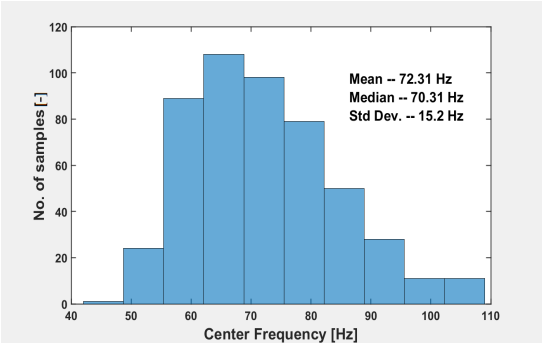


(b)

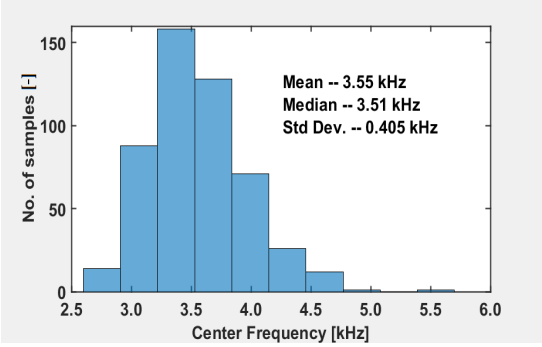


(c)

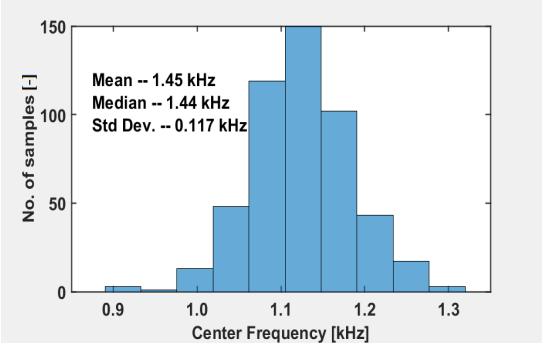
Fig. 5: Corner analysis for feedback signal as (a) FHPF (case-1, $A = -1.62$) (b) FLPF (case-2, $A = -2$) (c) FBPF (case-3, $A = 2$)



(a)



(b)



(c)

Fig. 6: Histogram for the variation of C_α on 500 samples, for feedback signal as (a) FHPF (case-1, $A = -1.62$) (b) FLPF (case-2, $A = -2$) (c) FBPF (case-3, $A = 2$)

REFERENCES

[1] Y.Chen, I.Petras and D.Xue, "Fractional order control - A tutorial" 2009 American Control Conf, pp. 1397-1411, 2009, doi:10.1109/ACC.2009.5160719

[2] H.Sheng, Y.Q.Chen and T.S.Qiu, "Fractional Processes and Fractional-Order Signal Processing: Techniques and Applications," Springer, New York, 2011.

[3] L.A.Geddes and L.E.Baker, "Principles of Applied Biomedical Instrumentation," 3rd edn. Wiely, New York. 1989.

[4] A.S.Elwakil, "Fractional-order circuits and systems: an emerging Interdisciplinary research area," IEEE Circ Syst Mag, vol. 10, no. 4, pp. 40–50, 2010.

[5] S.Das and I.Pan, "Fractional order signal processing—introductory concepts and applications," In: Springer Briefs in App. Sci and Tech., 2012.

[6] R.Kulkarni, J.Kim, H.Jeon, J.Xiao and J.Silva-Martinez,"UHF Receiver Front-End: Implementation and Analog Baseband Design Considerations," in IEEE Transactions on VLSI Systems, vol. 20, no. 2, pp. 197-210, Feb. 2012.

[7] A.G.Radwan, A.S.Elwakil and A.M.Soliman,"On the Generalization of second order filters to the fractional order domain," J of Circ, Syst and Computers, vol. 18, no. 2, pp. 361-386, 2009.

[8] G.Varshney, N.Pandey and R.Pandey, "Electronically tunable multifunction transadmittance-mode fractional-order filter," Arab J Sci Eng, vol. 46, pp. 1067–1078, 2021.

[9] L.Langhammer, R.Sotner, J.Dvorak, T.Dostal, "Fully-differential multifunctional electronically configurable fractional-order filter with electronically adjustable parameters," Elektronika IR Elektrotechnika, vol. 24, no. 5, pp. 42-45, 2018.

[10] R.Verma, N.Pandey, R.Pandey, "Electronically tunable fractional order filter," Arab J Sci & Eng, vol. 42, pp. 3409–3422, 2017.

[11] G.Tsirimokou, C.Psychalinos, A.S.Elwakil, "Fractional-order electronically controlled generalized filters," Int J of Circ Theo and Appl, vol. 45, no. 5, pp. 595–612, 2017.

[12] G.Tsirimokou, R.Sotner, J.Jerabek, J.Koton, C.Psychalinos, "Programmable analog array of fractional-order filters with CFOAs," in Proc. 40th Int. Conf. Telecommunications and Signal Processing (TSP 2017), Barcelona, Spain, pp. 706–709, 2017.

[13] O.I.Ahmed, H.M.Yassin, A.Said, C.Psychalinos, A.G.Radwan, "Implementation and analysis of tunable fractional-order band-pass filter of order 2α," AEU – Int J of Electron. and Commun, vol.124, 153343, 2020.

[14] J.Dvorak, L.Langhammer, J.Jerabek, J.Koton, R.Sotner, J.Polak, "Electronically tunable fractional-order low-pass filter with current followers," In: 2016 39th Int conf on Telecommunications and signal processing (TSP). IEEE, pp. 587–592, 2016, doi:10.1109/TSP.2016.7760949.

[15] L.Langhammer, R.Sotner, J.Dvorak, J.Jerabek, J.Polak, "Fully-differential tunable fractional-order filter with current followers and current amplifiers," 27th Int Conf Radioelektronika (RADIOELEKTRONIKA), IEEE, pp. 1–6, 2017, doi:10.1109/RADIOELEK.2017.7937576.

[16] G.Varshney, N.Pandey and R.Pandey, "Generalization of shadow filters in fractional domain," Int J Circ Theor Appl., pp. 1-18, May 2021.

[17] J.Ramirez-Angulo, M.Robinson and E.Sanchez-Sinecio, "Current modde continuous-time filters:Two design approaches," IEEE Tran on Circ and Sys II: Analog and Digital Signal Processing, vol. 39, no. 6, pp. 337-341, June 1992.

[18] A.Adhikary, P.Sen, S.Sen and K.Biswas, "Design and performance study of dynamic fractors in any of the four quadrants," Circ, Sys and Sig Proc, vol. 35, no. 6, pp. 1909-1932, 2016.

View publication stats

Hindawi
Bioinorganic Chemistry and Applications
Volume 2021, Article ID 1764647, 16 pages
https://doi.org/10.1155/2021/1764647

Review Article

Nanomaterials for Remediation of Environmental Pollutants

Arpita Roy,[1] **Apoorva Sharma**,[2] **Saanya Yadav**,[2] **Leta Tesfaye Jule**,[3,4] **and Ramaswamy Krishnaraj**[3,5]

[1]*Department of Biotechnology, School of Engineering & Technology, Sharda University, Greater Noida, India*
[2]*Department of Biotechnology, Delhi Technological University, Delhi, India*
[3]*Centre for Excellence-Indigenous Knowledge, Innovative Technology Transfer and Entrepreneurship, Dambi Dollo University, Dembi Dolo, Ethiopia*
[4]*Department of Physics, College of Natural and Computational Science, Dambi Dollo University, Dembi Dolo, Ethiopia*
[5]*Department of Mechanical Engineering, Dambi Dollo University, Dembi Dolo, Ethiopia*

Correspondence should be addressed to Arpita Roy; arbt2014@gmail.com and Ramaswamy Krishnaraj; dr.krishnarajdirectorcei@dadu.edu.et

Received 11 October 2021; Revised 11 December 2021; Accepted 14 December 2021; Published 28 December 2021

Academic Editor: Wilson Aruni

Copyright © 2021 Arpita Roy 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.

Today, environmental contamination is a big concern for both developing and developed countries. The primary sources of contamination of land, water, and air are extensive industrialization and intense agricultural activities. Various traditional methods are available for the treatment of different pollutants in the environment, but all have some limitations. Due to this, an alternative method is required which is effective and less toxic and provides better outcomes. Nanomaterials have attracted a lot of interest in terms of environmental remediation. Because of their huge surface area and related high reactivity, nanomaterials perform better in environmental clean-up than other conventional approaches. They can be modified for specific uses to provide novel features. Due to the large surface-area-to-volume ratio and the presence of a larger number of reactive sites, nanoscale materials can be extremely reactive. These characteristics allow for higher interaction with contaminants, leading to a quick reduction of contaminant concentration. In the present review, an overview of different nanomaterials that are potential in the remediation of environmental pollutants has been discussed.

1. Introduction

The world is on the edge of a major environmental calamity that will cost us our fortune. The current state of the present environment is forever deteriorating. Environmental issues are piling up across the globe, and we have to behave as in an emergency on our planet. We have to gain a new perspective and approach calamities beforehand with new concepts and strategies and with our full awareness and seriousness. Pollution in nature, such as in air, water, and soil, undergoes millions of years to eradicate. Industry and automobile exhaust emissions are the main contributing factors for most of the environmental pollution [1].

Air pollutants like NO(x), SO_2, highly reactive and toxic organic compounds, POPs like dioxins, and PAH (polycyclic aromatic hydrocarbons) are all hazardous for mankind [2]. When inhaled in high quantities, carbon monoxide (CO) can harm by immediate poisoning. When certain heavy metals such as Pb are taken into a living organism, they can vandalize by either immediate poisoning or chronic poisoning, depending on the different levels of exposure [3].

Pulmonary issues such as COPD, asthma, and bronchiolitis, cancers, heart diseases, CNS dysfunctions, and skin disorders are among the ailments caused by the aforementioned chemicals [4]. For ages, methods like condensation, flocculation, froth floatation, sand filtration, and AC adsorption are being widely used. However, these have their own limitations, such as inefficiency in scarping metal ions, high-energy input, and production of nonreusable chemicals. To overcome these problems, nanotechnology seems like an emerging option [5].

Nanotechnology possesses a prospect to significantly bestow in the development of cleaner, greener technologies that have major environmental and health advantages [6].

Nanotechnology techniques are being investigated for their potential to provide solutions for pollution management and mitigation, as well as to improve the performance of traditional environmental clean-up methods [7]. Nanotechnology can help the environment by reducing energy uptake during the manufacturing and production pathways, allowing products to get recycled after their usage, and developing and using environmentally friendly materials [8]. Nanotechnology presently holds great promise for addressing sustainability issues, but at the same time, we should also take into account the harm to the environment and human health [9]. Therefore, in this review, the importance of various nanomaterials that are capable of removing pollutants has been discussed.

2. Different Types of Pollutants

Pollutants are particles that cause damage to the environment by polluting them. When exposed to these pollutants, life can be damaged, and consequences on humans and other organisms are well recognized. Pollutants can get into the environment in a number of different ways, both naturally and through humankind [10]. The different kinds of pollutants include dyes, heavy metals, pesticides, and poly aromatic hydrocarbons. (Figure 1).

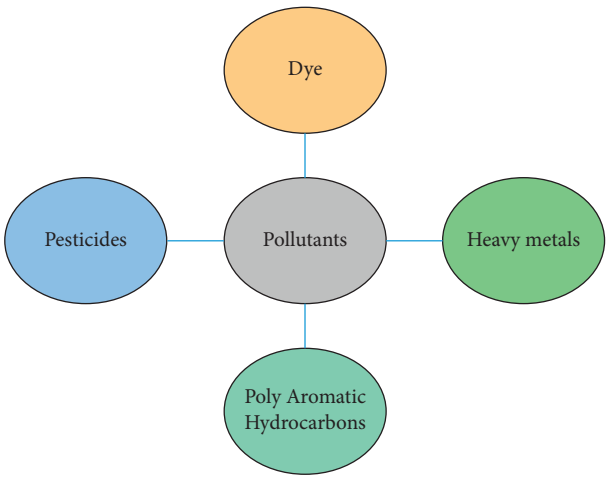


Figure 1: Types of pollutants.

2.1. Dyes. Dyes are a type of synthetic organic chemical that is used in a variety of industries, including textiles. As a result, during their manufacture and later during fabric dyeing, they have become frequent industrial environmental contaminants [11]. Synthetic dyes in textile wastewater decrease light penetration into rivers, altering the photosynthetic efficiency of aquatic vegetation and, as a result, the food supply of aquatic creatures. The thin coating of released dyes that might build on the receiving waters' surfaces reduces the quantity of dissolved oxygen, which has a serious impact on aquatic life. Biochemical oxygen demand (BOD) is increased by dye-containing industrial effluents. Dyes are basically persistent organic pollutants (POPs) that exist in the environment, and it has been a concern that these man-made chemicals are xenobiotic [12]. Dyes like Azure-B if introduced can affect the helical structure of an organism's DNA and duplex RNA. Disperse Red 1 and Disperse Orange 1 dyes are mutagenic in nature. Dyes such as Sudan I Dye (now banned), Basic Red 9 Dye, and Crystal Violet dye (induces tumor in fish) are carcinogenic in nature [13].

2.2. Heavy Metals. Heavy metal pollution is becoming more of an issue and a source of concern owing to the negative consequences it is causing all over the world [14]. Because of the rapidly increasing agriculture and metallurgy industries, inappropriate waste management, fertilizers, and pesticides, these come under inorganic pollutants and are being dumped in our waterways, land (soil), and the air (environment). Some of the most common heavy metals causing pollution are titanium (Ti)), iron (Fe), vanadium (V), chromium (Cr), manganese (Mn), cobalt (Co), copper (Cu), zinc (Zn), arsenic (As), molybdenum (Mo), nickel (Ni), cadmium (Cd), silver (Ag), tin (Sn), platinum (Pt), gold (Au), mercury (Hg), and lead (Pb) [14]. Heavy metals are classified as such because of their high gram atomic weight or higher densities. Heavy metals are now used to refer to poisonous synthetic metallic atoms or molecules and metalloids that are harmful to the environment and humans [15]. These toxic heavy metals are naturally available on the Earth's surface in their native state. The massive growth in their usage has resulted in an imminent influx of these metallic compounds in both the land and water natural resources. Heavy metal pollution has become an important component of anthropogenic activity, primarily due to metal mining, steelmaking, activities such as foundries and other metal-dependent industries, as well as metal leaching for extraction from various sources such as dumped landfills, excretion, domesticated poultry and chicken manure fertilizers, water runoffs, automobile industry, and road repair [16].

2.3. Pesticides. Pesticides are employed in agriculture for decades to safeguard crops and animals against insect infestations and to prevent the decrease in production. Pesticides are chemicals that are used to destroy unwanted organisms in gardens, agricultural farms, and other public locations. Pesticides, despite their use, may represent a threat to food safety, the environment, and all living things [17]. According to a survey conducted by the World Health Organization (WHO), around 1 million people are acutely poisoned as a result of pesticide exposure. Pesticides have a role in a variety of organic micropollutants with harmful environmental consequences. Biomagnification and bioconcentration are the two primary pollution processes [18]. Major interactions (receptor) in the human body for many pesticides are adipose fat tissue, causing these toxins to gather in lipids. DDT (dichlorodiphenyltrichloroethane) is an example of such a pesticide. If any human ingests any aquatic animal that has taken DDT with any food, it accumulates in the human body as DDT is lipophilic and dissolves in adipose tissues [19]. Inhibition or inability to reproduce, hormonal system disturbances, suppression of

the immune system, animal and fish tumors death, malignancies, effects on intergenerational impacts of teratogens, DNA, cell damage, and significant muck on fish gills and crusts, indicative of poor fish health, and other physiological impacts including the weakening of the eggshell are some of the problems caused by pesticides pollutants [20].

2.4. Poly Aromatic Hydrocarbons. Due to their toxic, mutagenic, and carcinogenic potential, poly aromatic hydrocarbons are a category of persistent and widespread environmental contaminants that are also hazardous to people. With the rise in demand for petroleum goods, their emissions have grown. Contributors include combustion in lack of oxygen of organic products, for example, coal and firewood [21]. Thermally, poly aromatic hydrocarbons are extremely stable and resistant to decomposition. The melting and boiling temperatures of these aromatic compounds are quite high. Poly aromatic hydrocarbons with a lower molecular weight are typically water soluble, but their hydrophobicity and aqueous insolubility increase with each additional ring [22]. The US Environmental Protection Agency has identified naphthalene, phenanthrene (C_4H_{10}), anthracene ($C_6H_4CH)_2$, fluoranthene ($C_{16}H_{10}$), pyrene ($C_{16}H_{10}$), chrysene ($C_{18}H_{12}$), benzo anthracene ($C_{18}H_{12}$), and benzo fluoranthene ($C_{20}H_{12}$) as pollutants that need to be kept in low concentration by constant monitoring.

3. Environmental and Health Consequences of Pollutants

Trace pollutants have been detected and analyzed during the recent decades with the establishment of refined techniques. However, subsequently with that emerging environmental concerns are to be reasoned with. The toxicological information regarding the majority of chemicals used is incomplete; long-term usage information or low-level exposure still lacks making the environmental challenges hidden. The path for identification of various sustainable as well as long-lasting techniques to counterattack the challenges of new emerging and identified pollutants with a detailed study of their consequences will be fruitful for human generations.

3.1. Environmental Consequences

3.1.1. Aquatic. One of the largest sources of aquatic pollution by environmental pollutants is Cultural Eutrophication. The increased rate of food production and population size has resulted in the discharge of nitrogen and phosphorus-rich compounds in water bodies [23]. This in turn has led to the expansion of harmful algal blooms toxic to the native aquatic plant and animal species and dysregulation of the biogeochemical cycle of persistently present organic pollutants in the natural environment [24]. Eutrophication has undesirably caused restriction of water use by fisheries, industries, and consumption as potable water because of decreased oxygen percentages and high ratios of poisonous compounds [25]. Another problem characterized by persistent organic pollutants is biomagnification, that is, the accumulation of poorly metabolizable, moderate to highly hydrophobic substances that are circulated in food webs [26]. The major effect caused due to these is a significant reduction in the numbers of biologically important local and native species which play an important role in the ecosystem.

The anthropogenic pollutants have threatened marine integrity; elevated carbon dioxide and trace metal levels have caused the imbalance of acid base homeostasis in organisms, thereby affecting their chance of survival and performance often also called Ocean Acidification [27]. It has caused an untimely calcification of coral reefs, microorganisms like zooplanktons, and shellfishes [28]. The aquatic toxicology reveals the presence of mixed toxicities including heavy metals [29] and persistent organic pollutants (POPs), including perfluoro alkylated acids (PFAS) worldwide [30]. Water is one of the most vulnerable and essential components of the environment and its pollution is a matter of great concern [31].

3.1.2. Terrestrial. The major contributors to environmental pollution are industries. The unauthorized dumping of pollutants such as aromatic hydrocarbons, plasticizers, and various xenobiotic compounds is a threat to the terrestrial ecosystem [32]. They are holders of mutagenic and carcinogenic properties and have decreased agricultural land production due to the accumulation of toxic compounds that have an impact on seed dormancy. The increased pollutant levels in the terrestrial domain by municipal sewage and extensive farm fertilization have led to the vulnerable quality of ground water by leaching of pollutants like heavy metals (zinc, cadmium, lead), chemical pesticides, fertilizers, and other micropollutants [33]. Nutrient secretion in land areas can lead to water nutrient enrichment through processes like eutrophication. Recent instances show the increasing risk of forest fires in zones contaminated with radioactive pollutants [34]. Increased deforestation due to various pollutants has directly or indirectly affected the increasing soil erosion rates and has displaced sediments and nutrient-rich soil surface causing loose soil [35] and thus promoting the events of flash floods.

3.1.3. Aerial. Air pollution is a threat to the whole human civilization and is the sixth leading cause of death in the world. Aerial pollutants like carbon dioxide, fuel soot particles, methane, halocarbons, ozone, and nitrous oxides have contributed to causing one of the biggest environmental challenges, i.e., global warming [36]. Global warming leads to ocean acidity, increased sea levels due to melting glaciers and snow, enhanced tropical calamities such as storms, loss of viable agricultural grounds and animal habitats, and dysregulated water supply [37]. Some pollutants like nitrogenous and sulfur oxides produced by combustion interact with the atmosphere and result in acid deposition in the form of rain, fog, or snow. It is known to destroy the stability of the ecosystem, threaten food reserves, and can be speculated to cause a socioeconomic implication at a larger

level socioeconomic implication [38]. The levels of air pollutants are increasing substantially in urban and peri-urban areas in developing countries like India. They in turn have affected the adjacent agricultural areas and have caused a threat to the crop production rates [39]. High particulate matter concentration can cause emission changes and physical, chemical changes in the atmosphere. The emissions of volatile organic compounds will be altered in the near future, which may further impact secondary organic aerosols [40].

3.2. Health Consequences. Air pollutants like oxides, particulate matter, and ground-level ozone are one of the most common health threats. Oxides such as CO (carbon monoxide) decrease oxygen delivery to body organs, impair vision, and at extreme concentrations are fatal. Sulfur oxides have been reported to be carcinogenic in nature if exposed for a long duration [41]. These pollutants have contributed significantly to problems such as cough, increased difficulty in breathing /decreased lung function, irritation in respiratory pathways, triggering asthma, and inflammation of bronchi and bronchioles. Besides adults, children also have been adversely affected; there have been increased cases of infant mortality, perinatal effects, allergy, cardiovascular and mental disorders, endothelial dysfunction, among them [42].

Traffic-Related Air Pollution. A mixture class of air pollutants can contribute to the generation of reactive oxygen species, creating an imbalance in free radical concentration that causes oxidative stress and initiates a risk of cancer development [43]. Increased oxidative stress and poor cardiovascular circulation cause inflammation in organs, which can cause mental health problems. Nitrogenous oxides and particulate fine and ultrafine matter have the capability to cross the blood-brain barrier and reach the brain, thus modulating Vaso regulation and even triggering neural inflammation and thus being a stepping stone for future neural dysregulations [44].

New pollutants are emerging and entering aquatic bodies: pesticides, chemically synthesized fertilizers, compounds such as dyes, detergents, and heavy metals have the potential to affect human health [45]. Eutrophic waters can cause skin-eye irritation, gastric inflammation, and vomiting in individuals. If high nitrogen levels are persistent, they may be a risk to infants <6 months of age, causing Blue Baby syndrome (decrease in the oxygen-carrying capacity of hemoglobin) [46]. Major pollutants even observed in drinking water, such as heavy metals—arsenic, lead, mercury, phthalates—have endocrine-disrupting properties. They can mimic, block, or interfere with the action of hormones present in the human body producing neurological or reproductive adverse developments. PFAS (Per- and polyfluoroalkyl substances), another type of pollutant, has been linked with infertility, thyroid dysfunction, cancer, and obesity [47]. Contaminants are known to cause problems like diarrhea, cholera, typhoid, malaria, and other congenital problems [48].

4. Traditional Methods of Removing Pollutants and Their Side Effects

Traditional methods like coagulation, adsorption, and advanced oxidation processes combined with electrochemical and biological processes (Table 1) have been used before the revealment of nanobiotechnology.

Coagulation and flocculation are traditional methods for the treatment of polluted water. They are used as a pretreatment of biological treatment, improving the biodegradability of waste and removing macropollutants [58]. Compounds like ferric chloride and polymers are added which destabilize the colloidal materials and small particles to agglomerate in large flocs that are settleable. It is a physiochemical approach to treat metal pollutants in industrial wastewater and remove soluble compounds by forming flocs [59]. It also aids in the elimination of suspended solids (SS) and organic material. Coagulation accompanied by ozonation has been used for waste sanitary landfill leachates [60]. The coagulation-flocculation process is also known to remove nutrient pollutants like P and nitrogen [61]. In comparison to this, adsorption is a surface phenomenon of organic as well as inorganic pollutant removal. It can be achieved through its subtypes like membrane adsorption where the contaminants are adsorbed to the membranes while the pollutant carrying source flows through it. Chromium(VI), zinc (II), and lead (II) are known to be successfully removed by polymer membrane-based adsorption methods and membrane-based adsorption methods, especially polypyrrole [62]. Activated carbon adsorption, as listed by the USEPA (Environmental Protection Agency), is one of the best control technologies for the removal of organic dyes and discharges from pharmaceutical companies [63]. Adsorption processes involving activated carbon have long been regarded as conventional treatment processes which can remove a multitude of pollutants [64]. *Zeolite Adsorption.* Natural zeolites known to exhibit high selectivity for heavy metals (cadmium, copper, zinc) and high abundance as compared to other chemicals synthesized for pollutant removal [64] have a high surface, large pores, thermally stable, cheap, chemically inert, and easily modified surface functional group that enables it to have excellent selectivity against pollutants [50]. *Biosorption.* Adsorbents are either prepared naturally or engineered artificially, giving them several advantages such as low cost, abundant raw material, high selectivity with regenerative ability, and performance comparable to conventional techniques. It can be used for the treatment of secondary and tertiary effluent pollutants [65].

Complete degradation of organic pollutants with mineralization can be achieved by advanced oxidation processes. Processes like ozonation, Fenton reaction, UV irradiation, peroxide handling, and their combination are being included in it. Polyphenols can be treated by Fenton and solar photo- Fentons. Photo-Fenton can also treat reoccurring organic compounds. AOP can aid in chemical oxygen demand and poly phenol removals up to 90 percent [66]. Pollutants such as toxic cyanide can be removed using a

Table 1: Side effects of using traditional processes for pollutant removal.

Pollutant removal method	Side effect	Reference
Biological processes	(i) Cannot treat toxic or refractory organic pollutants and have limitations.	[49]
Zeolite adsorption	(i) They might be highly hydrophilic on template removal because of their high surface silanol density, leading to low adsorption for hydrophobic pollutants.	[50]
Photocatalysis	(i) Economic constraints for the high level of mineralization. Postreaction products of photocatalysis could still remain toxic. (ii) Photo corrosion is a typical drawback, postseparation inorganic catalysts. Semiconductor photocatalysts are unstable under light irradiation.	[51, 52]
Electro kinetics	(i) There might be ionic motion, crystallization of metal salts, and dehydration which prevented removal of inorganic pollutants.	[53]
Electrochemical advanced oxidation processes	(i) High cost, high energy consumption for complete mineralization of pollutant.	[49]
Advanced oxidation process	(i) Considerably affected by pollutant nature, type, and concentration of oxidants and catalyst, reactor configuration. (ii) Release toxic and less biodegradable by-products in extreme cases.	[54, 55]
Electrocoagulation	(i) Myriad of designs for reactor formation.	[56]
Ozonation	(i) Difficult mineralization of pollutants due to the presence of hydroxyl ion scavengers.	[43]
Classical Fenton process	(i) Unsafe storage, transportation, and handling of hydrogen peroxide for large treatments.	[57]

combination of NaClO and hydrogen peroxide forming singlet oxygen and hydroxide ions [67]. AOPs tend to produce highly biodegradable and less toxic products [54]. Photolysis, i.e., degradation by light (photons) combined with hydrogen peroxide, can be used for the degradation of pesticides with the assistance of solvents like trichloromethane and carbon tetrachloride [68, 69].

Heterogenous photocatalysis by semiconductors such as titanium dioxide, zinc oxide, ferric oxide, copper sulfide, and cadmium sulfide can be efficient in degrading the pollutants into biodegradable compounds with a less toxic footprint as well as inorganic and halide ions [51], with additional benefits of low cost, easily crystallizable forms, phytochemical stable nature, and nontoxic behavior [55]. Ozonation based on the oxidation and/or degradation of pesticides (belonging to the classes organochlorine, organophosphorus, and triazine) allows for high reductions in chemical oxygen demand and significant removal of aromatic compounds [70]. Electrochemical processes have been also used as traditional methods for pollutant removal. Electrocoagulation uses a metal sacrificial anode which doses water. It has advantages in providing cations without increasing the saline nature of the solution, is easily operatable, has rapid sedimentation rates, reduced sludge production, and has less amount of coagulant ions required for treatment [56, 71]. Electrochemical Advanced Oxidation Processes have enabled the removal of pollutants and their complete mineralization with an environment-friendly approach. Extensive hydroxyl ions are produced even under optimal conditions of current and catalyst [72]. The current usage promises complete automation of the process on a large scale [49]. Electrokinetic remediation is a potential technique used for the separation and remediation of areas contaminated with organic and inorganic pollutants, for clean-up of low permeable matrices. It uses mechanisms like electromigration, electroosmosis, and electrophoresis combined with or without electrolysis [53, 73]. Electro Fenton is an electrochemical-based oxidation method [57]. The oxidation of organic compounds occurs via hydroxyl radical generated by electrochemical oxidation indirectly, and both hydrogen peroxide and ferrous ions can be electrogenerated under in situ conditions [74]. Electrochemical oxidation is a widely used popular technique for removing organic pollutants [75]. The oxidation is direct or indirect [76]. However, the direct oxidation process is in general not in use due to its ineffective nature in the degradation of organic pollutants, but its powerful oxidating agents produced could be used in indirect oxidation processes [57]. Biological methods have been adopted for pollutant removal due to their low cost and green nature [52]. They are widespread and conventional in nature, cost-effective, and well studied and can be modified according to the local needs [49]. They can be combined with other processes like wet air oxidation, ozonation for their increased efficiency, and better removal of pollutants that are toxic and refractory in nature [77].

5. Role of Nanotechnology and Nanomaterials for Remediation of Pollutants

Nanotechnology is a branch of science and engineering that is creating "materials" on the scale of atoms and molecules. Conventional principles of physical and chemical sciences do not apply to these sizes. The appearance, toughness, conductivity, and sensitivity of materials change drastically between the nano- and macroscales. For instance, carbon "nanotubes" are hundredfold stronger than steel, as well as sixfold lighter [78]. Various types of nanoparticles, their characterization mechanism, and how they can be applied in various sectors ranging from the health industry to the removal of environmental pollutants have been elucidated in this review.

For the elimination of environmental pollutants, the environmental remediation process employs a variety of methods (e.g., surface assimilation, absorption, synthetic reactions, catalysis using light, and filtering) [79]. Nanomaterials possess a high ratio of their surface-area-to-

volume, frequently resulting in greater biogeological and chemical reactivity compared to other macroscale materials. Their improved characteristics and efficacy make them particularly appropriate for carrying out such operations [80]. Nanoscale materials are being used in a wide array of fields, including science, the environment, industry, and medicine.

Over the last few years, an increase in the number of nanoscale products with environmental remedial applications has been created and deployed. Nanomaterials, for example, have been utilized to clean up contaminated soil and groundwater at various hazardous waste sites like those damaged by chlorinated solvents or in oil spills. Manufactured nanoparticles have physicochemical, surface, and optical-electronic characteristics that answer issues that were previously difficult to tackle with traditional methods. It can design unique ways for creating new methodologies, replacing current tools, and producing new materials and chemicals with high performance and low energy usage [81].

6. Different Methods for Nanoparticle Synthesis

Science and innovations have made rapid advances in the production of nanomaterials to obtain unique characteristics that differ from those of bulk materials. Nanoparticles can be prepared by chemical, physical, and biological methods (Figure 2).

6.1. Chemical Methods

6.1.1. Using Polyols for Synthesis of Nanoparticles. Polyols have numerous hydroxyl groups attached to any organic compound. Nonaqueous liquid (in which the solvent is any other than water),falling under polyols is used as a solvent and reducing agent in this technique. The use of nonaqueous solvents in this technique has the benefit of reducing surface oxidation and assemblage [82]. If the synthesis is carried out at a somewhat elevated temperature with precise particle growth control, the polyol process can be used as a sol-gel technique in the production of oxide [83]. Oxide-based nanoparticles of vanadium, yttrium, manganese, cobalt, tin, lead, and titanium can be produced by this method. Because of its unique properties of high reducing ability and high boiling point, ethylene glycol is the most broadly adopted solvent in the polyol technique in metal oxide nanoparticles manufacturing [84]. Bimetallic alloys and core-shell nanoparticles can also be produced by this method [85].

6.1.2. Reduction in Solution (Colloidal Method). The use of reducing agents such as $Na_3C_6H_5O_7$ is one of the most essential ways. Turkevich suggested this method for the first time for the manufacture of Au monodisperse colloidal solutions in 1951. This approach allows for the creation of spherical nanoparticles. Later, the same approach was utilized to make silver nanoparticles, and the spread of these particles was bigger, ranging from 60 to 200 nm. In this method, a metal salt gives nanoparticles as residues upon reduction in a liquid solution. This method is generally used for the synthesis of gold, iron, molybdenum, and silver nanoparticles. For gold, a solution of chloroauric acid is heated to a boil. Trisodium citrate is used as a reducing agent. After mixing and stirring for 10 mins, seed crystals changed to gold nanoparticles.

6.1.3. Electrochemical Synthesis. The capacity to precisely tune the chosen potential and the rejection of the potentially wasteful alternative half-reaction are the key advantages of this technology over a conventional chemical reaction [86]. The electrochemical production of Ag nanoparticles has received a lot of attention lately. The electrochemical process employed consisted of dissolving a metallic anode in an aprotic liquid [87]. Adding electrons to oxidize the positive electrode of the cell for solvable silver ions in acetonitrile containing tetra butyl ammonium yields Ag nanoparticles ranging in size from 2 to 7 nm. By adjusting the current density, the particle size can be determined. The electrochemical production of styrene-stabilized red luminous Si nanoparticles is also commercialized. Under UV stimulation, Si nanoparticles glow, which is useful in optics [88]. Platinum nanoparticles can also be synthesized by the electrochemical method. The desired Pt nanoparticles are deposited on the surface of the electrode with a size of about 10 nm [89].

6.2. Physical Methods

6.2.1. Plasma Method. Another approach for producing nanoparticles is the plasma process. Radiofrequency (RF) heating coils create the plasma. The first metal is placed in a pestle, which is then placed in an evacuated chamber. High-voltage RF coils are wrapped around the evacuated chamber; then, the metal is heated above its evaporation point [90]. Helium is the gas employed in the technique, which after flowing into the system generates a high-temperature plasma in the region of the coils. Metal vapor forms on helium gas atoms and diffuses up to a cold collecting rod, where nanoparticles are collected and passivated by oxygen gas [86].

6.2.2. Microwave Irradiation. There is the usage of microwaves to heat samples, which reduces reaction time and requires fewer materials. Furthermore, the use of microwaves allows for better control of reaction processes. The higher the effectiveness of nanoparticles with greater applicability, whether as ferrofluid, cell separation, or pollution removal, the closer they are to the spherical form and the more homogeneity between the forms [91]. Microwave-assisted irradiation can be used to make ZnS nanoparticles. XRD, SEM, and UV-Vis spectroscopy were used to analyze the generated ZnS nanoparticles. According to the XRD spectrum, the average size of the nanocrystallites was measured using the Debye-Scherrer formula, and they were determined to be around 6 nm [84].

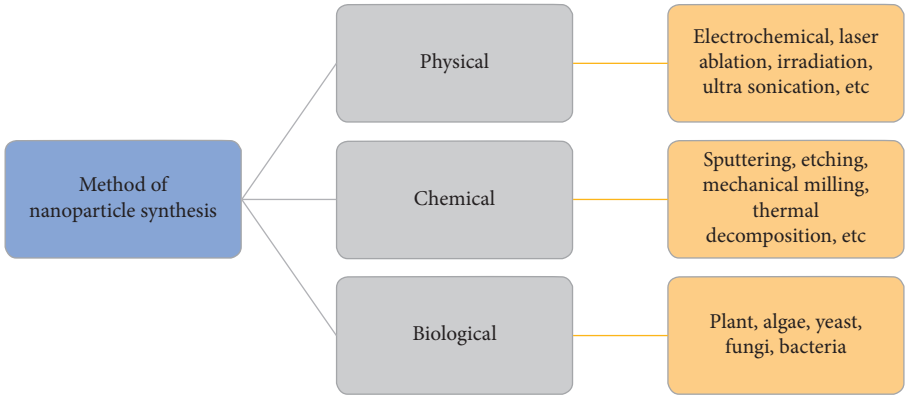


Figure 2: Method of nanoparticle synthesis.

6.2.3. Radiation Dependent. Gamma radiation is repeatable, can control particle shape, yields monodisperse metallic nanoparticles, is simple, is inexpensive, and uses fewer toxins as a precursor: it uses the fewest reagents, uses a reaction temperature close to room temperature with as few synthetic steps as possible, and minimizes the amount of generated by-products and waste [92]. The radiolytic reduction has been shown to be an effective method for producing monosized and widely distributed metallic clusters [93]. The excitation and ionization of the solvent are the major consequences of high-energy gamma rays interacting with a solution of metal ions [94].

6.3. Biological Method. Green synthesis is necessary to avoid the creation of undesired or hazardous by-products by developing dependable, long-term, and environmentally friendly manufacturing processes. To attain this aim, suitable solvent systems and natural resources are required. To accommodate diverse biological components, green production of metallic nanoparticles is being used.

6.3.1. Plant. Plant leaf extracts contain important elements to be used in the process of synthesis of nanoparticles. At various experimental parameters, the plant extract is combined with metal precursor solutions [95]. The parameters that determine the settings of plant leaf extracts are acknowledged to affect the pace of nanoparticle production, as well as the output and stability of the particles [96]. Flavones, terpenoids, sugars, ketones, aldehydes, carboxylic acids, and amides are the major phytochemicals involved in nanoparticle bioreduction [97]. The enol-to-keto-transformation regulates the generation of biogenic Ag nanoparticles in sweet basil extracts [98]. Sugars found in plant extracts, such as glucose and fructose, also have a role in the creation of metallic nanoparticles [99]. Plant extracts are biomolecules composed of carbohydrates and proteins that function as a reducing agent in the synthesis of metallic nanoparticles [100]. Geranium leaf extract contains terpenoids that actively participate in the conversion of Ag ions into nanoparticles [101]. Eugenol is a major terpenoid found in cinnamon extracts and is required for the bioreduction of the metal ions $HAuCl_4$ and $AgNO_3$ [102].

6.3.2. Bacteria. Commercial biotechnological applications such as environmental remediation, genetic manipulation, and biosorption have all made use of bacterial species [103]. Bacteria have the ability to decrease metal ions and are important candidates in the creation of nanoparticles [104]. Due to its economic potential and viability, green synthesis technology delivers a clean, nontoxic, and environmentally acceptable technique for the production of metallic nanoparticles [105]. Bacteria like *Delftia acidovorans* can synthesize pure Au nanoparticles [106]. Delftibactin is considered a nonribosomal peptide (NRP) that has been linked to the production of Au nanoparticles due to its ability to induce resistance to toxic Au ions. Due to the formation of inert gold nanoparticles bound to delftibactin, the transition metal gold had no toxicity toward bacteria [107]. *Bacillus licheniformis* is said to have produced Ag nanoparticles within its cells [108].

6.3.3. Fungi. Green synthesis of metal/metal-oxide nanomaterials by fungi is also a very rapid method. Because they include a range of enzymes, they are better green means for the synthesis of metal-based nanoparticles [109]. In comparison to bacteria, the competent fungus can generate greater volumes of nanoparticles [110]. *Trichoderma reesei* has an advantage over other fungi in that it is a well-studied organism that can be controlled to produce a large number of enzymes and may aid in increasing the rate of nanoparticle synthesis [111]. The fungus *Neurospora crassa* can be used in the biological manufacture of platinum nanoparticles. Single PtNPs with diameters ranging from 4 to 35 nm were generated intracellularly [112].

6.3.4. Algae. Heavy metals are thought to accumulate in algae, which could be used in the biogenic manufacture of metallic nanoparticles [113]. The dried unicellular alga *Chlorella vulgaris* could produce tetrahedral, decahedral, and icosahedral nanoparticles that aggregated near the surface. The proteins in the algal extract serve as a stabilizing agent, a reducing agent, and a shape-control modifier, among other things [114]. *Sargassum wightii*, a marine alga, has the ability to produce extracellular Ag and Au nanoparticles [115]. The gold nanoparticles created are of real

yield since they are made from living *Euglena gracilis* microalga cells that have been cultivated under either mixotrophic or autotrophic conditions [116].

7. Types of Nanoparticles and Their Characterization Methods

Nanoparticles can be divided further into many subtypes according to broader categories like material type, dimensions, or 3D structure and origin.

7.1. Basis of Origin. Nanoparticles can be either natural in origin or engineering (synthetic). Natural nanoparticles are produced in nature by biological processes and organisms or induced by human activity [117]. Synthetic or engineered nanoparticles as their name describes are nanomaterials produced by mechanical processes, exhaustive mechanisms, or chemical, biological, and physical processes individually or in combination [118].

7.2. Basis of 3D Structure. On the basis of structure, they can be classified into three categories: one dimension, two dimensions, and three dimensions. In zero D, the electrons cannot move in any axis; i.e., they are dimensionless; in 1 D, they can move along a single axis followed on by 2 D and 3 D where they can move in two and three axes, respectively, which is irrespective of the axis combination [119]. One-dimension nanoparticles are thin films of sizes ranging from 1 to 100 nm. Two-dimension nanoparticles generally include nanotubes, e.g., carbon nanotubes. Three-dimension nanoparticles consist of dendrimers, quantum dots, and fullerenes (like carbon 60) [120].

7.3. Based on Material Type. Nanomaterials can be subclassified according to the material they are made of such as inorganic, organic, carbon-containing, or composite substances. Inorganic-based nanomaterials are made of metals or their oxides, e.g., gold, silver [121], titanium dioxide, zinc oxide, or silicon-based compounds, etc. [119]. Organic nanomaterials are obtained from organic matter excluding carbon-based or inorganic-based sources; they are present in forms like micelles, liposomes, and polymeric particles. Composite nanoparticles are multiphasic in nature such that one of their phases can combine with other similar or dissimilar nanoparticles in bulk, complex formations, e.g., a gold nanoparticle in composition with ceramic or a polymer. Carbon-based nanoparticles consist of fullerenes, nanotubes, and nanofibers found in generally two-dimensional morphological structures [122–124].

7.4. Characterization. Various techniques (Figure 3) have been used for characterization, including UV-visible spectroscopy, that is, based on the principle of surface plaque resonance (SPR) [125]. Another technique is X-ray diffraction (XRD) to determine the crystalline phase [126–128]. Fourier transform infrared (FTIR) spectroscopy helps to identify the functional groups present in the sample. Characterization of nanoparticles based on size, morphology, and surface charge was performed using atomic force microscopy (AFM), scanning electron microscopy, and transmission electron microscopy [129–131]. Dynamic light scattering is used to measure nanoparticles size.

8. Different Application of Nanoparticles

The usage of nanoparticles/materials has shown tremendous growth worldwide from medical grafts to sensing tools. They have revolutionized how various techniques happen. Green synthesized particles especially silver nanoparticles have antibiotic activities; they have been used for medical implantations as silicone artificial heart valves, catheters, bone prosthetics, dental sets, and bandages [132]. The key ability of particles at the nanoscale has promoted the formation of biosensors for the easier diagnosis of diseases, checking the progression of the problem, as well as tracking the progress of the given therapy [133]. Green synthesized biodegradable nanoscale particles can have a sustained drug release for a long period of time without reporting the risk of infections [134]. Many studies have reported the multifactorial antimicrobial property of green nanoparticles. Generally, nanoparticles either function extracellularly or intracellularly to alter the bacterial growth rates such as inhibition of cytological transport and signaling [135]. Antibiotic activities of NP antifungal properties have also been reported. These particles can degrade cell walls and have been found to be as effective as 87% in studies with concentrations as low as 6 mg/ml [136].

Through the usage of gold nanoclusters and in vivo studies, carcinogenic cells were bioimaged with the help of fluorescent objects in cell lines such as hepatocarcinoma cell lines and leukemia cell lines. These nanoparticles are in clustered form with fluorescence activity [137]. Huang et al. created nanoparticles of carbon which were negligibly cytotoxic and highly biocompatible with hydrophilic nature and compatible nature. The blue photoluminescence cells were used to stain human HeLa cells and green fluoresced imaging [138]. Through recent advances in technology, antineoplastic traits of synthesized nanoparticles, which can have much lower side effects, have been the topic of research. These nanoparticles are very effective, provide a small amount of negative impact and target tumorigenic cell lines, and can also be used in combination with high-dose drugs to achieve the correct therapy [139]. Target therapies with a combination of anticarcinogenic nanomaterials use carriers like polymeric micelles, dendrimers, and liposomes. Inorganic metallic NM including metals and metal oxides is a promising approach for the future in cell imaging, using it as sensors, and drug delivery [140].

9. Nanomaterials in Remediation of Soil

Soil contamination and degradation continue to be a significant environmental problem, with remediation becoming a worldwide challenge. Degrading soil reservoirs throughout the world have a great influence on agricultural production and food safety; therefore, this needs to be addressed right away [141]. Heavy metals (remains that become toxins), herbicides, and POPs-persistent organic

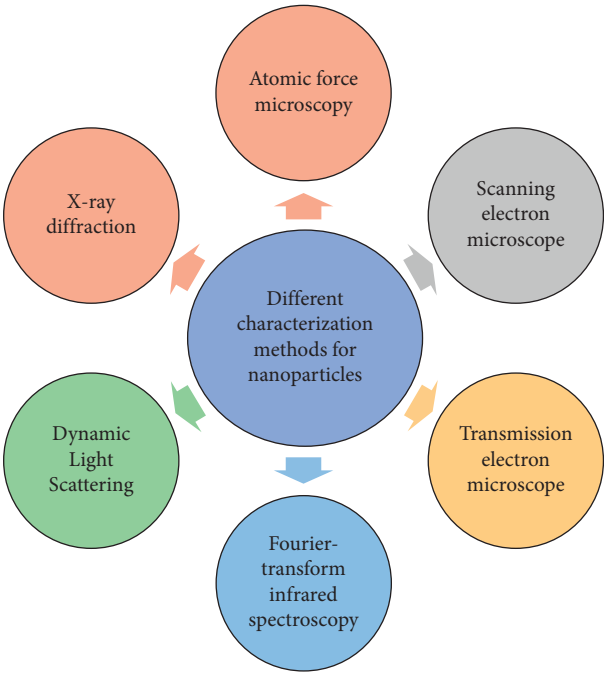


Figure 3: Different characterization methods for nanoparticles.

pollutants contaminate soil, adding to the issue. Because of biomagnification, soil contaminated with these chemicals provides an increased chance of intoxication of the food chain. Both the urgent production to feed more people and the difficulty of avoiding additional land deterioration have a significant impact on agricultural production [141]. The immobilization or adsorption procedure is one of the most commonly utilized treatment techniques. The effectiveness, low cost, and environmentally favorable approach for removing metal pollutants from soils are some of the most prominent benefits connected with this sort of remediation technology [142]. Carbon-based nanomaterials, for example, carbon nanotubes, oxides of metal (ferric oxide and titanium oxide), and various nanocomposites are some of the engineered nanomaterials that have been utilized to immobilize soil pollutants [143]. For example, ferric oxide nanoparticles have an extraordinary potential to absorb and immobilize heavy metals such as Cd and As from various media samples [144]. Oily sewage is an adverse problem affecting the aquatic local life [145]. Iron nanoparticles are successful in removal of Total Petroleum Hydrocarbons (TPHs) from water with enhanced results of 88.34% [146]. Thus, nanotechnology-based treatments are able to provide high-performance treated water which has less impurities and toxic substances combined with the removal of heavy metals. Nanoscience's can be combined with biological treatments for reducing the hazards of environmental substances. Organic dyes can be efficiently decolorized through industrial mechanisms backboned by nanoparticles [147]. Gold and silver nanoparticles are very successful catalyst in dye degradation, the basic process being two stepped; first the accumulation of electrons on particle surface and second the diffusion of dye for a reduction reaction [148].

Magnetic nanoparticles have an extended range of use within adsorption and catalytic pollution treatment [149]. Polyacrylamide-modified magnetite nanomaterials can be used, as they provide soil erosion and arsenate leaching [150]. On-site immobilization of pollutants has received a lot of attention as a feasible and economical method of remediating polluted soils. Various NM-based full form modifications have been studied, to find a relevant apt material that takes into account low cost, high efficiency, better stability, least harmful environmental effect, and maximum reusability. By reducing the exchangeable percentage of metals in polluted sediments and soils, nanohydroxyapatite particles successfully immobilized metal concentrations in pore water [151]. This research stated that utilizing sodium carboxymethyl cellulose-stabilized nZVI, they were able to remove 80 percent of soil-bound Cr (VI). Redox-sensitive elements as pollutants can be treated by abiotic reductive reactions directly impacting their mobility and toxicity in soil alone or in combination with immobilization techniques. The addition of nZVI to pyrene-adulterated soil, for example, led to strong reducing conditions, allowing pyrene to be removed through reduction. Nanomaterials are also used in advanced oxidation processes (AOPs) that employ various oxidants for the breakdown of organic contaminants in polluted soil [152]. Highly reactive radicals (water radical, sulfur dioxide radical) can be generated by combining hydrogen peroxide or persulfate with soluble Fe (II) for the efficient oxidative breakdown of organic contaminants. However, acidic pH (which is necessary to avoid Fe (II) precipitation) and related drawbacks, such as the expense of early acidification and negative impacts on soil quality and microorganisms, restrict these activities. As a result, solid iron phases have been proposed instead of

soluble Fe (II) to allow chemical oxidation without pH alteration. Iron particles when used with chelating ligands act as an enhancement in the process [153]. Green nanoparticles have a very large scope for providing solutions against environmental pollutants and provide scope for the treatment of localized environments contaminated with metal ions toxic in nature. These materials can be useful in remediation and cleaning waste sites. Ferrous NP can be applied for disinfection of heavy metal contaminated water as well as heavy metal-containing soil, additionally being effective eco-friendly fertilizers. These nanofertilizers can synchronize plant nutrient uptake while decreasing groundwater contamination as they are immobilized and will not change into permanent chemical or gaseous forms that are unattainable for the growing plant [154]. Magnetite (Fe_3O_4) andgreigite (Fe_3S_4) with siliceous material produced using bacteria have been used in optical coatings for solar energy applications [155].

10. Nanomaterials in Remediation of Water Pollution

Wastewater discharge from commercial and industrial factories, as well as untreated residential sludge and synthetic contaminants influx to aquatic resources, is extremely harmful to humans as well as for the ecosystem. Heavy metal ions, organic compounds, and oils are the primary water pollutants in this regard, and they can render any water stream unfit for consumption [156]. Through the advancement of nanotechnology, a wide spectrum of novel nanomaterial-based technologies has emerged. Nanomaterials, as adsorbent systems, have a huge surface area for reactions at a very smaller relative weight, are manufactured at a lower price than activated carbon, and can effectively withdraw contaminants. A wide array of nanomaterial-based adsorbents in various shapes and parameters like dimensionalities have been studied in this respect, including zero-dimensional nanoparticles, one-dimensional nanofibers and nanotubes, two-dimensional nanosheets, and three-dimensional nanoflowers [157]. In terms of composition, there is a lot of variety, and various carbon-based and noncarbon-based nanomaterials can assist in cleansing aquatic bodies. The segregation process depends on the contaminant's chemical/physical interaction with the surface of the nanomaterial. Nanomaterials can also function as a major manufacturing component of the permeable segregator unit as a nanocomposite supreme filtration membrane, or as a thin film addition to enhancing the hydrophilicity and thermomechanical characteristics of the membrane [158]. Zinc and titanium oxides, ceramic, nanowire, and polymeric membranes, carbon nanotubes, and submicron particles are used in various remediation ways such as lysis, filtration, adsorption, and oxidation [159].

Dye pollutants are deadly because of their carcinogenic nature. Dyes have been used in textiles, paintings, pigments, and other products for thousands of years. Around 0.1 million different synthetic dyes are manufactured on an industrial level today. Talking about yearly usage, roughly 1.6 million tonnes of dyes are utilized. During usage, fifteen percent of the total is squandered [160]. Crocein orange G and acid green 25 adsorption capability is high in chitosan sugar-coated magnetite nanoparticles. A magnetic field, interestingly, could easily recover the adsorbent nanoparticles [161]. Fe_3O_4/activated carbon nanoparticles capable of separating 138 mg/g and 166.6 mg/g dyes (methylene blue and brilliant green) have been produced based on this idea [162]. In removing methylene blue dye from water, electrospun polyether sulfone nanofibers having vanadium nanoparticles are used. Because the nanofibers have a low isoelectric point, they develop a large, highly hydroxylated surface area in an alkaline medium, which promotes the adsorption of cationic methylene blue molecules [163]. Adsorption has the disadvantage of not being able to totally "eliminate" or "neutralize" dye contaminants. Instead, it collects and gathers dye constituents exclusively by moving them at different molecular phases. This functionality may be difficult to implement since the dye-based sewage is difficult to remove and is adsorbed on the rarely reused membrane [158]. A number of progressive oxidation processes that release hydroxyl radicals (ȮH) hold promise for the process of removing colors from textile effluents in this respect. Hydroxyl radical is highly reactive and oxidizes refractory organic contaminants due to its unpaired electrons. The increasing rate of photoreaction, which is photocatalysis, has become one of the most investigated advanced oxidative remediation methods and is regarded as a realistic destructive method for dyes (organic) and pesticides due to less cost. Metal oxides that have properties of semiconductors such as zinc oxide and titanium oxide demonstrated impressive photodecomposition capability when used to degrade dye contaminants [164]. Despite semiconductors' significant potential for decomposition in presence of the light of different dyes (organic) contaminants, industries with affiliated visible light photocatalysts have low steadiness or efficiency when exposed to light semiconductors having their energy gaps falling in the visible light spectrum that belongs to groups II–VI that have been offered as majestic, feasible options to address such issues [165].

11. Nanomaterials in Remediation of Air Pollution

Nanotechnology can help to clean up the air in a variety of ways. For gaseous reactions, one method is to use nanocatalysts with a larger surface area. Catalysts accelerate chemical processes that convert hazardous vapors from automobiles and industrial facilities into innocuous gases. The removal of air waste to the ground is ineffective; therefore, the filtered chemicals posed a difficulty for disposal [166]. Filters are beneficial in the case of air pollution. A filter has a small porous construction that allows gas to flow through while particles stay trapped inside. The membrane of the filter treats toxins using three mechanisms: filter structure surface contact, inertial force applied while moving the direction of gas, and electric charge generated by particle and filter construction [167]. The primary flaw of filters is that they cause a pressure drop, which necessitates a

lot of energy usage. Nanofilters are more efficient than traditional filters, as their pores range from 1 to 10 nm, thereby efficiently removing a variety of microorganisms and organic pollutants. Carbon nanotube-based membranes have a better capability for separating CO_2 from other gases [168]. Carbon nanotubes can capture gases at a pace that is 100-fold faster than current gas separation techniques. As a result, they are suitable for large-scale applications [169]. There seems to be a reverse connection between the efficiency of the gas separation process and the related volume flowing through traditional layers, but carbon-based nanotube membranes cover this. Nanomembrane technology may be utilized to separate and purify gases and pollutant vapors in a variety of sectors, as well as to prevent their discharge into the environment [170]. Nanosensors are a cutting-edge device which can identify and take action for a detectable physical change on such a nanometre scale. Nanoparticles, which are small 3D circular substances, are used as sensors and provide information when a toxin is present in a space [171]. One of the hazards in the commercial sector is the risk of obnoxious and hazardous spillage. Sensors used mostly at commercial levels have recently discovered small amounts of these gases. This requires the use of technical advancements in the creation of faster and more accurate sensors. This type of biosensor consists of single or multimembrane nanotubes that can capture dangerous gas molecules. Smart dust is a new kind of nanosensor. The ultimate goal is to develop a series of advanced sensors in the form of ultralight nanocomputers. Such silicon-based ultrafine particles have a structure that permits them to wirelessly transfer data collected to the main server. The crucial data transportation rate with this sort of sensor is about 1 kbps or more. These solar-powered nanosensors can easily stay hung in the air for hours. Smart dust can convey information such as climate conditions, humidity, and amount of pollutants for up to 20 kilometers and can provide conditions to continuously manage air pollution in a given area [171].

12. Conclusion and Future Prospects

Nanoparticles have revolutionized the future approach for litigation of environment with biotechnology. With the increased progress of humans, industrialization and consumption of reserves are present in nature. There is an urgent requirement for the removal of these pollutants with the least possible side effects to the Earth, though not lamenting the side effects of earlier existing processes like Fenton, coagulation, adsorption, advanced oxidation etc. and their combinations as well. A more sustainable, holistic approach to the tackling of this grave issue can be nanoparticles that may be active while working on both small- and large-scale levels. Diversified nanoparticles have become handy for removing specified pollutants in their native area like sewage spills, landmines, etc. Moreover, they cause a minimum toxic effect during production rates and thus can be well researched forward so that they are more compatible in the future with decreased production costs.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] A. S. Adeleye, J. R. Conway, K. Garner, Y. Huang, Y. Su, and A. A. Keller, "Engineered nanomaterials for water treatment and remediation: costs, benefits, and applicability," *Chemical Engineering Journal*, vol. 286, pp. 640–662, 2016.

[2] S. E. Mancebo and S. Q. Wang, "Recognizing the impact of ambient air pollution on skin health," *Journal of the European Academy of Dermatology and Venereology*, vol. 29, no. 12, pp. 2326–2332, 2015.

[3] A. Roy and N. Bharadvaja, "Efficient removal of heavy metals from artificial wastewater using biochar," *Environmental Nanotechnology, Monitoring & Management*, vol. 16, Article ID 100602, 2021.

[4] I. Manisalidis, E. Stavropoulou, A. Stavropoulos, and E. Bezirtzoglou, "Environmental and health impacts of air pollution: a review," *Frontiers in Public Health*, vol. 8, 2020.

[5] A. Roy, A. Elzaki, V. Tirth et al., "Biological synthesis of nanocatalysts and their applications," *Catalysts*, vol. 11, no. 12, p. 1494, 2021.

[6] H. M. Ahmed, A. Roy, M. Wahab et al., "Applications of nanomaterials in agrifood and pharmaceutical industry," *Journal of Nanomaterials*, vol. 2021, Article ID 1472096, 10 pages, 2021.

[7] J. R. Marlon, M. T. Ballew, C. Roser-Renouf, A. Leiserowitz, and E. Maibach, "How hope and doubt affect climate change mobilization," *Frontiers in Communication*, vol. 4, p. 20, 2019.

[8] S. Kaur and A. Roy, "Bioremediation of heavy metals from wastewater using nanomaterials," *Environment, Development and Sustainability*, vol. 23, no. 7, pp. 9617–9640, 2021.

[9] S. H. Khan, "Green nanotechnology for the environment and sustainable development," *Green Materials for Wastewater Treatment*, vol. 13, pp. 13–46, 2020.

[10] C. L. Seewagen, "The threat of global mercury pollution to bird migration: potential mechanisms and current evidence," *Ecotoxicology*, vol. 29, no. 8, pp. 1254–1267, 2020.

[11] R. Kant, "Textile dyeing industry an environmental hazard," *Natural Science*, vol. 04, no. 01, pp. 22–26, 2012.

[12] L. Pereira and M. Alves, "Dyes—environmental impact and remediation," in *Environmental Protection Strategies for Sustainable Development*, pp. 111–162, Springer, Berlin, Germany, 2012.

[13] B. Lellis, C. Z. Fávaro-Polonio, J. A. Pamphile, and J. C. Polonio, "Effects of textile dyes on health and the environment and bioremediation potential of living organisms," *Biotechnology Research and Innovation*, vol. 3, no. 2, pp. 275–290, 2019.

[14] J. Briffa, E. Sinagra, and R. Blundell, "Heavy metal pollution in the environment and their toxicological effects on humans," *Heliyon*, vol. 6, no. 9, Article ID e04691, 2020.

[15] P. B. Tchounwou, C. G. Yedjou, A. K. Patlolla, and D. J. Sutton, "Heavy metal toxicity and the environment," *Experientia Supplementum*, vol. 101, pp. 133–164, 2012.

[16] P. K. Gautam, R. K. Gautam, M. C. Chattopadhyaya, S. Banerjee, M. C. Chattopadhyaya, and J. D. Pandey, "Heavy metals in the environment: fate, transport, toxicity and remediation technologies," *Nova Science Publishers*, vol. 60, pp. 101–130, 2016.

[17] M. A. Hassaan and A. El Nemr, "Pesticides pollution: classifications, human health impact, extraction and treatment techniques," *The Egyptian Journal of Aquatic Research*, vol. 46, no. 3, pp. 207–220, 2020.

[18] X. Zhen, L. Liu, X. Wang, G. Zhong, and J. Tang, "Fates and ecological effects of current-use pesticides (CUPs) in a typical river-estuarine system of Laizhou Bay, North China," *Environmental Pollution*, vol. 252, pp. 573–579, 2019.

[19] Z.-Q. Jia, Y.-C. Zhang, Q.-T. Huang, A. K. Jones, Z.-J. Han, and C.-Q. Zhao, "Acute toxicity, bioconcentration, elimination, action mode and detoxification metabolism of broflanilide in zebrafish, *Danio rerio*," *Journal of Hazardous Materials*, vol. 394, Article ID 122521, 2020.

[20] J.-I. Hwang, A. R. Zimmerman, and J.-E. Kim, "Bioconcentration factor-based management of soil pesticide residues: endosulfan uptake by carrot and potato plants," *The Science of the Total Environment*, vol. 627, pp. 514–522, 2018.

[21] R. Pandey, F. Masood, H. P. Singh, and D. R. Batish, "Polycyclic aromatic hydrocarbons as environmental pollutants: a review," *International Journal of Advanced Research in Science, Engineering and Technology*, vol. 6, pp. 1361–1369, 2017.

[22] J. Masih, A. Masih, A. Kulshrestha, R. Singhvi, and A. Taneja, "Characteristics of polycyclic aromatic hydrocarbons in indoor and outdoor atmosphere in the North central part of India," *Journal of Hazardous Materials*, vol. 177, no. 1–3, pp. 190–198, 2010.

[23] S. J. Tilley, E. V. Orlova, R. J. C. Gilbert, P. W. Andrew, and H. R. Saibil, "Structural basis of pore formation by the bacterial toxin pneumolysin," *Cell*, vol. 121, no. 2, 2005.

[24] J. Dachs, S. J. Eisenreich, and R. M. Hoff, "Influence of eutrophication on air-water exchange, vertical fluxes, and phytoplankton concentrations of persistent organic pollutants," *Environmental Science & Technology*, vol. 34, no. 6, pp. 1095–1102, 2000.

[25] M. N. Khan and F. Mohammad, *Eutrophication: Challenges and Solutions BT - Eutrophication: Causes, Consequences and Control*, Ansari and S. S. Gill, Eds., Springer, Dordrecht, Netherlands, 2014.

[26] B. C. Kelly, M. G. Ikonomou, J. D. Blair, A. E. Morin, and F. A. P. C. Gobas, "Food web–specific biomagnification of persistent organic pollutants," *Science*, vol. 317, no. 5835, pp. 236–239, 2007.

[27] A. V. Ivanina and I. M. Sokolova, "Interactive effects of metal pollution and ocean acidification on physiology of marine organisms," *Current Zoology*, vol. 61, no. 4, pp. 653–668, 2015.

[28] R. Billé, R. Kelly, A. Biastoch et al., "Taking action against Ocean Acidification: a review of management and policy options," *Environmental Management*, vol. 52, no. 4, pp. 761–779, 2013.

[29] R. Altenburger, W. Boedeker, M. Faust, and L. H. Grimme, "Regulations for combined effects of pollutants: consequences from risk assessment in aquatic toxicology," *Food and Chemical Toxicology*, vol. 34, no. 11-12, pp. 1155–1157, 1996.

[30] T. Groffen, "Preliminary study on the distribution of metals and persistent organic pollutants (POPs), including perfluoroalkylated acids (PFAS), in the aquatic environment near Morogoro, Tanzania, and the potential health risks for humans," *Environmental Research: Anthropology and Aesthetics*, vol. 192, Article ID 110299, 2021.

[31] I. C. Vasilachi, D. M. Asiminicesei, D. I. Fertu, and M. Gavrilescu, "Occurrence and fate of emerging pollutants in water environment and options for their removal," *Water*, vol. 13, no. 2, 2021.

[32] P. S. Phale, A. Sharma, and K. Gautam, "Microbial degradation of xenobiotics like aromatic pollutants from the terrestrial environments," *Pharmaceuticals and Personal Care Products: Waste Management and Treatment Technology*, pp. 259–278, 2019.

[33] M. Waldman and Y. Shevah, "Biodegradation and leaching of pollutants: monitoring aspects (technical report)," *Pure and Applied Chemistry*, vol. 65, no. 7, pp. 1595–1603, 1993.

[34] X. Álvarez, E. Valero, R. M. Santos, S. G. P. Varandas, L. S. Fernandes, and F. A. L. Pacheco, "Anthropogenic nutrients and eutrophication in multiple land use watersheds: Best management practices and policies for the protection of water resources," *Land Use Policy*, vol. 69, pp. 1–11, 2017.

[35] J. F. Artiola, J. L. Walworth, S. A. Musil, and M. A. Crimmins, "Soil and land pollution," in *Environmental and Pollution Science*, pp. 219–235, Elsevier, Amsterdam, Netherlands, 2019.

[36] P. Michaelis, "Global warming: efficient policies in the case of multiple pollutants," *Environmental and Resource Economics*, vol. 2, no. 1, pp. 61–77, 1992.

[37] M. Z. Jacobson, "Review of solutions to global warming, air pollution, and energy security," *Energy & Environmental Science*, vol. 2, no. 2, pp. 148–173, 2009.

[38] A. Singh and M. Agrawal, "Acid rain and its ecological consequences," *Journal of Environmental Biology*, vol. 29, no. 1, p. 15, 2007.

[39] M. Agrawal, B. Singh, M. Rajput, F. Marshall, and J. N. B. Bell, "Effect of air pollution on peri-urban agriculture: a case study," *Environmental Pollution*, vol. 126, no. 3, pp. 323–329, 2003.

[40] H. Orru, K. L. Ebi, and B. Forsberg, "The interplay of climate change and air pollution on health," *Current Environmental Health Reports*, vol. 4, no. 4, pp. 504–513, 2017.

[41] R. Kelishadi and P. Poursafa, "Air pollution and non-respiratory health hazards for children," *Archives of Medical Science*, vol. 4, pp. 483–495, 2010.

[42] S. Siddique, M. R. Ray, and T. Lahiri, "Effects of air pollution on the respiratory health of children: a study in the capital city of India," *Air Quality, Atmosphere & Health*, vol. 4, no. 2, pp. 95–102, 2011.

[43] I. N. Ismail, J. Jalaludin, S. A. Bakar, N. H. Hisamuddin, and N. F. Suhaimi, "Association of traffic-related air pollution (TRAP) with DNA damage and respiratory health symptoms among primary school children in selangor," *Asian Journal of Atmospheric Environment*, vol. 13, no. 2, 2019.

[44] S. Roberts, "Exploration of NO_2 and PM2.5 air pollution and mental health problems using high-resolution data in London-based children from a UK longitudinal cohort study," *Psychiatry Research*, vol. 272, pp. 8–17, 2019.

[45] K. Jayaswal, V. Sahu, and B. R. Gurjar, "Water pollution, human health and remediation," in *Energy, Environment, and Sustainability*, S. Bhattacharya, A. B. Gupta, A. Gupta, and A. Pandey, Eds., Springer, Singapore, pp. 11–27, 2018.

[46] M. M. Ghangrekar and P. Chatterjee, "Water pollutants classification and its effects on environment," in *Carbon*

Nanostructures, pp. 11–26, Springer International Publishing, Cham, Switzerland, 2018.

[47] N. Quinete and R. A. Hauser-Davis, "Drinking water pollutants may affect the immune system: concerns regarding COVID-19 health effects," *Environmental Science and Pollution Research*, vol. 28, no. 1, pp. 1235–1246, 2021.

[48] R. Qadri and M. A. Faiq, "Freshwater pollution: effects on aquatic life and human health," in *Fresh Water Pollution Dynamics and Remediation*, pp. 15–26, Springer, Singapore, 2020.

[49] O. Ganzenko, D. Huguenot, E. D. van Hullebusch, G. Esposito, and M. A. Oturan, "Electrochemical advanced oxidation and biological processes for wastewater treatment: a review of the combined approaches," *Environmental Science and Pollution Research*, vol. 21, no. 14, pp. 8493–8524, 2014.

[50] P. N. E. Diagboya and E. D. Dikio, "Silica-based mesoporous materials; emerging designer adsorbents for aqueous pollutants removal and water treatment," *Microporous and Mesoporous Materials*, vol. 266, pp. 252–267, 2018.

[51] Y. Pi, "Adsorptive and photocatalytic removal of Persistent Organic Pollutants (POPs) in water by metal-organic frameworks (MOFs)," *Chemical Engineering Journal*, vol. 337, pp. 351–371, 2018.

[52] C. Zhang, Y. Li, H. Shen, and D. Shuai, "Simultaneous coupling of photocatalytic and biological processes: a promising synergistic alternative for enhancing decontamination of recalcitrant compounds in water," *Chemical Engineering Journal*, vol. 403, Article ID 126365, 2021.

[53] D. Wen, R. Fu, and Q. Li, "Removal of inorganic contaminants in soil by electrokinetic remediation technologies: a review," *Journal of Hazardous Materials*, vol. 401, Article ID 123345, 2021.

[54] A. R. Ribeiro, O. C. Nunes, M. F. R. Pereira, and A. M. T. Silva, "An overview on the advanced oxidation processes applied for the treatment of water pollutants defined in the recently launched Directive 2013/39/EU," *Environment*, vol. 75, pp. 33–51, 2015.

[55] R. Dewil, D. Mantzavinos, I. Poulios, and M. A. Rodrigo, "New perspectives for advanced oxidation processes," *Journal of Environmental Management*, vol. 195, pp. 93–99, 2017.

[56] P. Holt, G. Barton, and C. Mitchell, "Electrocoagulation as a wastewater treatment," *Environmental Engineering Research*, vol. 1000, pp. 41–46, 1999.

[57] A. Fernandes, M. J. Pacheco, L. Ciríaco, and A. Lopes, "Review on the electrochemical processes for the treatment of sanitary landfill leachates: present and future," *Applied Catalysis B: Environmental*, vol. 176-177, pp. 183–200, 2015.

[58] O. S. Amuda and I. A. Amoo, "Coagulation/flocculation process and sludge conditioning in beverage industrial wastewater treatment," *Journal of Hazardous Materials*, vol. 141, no. 3, pp. 778–783, 2007.

[59] K. O. Iwuozor, "Prospects and challenges of using coagulation-flocculation method in the treatment of effluents," *Advanced Journal of Chemistry-Section A*, vol. 2, pp. 105–127, 2019.

[60] T. Poznyak, G. L. Bautista, I. Chaírez, R. I. Córdova, and L. E. Ríos, "Decomposition of toxic pollutants in landfill leachate by ozone after coagulation treatment," *Journal of Hazardous Materials*, vol. 152, no. 3, pp. 1108–1114, 2008.

[61] M. Aguilar, J. Sáez, M. Lloréns, A. Soler, and J. F. Ortuño, "Nutrient removal and sludge production in the coagulation-flocculation process," *Water Research*, vol. 36, no. 11, pp. 2910–2919, 2002.

[62] K. C. Khulbe and T. Matsuura, "Removal of heavy metals and pollutants by membrane adsorption techniques," *Applied Water Science*, vol. 8, no. 1, p. 19, 2018.

[63] M. Zambianchi, M. Durso, A. Liscio et al., "Graphene oxide doped polysulfone membrane adsorbers for the removal of organic contaminants from water," *Chemical Engineering Journal*, vol. 326, pp. 130–140, 2017.

[64] M. S. El-Geundi, "Adsorbents for industrial pollution control," *Adsorption Science and Technology*, vol. 15, no. 10, pp. 777–787, 1997.

[65] N. Chaukura, W. Gwenzi, N. Tavengwa, and M. M. Manyuchi, "Biosorbents for the removal of synthetic organics and emerging pollutants: opportunities and challenges for developing countries," *Environmental Development*, vol. 19, pp. 84–89, 2016.

[66] P. C. Lindholm-Lehto, J. S. Knuutinen, H. S. J. Ahkola, and S. H. Herve, "Refractory organic pollutants and toxicity in pulp and paper mill wastewaters," *Environmental Science and Pollution Research*, vol. 22, no. 9, pp. 6473–6499, 2015.

[67] H. Suty, C. De Traversay, and M. Cost, "Applications of advanced oxidation processes: present and future," *Water Science and Technology*, vol. 49, no. 4, pp. 227–233, 2004.

[68] C. P. James, E. Germain, and S. Judd, "Micropollutant removal by advanced oxidation of microfiltered secondary effluent for water reuse," *Separation and Purification Technology*, vol. 127, pp. 77–83, 2014.

[69] A. G. De Oliveira, J. P. Ribeiro, J. T. De Oliveira, D. De Keukeleire, M. S. Duarte, and R. F. Do Nascimento, "Degradation of the pesticide chlorpyrifos in aqueous solutions with UV/H_2O_2: optimization and effect of interfering anions," *Journal of Advanced Oxidation Technologies*, vol. 17, no. 1, pp. 133–138, 2014.

[70] O. Chedeville, M. Debacq, and C. Porte, "Removal of phenolic compounds present in olive mill wastewaters by ozonation," *Desalination*, vol. 249, no. 2, pp. 865–869, 2009.

[71] M. Elazzouzi, K. Haboubi, and M. S. Elyoubi, "Electrocoagulation flocculation as a low-cost process for pollutants removal from urban wastewater," *Chemical Engineering Research and Design*, vol. 117, pp. 614–626, 2017.

[72] A. Dirany, I. Sirés, N. Oturan, A. Özcan, and M. A. Oturan, "Electrochemical treatment of the antibiotic sulfachloropyridazine: kinetics, reaction pathways, and toxicity evolution," *Environmental Science & Technology*, vol. 46, no. 7, pp. 4074–4082, 2012.

[73] I. Hassan, E. Mohamedelhassan, and E. K. Yanful, "Solar powered electrokinetic remediation of Cu polluted soil using a novel anode configuration," *Electrochimica Acta*, vol. 181, pp. 58–67, 2015.

[74] C. A. Martínez-Huitle and M. Panizza, "Electrochemical oxidation of organic pollutants for wastewater treatment," *Current Opinion in Electrochemistry*, vol. 11, pp. 62–71, 2018.

[75] C. A. Martínez-Huitle and S. Ferro, "Electrochemical oxidation of organic pollutants for the wastewater treatment: direct and indirect processes," *Chemical Society Reviews*, vol. 35, no. 12, pp. 1324–1340, 2006.

[76] M. Panizza and G. Cerisola, "Application of diamond electrodes to electrochemical processes," *Electrochimica Acta*, vol. 51, no. 2, pp. 191–199, 2005.

[77] M. Jeworski and E. Heinzle, "Combined chemical-biological treatment of wastewater containing refractory pollutants," *Biotechnology Annual Review*, vol. 6, pp. 163–196, 2000.

[78] D. Janas and K. K. Koziol, "Carbon nanotube fibers and films: synthesis, applications and perspectives of the direct-spinning method," *Nanoscale*, vol. 8, no. 47, pp. 19475–19490, 2016.

[79] D. Jiang, G. Zeng, D. Huang et al., "Remediation of contaminated soils by enhanced nanoscale zero valent iron," *Environmental Research*, vol. 163, pp. 217–227, 2018.

[80] F. Guerra, M. Attia, D. Whitehead, and F. Alexis, "Nanotechnology for environmental remediation: materials and applications," *Molecules*, vol. 23, no. 7, p. 1760, 2018.

[81] E. F. Mohamed, "Nanotechnology: future of environmental air pollution control," *Environmental Management and Sustainable Development*, vol. 6, no. 2, p. 429, 2017.

[82] E. Luxa, J. Hartman, T. Antonatos, N. Shaban, P. Oparin, and Z. Sofer, "Non-aqueous solution-processed phosphorene by controlled low-potential electrochemical exfoliation and thin film preparation," *Nanoscale*, vol. 12, no. 4, pp. 2638–2647, 2020.

[83] C. Quievryn, S. Bernard, and P. Miele, "Polyol-based synthesis of praseodymium oxide nanoparticles," *Nanomaterials and Nanotechnology*, vol. 4, p. 7, 2014.

[84] T. Cele, "Preparation of nanoparticles," in *Engineered Nanomaterials - Health and Safety*IntechOpen, London, UK, 2020.

[85] D. I. Garcia-Gutierrez, C. E. Gutierrez-Wing, L. Giovanetti, J. M. Ramallo-López, F. G. Requejo, and M. Jose-Yacaman, "Temperature effect on the synthesis of Au-Pt bimetallic nanoparticles," *The Journal of Physical Chemistry B*, vol. 109, no. 9, pp. 3813–3821, 2005.

[86] P. D. Cavaliere, A. Perrone, and A. Silvello, "Water electrolysis for the production of hydrogen to be employed in the ironmaking and steelmaking industry," *Metals*, vol. 11, no. 11, p. 1816, 2021.

[87] L. Rodríguez-Sánchez, M. C. Blanco, and M. A. López-Quintela, "Electrochemical synthesis of silver nanoparticles," *The Journal of Physical Chemistry B*, vol. 104, no. 41, pp. 9683–9688, 2000.

[88] J. Choi, K. Kim, H.-S. Han, M. P. Hwang, and K. H. Lee, "Electrochemical synthesis of red fluorescent silicon nanoparticles," *Bulletin of the Korean Chemical Society*, vol. 35, no. 1, pp. 35–38, 2014.

[89] M. Islam and M. S. Islam, "Electro-deposition method for platinum nano-particles synthesis," *Electro-deposition Method for Platinum Nano-Particles Synthesis*, vol. 1, 2013.

[90] B. P. Kafle, "Introduction to nanomaterials and application of UV-Visible spectroscopy for their characterization," in *Chemical Analysis and Material Characterization by Spectrophotometry*, pp. 147–198, Elsevier, Amsterdam, Netherlands, 2020.

[91] H. Wang, J. Z. Xu, J. J. Zhu, and H. Y. Chen, "Preparation of CuO nanoparticles by microwave irradiation," *Journal of crystal growth*, vol. 244, no. 1, pp. 88–94, 2002.

[92] Y. N. Rao, D. Banerjee, A. Datta, S. K. Das, R. Guin, and A. Saha, "Gamma irradiation route to synthesis of highly re-dispersible natural polymer capped silver nanoparticles," *Radiation Physics and Chemistry*, vol. 79, no. 12, pp. 1240–1246, 2010.

[93] J. L. Marignier, J. Belloni, M. O. Delcourt, and J. P. Chevalier, "Microaggregates of non-noble metals and bimetallic alloys prepared by radiation-induced reduction," *Nature*, vol. 317, no. 6035, pp. 344-345, 1985.

[94] W. Abidi and H. Remita, "Gold based nanoparticles generated by radiolytic and photolytic methods," *Recent Patents on Engineering*, vol. 4, no. 3, pp. 170–188, 2010.

[95] A. K. Mittal, Y. Chisti, and U. C. Banerjee, "Synthesis of metallic nanoparticles using plant extracts," *Biotechnology Advances*, vol. 31, no. 2, pp. 346–356, 2013.

[96] A. D. Dwivedi and K. Gopal, "Biosynthesis of silver and gold nanoparticles using *Chenopodium album* leaf extract," *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, vol. 369, no. 1–3, pp. 27–33, 2010.

[97] L. Mathew, N. Chandrasekaran, and A. Mukherjee, "Biomimetic synthesis of nanoparticles: science, technology & applicability," in *Biomimetics Learning from Nature*InTech, London, UK, 2010.

[98] N. Ahmad, S. Sharma, M. K. Alam et al., "Rapid synthesis of silver nanoparticles using dried medicinal plant of basil," *Colloids and Surfaces B: Biointerfaces*, vol. 81, no. 1, pp. 81–86, 2010.

[99] S. Panigrahi, S. Kundu, S. Ghosh, S. Nath, and T. Pal, "General method of synthesis for metal nanoparticles," *Journal of Nanoparticle Research*, vol. 6, no. 4, pp. 411–414, 2004.

[100] S. Iravani, "Green synthesis of metal nanoparticles using plants," *Green Chemistry*, vol. 13, no. 10, p. 2638, 2011.

[101] S. S. Shankar, A. Ahmad, R. Pasricha, and M. Sastry, "Bioreduction of chloroaurate ions by geranium leaves and its endophytic fungus yields gold nanoparticles of different shapes," *Journal of Materials Chemistry*, vol. 13, no. 7, p. 1822, 2003.

[102] A. K. Singh, M. Talat, D. P. Singh, and O. N. Srivastava, "Biosynthesis of gold and silver nanoparticles by natural precursor clove and their functionalization with amine group," *Journal of Nanoparticle Research*, vol. 12, no. 5, pp. 1667–1675, 2010.

[103] M. Gericke and A. Pinches, "Microbial production of gold nanoparticles," *Gold Bulletin*, vol. 39, no. 1, pp. 22–28, 2006.

[104] S. Iravani, "Bacteria in nanoparticle synthesis: current status and future prospects," *International Scholarly Research Notices*, vol. 2014, Article ID 359316, 18 pages, 2014.

[105] D. Zhang, X. Ma, Y. Gu, H. Huang, and G. Zhang, "Green synthesis of metallic nanoparticles and their potential applications to treat cancer," *Frontiers in Inside Cosmetics*, vol. 8, 2020.

[106] C. W. Johnston, M. A. Wyatt, X. Li et al., "Gold biomineralization by a metallophore from a gold-associated microbe," *Nature Chemical Biology*, vol. 9, no. 4, pp. 241–243, 2013.

[107] N. Pantidos, "Biological synthesis of metallic nanoparticles by bacteria, fungi and plants," *Journal of Nanomedicine and Nanotechnology*, vol. 05, no. 05, 2014.

[108] K. Kalimuthu, R. Suresh Babu, D. Venkataraman, M. Bilal, and S. Gurunathan, "Biosynthesis of silver nanocrystals by Bacillus licheniformis," *Colloids and Surfaces B: Biointerfaces*, vol. 65, no. 1, pp. 150–153, 2008.

[109] Y.-L. Chen, H.-Y. Tuan, C.-W. Tien, W.-H. Lo, H.-C. Liang, and Y.-C. Hu, "Augmented biosynthesis of cadmium sulfide nanoparticles by genetically engineered *Escherichia coli*," *Biotechnology Progress*, vol. 25, no. 5, pp. 1260–1266, 2009.

[110] P. Mohanpuria, N. K. Rana, and S. K. Yadav, "Biosynthesis of nanoparticles: technological concepts and future applications," *Journal of Nanoparticle Research*, vol. 10, no. 3, pp. 507–517, 2008.

[111] K. Vahabi, G. A. Mansoori, and S. Karimi, "Biosynthesis of silver nanoparticles by fungus trichoderma reesei (a route for large-scale production of AgNPs)," *Insciences Journal*, vol. 1, pp. 65–79, 2011.

[112] L. Castro, M. L. Blázquez, J. A. Muñoz, F. González, and A. Ballester, "Biological synthesis of metallic nanoparticles using algae," *IET Nanobiotechnology*, vol. 7, no. 3, pp. 109–116, 2013.

[113] T. Luangpipat, I. R. Beattie, Y. Chisti, and R. G. Haverkamp, "Gold nanoparticles produced in a microalga," *Journal of Nanoparticle Research*, vol. 13, no. 12, pp. 6439–6445, 2011.

[114] N. Kahzad and A. Salehzadeh, "Green synthesis of $CuFe_2O_4$@Ag nanocomposite using the Chlorella vulgaris and evaluation of its effect on the expression of norA efflux pump gene among *Staphylococcus aureus* Strains$CuFe_2O_4$@Ag nanocomposite using the Chlorella vulgaris and evaluation of its effect on the expression of norA efflux pump gene among *Staphylococcus aureus* strains," *Biological Trace Element Research*, vol. 198, no. 1, pp. 359–370, 2020.

[115] K. Govindaraju, V. Kiruthiga, V. G. Kumar, and G. Singaravelu, "Extracellular synthesis of silver nanoparticles by a marine alga, sargassum wightii grevilli and their antibacterial effects," *Journal of Nanoscience and Nanotechnology*, vol. 9, no. 9, pp. 5497–5501, 2009.

[116] S. A. Dahoumane, C. Yéprémian, C. Djédiat et al., "Improvement of kinetics, yield, and colloidal stability of biogenic gold nanoparticles using living cells of Euglena gracilis microalga," *Journal of Nanoparticle Research*, vol. 18, no. 3, p. 79, 2016.

[117] M. F. Hochella, M. G. Spencer, and K. L. Jones, "Nanotechnology: nature's gift or scientists' brainchild?" *Environmental Sciences: Nano*, vol. 2, no. 2, pp. 114–119, 2015.

[118] S. Wagner, A. Gondikas, E. Neubauer, T. Hofmann, and F. von der Kammer, "Spot the difference: engineered and natural nanoparticles in the environment—release, behavior, and fate," *Angewandte Chemie*, vol. 53, no. 46, pp. 12398–12419, 2014.

[119] J. Jeevanandam, A. Barhoum, Y. S. Chan, A. Dufresne, and M. K. Danquah, "Review on nanoparticles and nanostructured materials: history, sources, toxicity and regulations," *Beilstein Journal of Nanotechnology*, vol. 9, pp. 1050–1074, 2018.

[120] S. Bhatia, "Nanoparticles types, classification, characterization, fabrication methods and drug delivery applications," in *Natural Polymer Drug Delivery Systems*, pp. 33–93, Springer, Berlin, Germany, 2016.

[121] S. Hasan, "A review on nanoparticles: their synthesis and types," *Research Journal of Recent Sciences*, vol. 2277, p. 2502, 2015.

[122] N. Kumar and S. Kumbhat, "Carbon-based nanomaterials," *Essentials in Nanoscience and Nanotechnology*, pp. 189–236, 2016.

[123] B. Magenheim, M. Y. Levy, and S. Benita, "A new in vitro technique for the evaluation of drug release profile from colloidal carriers-ultrafiltration technique at low pressure," *International Journal of Pharmacy Times*, vol. 94, no. 1–3, pp. 115–123, 1993.

[124] P. Nagore, S. Ghotekar, K. Mane, A. Ghoti, M. Bilal, and A. Roy, "Structural properties and antimicrobial activities of polyalthia longifolia leaf extract-mediated CuO nanoparticles," *BioNanoScience*, vol. 11, no. 2, pp. 579–589, 2021.

[125] A. Roy, "Plant derived silver nanoparticles and their therapeutic applications," *Current Pharmaceutical Biotechnology*, vol. 22, no. 14, pp. 1834–1847, 2021.

[126] S. Raina, A. Roy, and N. Bharadvaja, "Degradation of dyes using biologically synthesized silver and copper nanoparticles," *Environmental Nanotechnology, Monitoring & Management*, vol. 13, Article ID 100278, 2020.

[127] S. Mittal and A. Roy, "Fungus and plant-mediated synthesis of metallic nanoparticles and their application in degradation of dyes," in *Photocatalytic Degradation of Dyes*, pp. 287–308, Elsevier, Amsterdam, Netherlands, 2021.

[128] A. Verma, A. Roy, and N. Bharadvaja, "Remediation of heavy metals using nanophytoremediation," in *Advanced Oxidation Processes for Effluent Treatment Plants*, pp. 273–296, Elsevier, Amsterdam, Netherlands, 2021.

[129] A. Roy and N. Bharadvaja, "Silver nanoparticle synthesis from Plumbago zeylanica and its dye degradation activity," *Bioinspired, Biomimetic and Nanobiomaterials*, vol. 8, no. 2, pp. 130–140, 2019.

[130] A. Roy and N. Bharadvaja, "Silver nanoparticles synthesis from a pharmaceutically important medicinal plant *Plumbago zeylanica*," *MOJ Bioequivalence & Bioavailability*, vol. 3, no. 5, p. 46, 2017.

[131] A. Roy and N. Bharadvaja, "Qualitative analysis of phytocompounds and synthesis of silver nanoparticles from *Centella asiatica*," *Innovative Techniques in Agriculture*, vol. 1, no. 2, pp. 88–95, 2017.

[132] M. Rafique, I. Sadaf, M. S. Rafique, and M. B. Tahir, "A review on green synthesis of silver nanoparticles and their applications," *Artificial Cells, Nanomedicine, and Biotechnology*, vol. 45, no. 7, pp. 1272–1291, 2017.

[133] L. Marchiol, "Synthesis of metal nanoparticles in living plants," *Italian Journal of Agronomy*, vol. 7, no. 3, p. e37, 2012.

[134] K. Parveen, V. Banse, and L. Ledwani, "Green synthesis of nanoparticles: their advantages and disadvantages," *AIP Conference Proceedings*, vol. 1724, no. 1, Article ID 20048, 2016.

[135] R. Banasiuk, M. Krychowiak, D. Swigon et al., "Carnivorous plants used for green synthesis of silver nanoparticles with broad-spectrum antimicrobial activity," *Arabian Journal of Chemistry*, vol. 13, no. 1, pp. 1415–1428, 2020.

[136] T. Alam, R. A. A. Khan, A. Ali, H. Sher, Z. Ullah, and M. Ali, "Biogenic synthesis of iron oxide nanoparticles via Skimmia laureola and their antibacterial efficacy against bacterial wilt pathogen Ralstonia solanacearum," *Materials Science and Engineering: C*, vol. 98, pp. 101–108, 2019.

[137] J.-G. Wang, S.-S. Jang, D. S.-H. Wong, S.-S. Shieh, and C.-W. Wu, "Soft-sensor development with adaptive variable selection using nonnegative garrote," *Control Engineering Practice*, vol. 21, no. 9, pp. 1157–1164, 2013.

[138] H. Huang, Y. Xu, C.-J. Tang, J.-R. Chen, A.-J. Wang, and J.-J. Feng, "Facile and green synthesis of photoluminescent carbon nanoparticles for cellular imaging," *New Journal of Chemistry*, vol. 38, no. 2, pp. 784–789, 2014.

[139] P. S. Tourinho, C. A. M. Van Gestel, S. Lofts, C. Svendsen, A. M. V. M. Soares, and S. Loureiro, "Metal-based nanoparticles in soil: fate, behavior, and effects on soil invertebrates," *Environmental Toxicology & Chemistry*, vol. 31, no. 8, pp. 1679–1692, 2012.

[140] Z. Sanaeimehr, I. Javadi, and F. Namvar, "Antiangiogenic and antiapoptotic effects of green-synthesized zinc oxide nanoparticles using *Sargassum muticum* algae extraction," *Cancer Nanotechnology*, vol. 9, no. 1, pp. 1–16, 2018.

[141] M. Bakshi and P. C. Abhilash, "Nanotechnology for soil remediation: revitalizing the tarnished resource," in *Nano-Materials as Photocatalysts for Degradation of Environmental Pollutants*, pp. 345–370, Elsevier, Amsterdam, Netherlands, 2020.

[142] R. Bhadouria, D. Mishra, V. K. Singh et al., "Nanocatalyst types and their potential impacts in agroecosystems: an

overview," in *Nano-Materials as Photocatalysts for Degradation of Environmental Pollutants*, pp. 323–344, Elsevier, Amsterdam, Netherlands, 2020.

[143] Z. Arif, N. K. Sethy, L. Kumari, P. K. Mishra, and S. N. Upadhyay, "Recent advances in functionalized polymer-based composite photocatalysts for wastewater treatment," in *Nano-Materials as Photocatalysts for Degradation of Environmental Pollutants*, pp. 39–64, Elsevier, Amsterdam, Netherlands, 2020.

[144] N. Gaur, G. Flora, M. Yadav, and A. Tiwari, "A review with recent advancements on bioremediation-based abolition of heavy metals," *Environmental Sciences: Processes Impacts*, vol. 16, no. 2, pp. 180–193, 2014.

[145] Y. Peng, Y. Liu, J. Dai, L. Cao, and X. Liu, "A sustainable strategy for remediation of oily sewage: clean and safe," *Separation and Purification Technology*, vol. 240, Article ID 116592, 2020.

[146] E. Murgueitio, L. Cumbal, M. Abril, A. Izquierdo, A. Debut, and O. Tinoco, "Green synthesis of iron nanoparticles: application on the removal of petroleum oil from contaminated water and soils," *Journal of Nanotechnology*, vol. 2018, Article ID 4184769, 8 pages, 2018.

[147] K. Sharma, G. Singh, G. Singh, M. Kumar, and V. Bhalla, "Silver nanoparticles: facile synthesis and their catalytic application for the degradation of dyes," *RSC Advances*, vol. 5, no. 33, pp. 25781–25788, 2015.

[148] N. G. Bastús, F. Merkoçi, J. Piella, and V. Puntes, "Synthesis of highly monodisperse citrate-stabilized silver nanoparticles of up to 200 nm: kinetic control and catalytic properties," *Chemistry of Materials*, vol. 26, no. 9, pp. 2836–2846, 2014.

[149] Z. Ajmal, A. Muhmood, M. Usman et al., "Phosphate removal from aqueous solution using iron oxides: adsorption, desorption and regeneration characteristics," *Journal of Colloid and Interface Science*, vol. 528, pp. 145–155, 2018.

[150] M. Zheng, "Simultaneous control of soil erosion and arsenic leaching at disturbed land using polyacrylamide modified magnetite nanoparticles," *Science of the Total Environment*, vol. 702, Article ID 134997, 2020.

[151] Y. Wang, Z. Fang, B. Liang, and E. P. Tsang, "Remediation of hexavalent chromium contaminated soil by stabilized nanoscale zero-valent iron prepared from steel pickling waste liquor," *Chemical Engineering Journal*, vol. 247, pp. 283–290, 2014.

[152] M. Usman, M. Farooq, A. Wakeel et al., "Nanotechnology in agriculture: current status, challenges and future opportunities," *The Science of the Total Environment*, vol. 721, Article ID 137778, 2020.

[153] H. Fu, Y. Yang, R. Zhu et al., "Superior adsorption of phosphate by ferrihydrite-coated and lanthanum-decorated magnetite," *Journal of Colloid and Interface Science*, vol. 530, pp. 704–713, 2018.

[154] R. Prasad, V. K. Jain, and A. Varma, "Role of nanomaterials in symbiotic fungus growth enhancement," *Current Science*, vol. 99, no. 9, pp. 1189–1191, 2010.

[155] I. Hussain, N. B. Singh, A. Singh, H. Singh, and S. C. Singh, "Green synthesis of nanoparticles and its potential application," *Biotechnology Letters*, vol. 38, no. 4, pp. 545–560, 2016.

[156] H. Lu, J. Wang, M. Stoller, T. Wang, Y. Bao, and H. Hao, "An overview of nanomaterials for water and wastewater treatment," *Advances in Materials Science and Engineering*, vol. 2016, Article ID 4964828, 10 pages, 2016.

[157] S. Homaeigohar, "The nanosized dye adsorbents for water treatment," *Nanomater (Basel, Switzerland)*, vol. 10, no. 2, 2020.

[158] P. Bansal, G. R. Chaudhary, and S. K. Mehta, "Comparative study of catalytic activity of ZrO_2 nanoparticles for sonocatalytic and photocatalytic degradation of cationic and anionic dyes," *Chemical Engineering Journal*, vol. 280, pp. 475–485, 2015.

[159] P. K. T. R. Singh, S. Vats, D. Kumar, and S. Tyagi, "Nanomaterials use in wastewater treatment," in *Proceedings of the International Conference on Nanotechnology and Chemical Engineering*, vol. 21, Bangkuk, Thailand, 2012.

[160] K. B. Tan, M. Vakili, B. A. Horri, P. E. Poh, A. Z. Abdullah, and B. Salamatinia, "Adsorption of dyes by nanomaterials: recent developments and adsorption mechanisms," *Separation and Purification Technology*, vol. 150, pp. 229–242, 2015.

[161] Y.-C. Chang and D.-H. Chen, "Adsorption kinetics and thermodynamics of acid dyes on a carboxymethylated chitosan-conjugated magnetic nano-adsorbent," *Macromolecular Bioscience*, vol. 5, no. 3, pp. 254–261, 2005.

[162] S. Joshi and V. K. Garg, "Applications of Fe_3O_4@AC nanoparticles for dye removal from simulated wastewater," *Chemosphere*, vol. 236, 2019.

[163] S. Homaeigohar, A. Zillohu, R. Abdelaziz, M. Hedayati, and M. Elbahri, "A novel nanohybrid nanofibrous adsorbent for water purification from dye pollutants," *Materials*, vol. 9, no. 10, p. 848, 2016.

[164] Q. Li, T. Zhao, M. Li et al., "One-step construction of Pickering emulsion via commercial TiO_2 nanoparticles for photocatalytic dye degradation," *Applied Catalysis B: Environmental*, vol. 249, pp. 1–8, 2019.

[165] M. Ghadimi, S. Zangenehtabar, and S. Homaeigohar, "An overview of the water remediation potential of nanomaterials and their ecotoxicological impacts," *Water (Switzerland)*, vol. 12, no. 4, 2020.

[166] R. D. Singh, "Technological options for the removal of arsenic with special reference to South East Asia," *Journal of environmental management*, vol. 107, pp. 1–18, 2012.

[167] F. Shao, W. Xiong, R. Zhang, and C. Huang, "Electrospun nanofibers membranes for effective air filtration," *Macromolecular Materials and Engineering*, vol. 302, no. 1, Article ID 1600353, 2016.

[168] A. I. Skoulidas, D. S. Sholl, and J. K. Johnson, "Adsorption and diffusion of carbon dioxide and nitrogen through single-walled carbon nanotube membranes," *The Journal of Chemical Physics*, vol. 124, no. 5, Article ID 054708, 2006.

[169] S. Chaturvedi, P. N. Dave, and N. K. Shah, "Applications of nano-catalyst in new era," *Journal of Saudi Chemical Society*, vol. 16, no. 3, pp. 307–325, 2012.

[170] L. Zhang, B. Zhao, C. Jiang, J. Yang, and G. Zheng, "Preparation and transport performances of high-density, aligned carbon nanotube membranes," *Nanoscale Research Letters*, vol. 101, no. 10, pp. 1–8, 2015.

[171] F. Peng, Y. Su, Y. Zhong, C. Fan, S.-T. Lee, and Y. He, "Silicon nanomaterials platform for bioimaging, biosensing, and cancer therapy," *Accounts of Chemical Research*, vol. 47, no. 2, pp. 612–623, 2014.

J. Inst. Eng. India Ser. B
https://doi.org/10.1007/s40031-021-00708-0

ORIGINAL CONTRIBUTION

New Multicarrier Modulation Scheme for Harmonics Mitigation of T-Type Solar Multilevel Inverter

Sanjay Upreti[1] · Shivam Kumar Yadav[2] · Bhim Singh[2] · Narendra Kumar[3]

Received: 16 May 2021 / Accepted: 6 December 2021
© The Institution of Engineers (India) 2021

Abstract This paper presents a new low-frequency modulation scheme for the power quality improvement of a T-type multilevel inverter (T-MLI) based solar photovoltaic (SPV) system. The T-type converter is employed for DC-AC conversion, generating five voltage levels with six switches per phase. The new modulation scheme is generated with two different functions i.e., sine and pulse generator, to have a modified carrier wave. This unique modified shape reduces the total harmonics distortion (THD) of the five-level output voltage compared to well-known pulse width modulation (PWM) techniques. The new modulation is used for closed-loop switching of a three-phase multilevel converter with designated maximum power point trackers (MPPT). The unit template control is utilized in this paper to generate reference grid currents. This system feeds active power in variable solar irradiances with enhanced power quality. Simulated results show the effectiveness of the system, and real-time implementation validates its practicability.

Keywords Solar Inverter · Harmonics Analysis · Modulation Scheme · DC-AC Conversion

✉ Shivam Kumar Yadav
shivamyadaviitdelhi@gmail.com

[1] Department of Electrical Engineering, Bhagwan Parshuram Institute of Technology, Rohini, New Delhi 110089, India

[2] Department of Electrical Engineering, Indian Institute of Technology, New Delhi 110016, India

[3] Department of Electrical Engineering, Delhi Technological University, New Delhi 110042, India

Introduction

Multilevel inverters (MLI) are the new trend for grid-integrated renewable energy systems [1]. They have advantages of lower harmonics distortion, lower rating of power electronic devices to generate high voltage levels, lower size filter components, and low electromagnetic interference. MLI provides more than two voltage levels in the stair-case form with single or multiple voltage sources. They are better than traditional two-level voltage source inverters in terms of reduced harmonics in their output voltage and elimination of bulky filters [2]. The low-frequency switching decreases the conduction loss and enhances the inverter efficiency. The multilevel converters divide the total rated power into multiple power converter cells. The power burden of each switch is reduced, and power regulation becomes easier due to the modular approach. Flying capacitor (FC) [3], cascaded H-bridge (CHB) [4], and neutral point clamped (NPC) [5] are the most popular MLI topologies. Various advanced topologies have come to overcome the disadvantages of basic MLI [6].

Traditional MLI configuration employs a large number of switches to generate the voltage levels. Various MLI topologies such as neutral point clamped, flying capacitor, packed U-cell, and cascaded H-bridge inverters have been reported in the literature for different applications. The T-type MLI topology is the latest, most suitable for an induction motor drive and integrating a grid-tied system [7, 8]. This configuration does not require a clamping diode as required in basic NPC inverter topology. It is simple and easier to implement and uses fewer switches. The reduction in the count of switches leads to few driver circuits that reduce the system cost. The well-known SPWM switching scheme requires a high switching frequency to reduce

Published online: 04 January 2022

harmonics in the output voltage [9–11]. Some of the research work has reported selective harmonic elimination (SHE) and space vector modulation (SVM), but these schemes lead to difficult implementation at higher output voltage levels [12, 13]. The reduced carrier and multi-reference techniques are also popular, as discussed in the literature [14]. In the researches [15], a nine-level T-type nested configuration is shown, which has five levels in its phase voltage. The generation of five voltage levels, is possible due to multiple floating capacitors. The voltage balancing of each floating capacitor is essential to have a proper generation of each level. So the advanced PWM technologies have come to regulate them but require extra PI controllers or complex modulation algorithms.

A three-level T-type is used for mitigating the issues of pulse dropping and skipping for motor-generator applications [16]. In the literature [17], two classic H-bridge cells are cascaded to have a 5 level H-NPC converter capable of operating at higher voltage levels. However, this topology requires several isolated DC sources to generate the five levels in the inverter output. Cascaded H-bridge MLI topology exhibits a modular structure that is reliable and simple. CHB MLI can be realized by an asymmetrical structure that requires DC sources of unequal magnitude [18]. These CHB configurations require more DC sources, but the count of conducting switches for realizing various voltage levels is also higher, leading to higher conduction losses. NPC topology for inverters is widely used as it generates multiple voltage levels at the inverter output from the single voltage source. Diodes are used to clamp the voltage to generate multiple voltage levels. Recently a T-type configuration is reported that replaces the clamping diode in the NPC inverter with a bidirectional switch. Modulation techniques have been devloped, which generate the necessary pulse train for switching the inverter switches. Level shifted multicarrier PWM technique is the most widely used technique to generate multiple output levels. In a level shifted PWM, reference sinusoidal voltage is compared with triangular carrier waves. In the investigations [19], triangular carrier waves are replaced by the waves of parabolic function, which has shown better THD performance. Parabolic function for carrier wave is generated by step and sinusoidal functions. A new modified carrier wave for dominant harmonics shift is covered in the literature [20].

In this paper, two different function generators at designated frequencies and gain factors are taken together to create a new wave shape. The harmonics analysis is conducted for a closed-loop operated solar-powered T-type converter. The dynamic performance is analyzed, and the system is validated in real time. Obtained results meet the IEEE 519 standard, and the new switching scheme provides efficient power conversion.

System Configuration

Figure 1 shows a T-type solar converter configuration. This topology uses six switches in each phase. Switches S_1-S_4 form an H-bridge cell, while S_5 & S_6 are connected in series to form a bi-directional switch (Suffix a, b and c are used for switches in phase a, b & c, respectively, as S_{a1}-S_{a6} are used for phase a. This MLI converter generates five voltage levels in each phase. Obtained voltage output is made sinusoidal through the harmonic filter. Here, two PV arrays in each phase are used for solar input, and an inverter output is integrated into the three-phase grid. Two solar panels generate the voltage levels (V_{pv1}+V_{pv2}), V_{pv1}, 0, -V_{pv1}, and – (V_{pv1}+V_{pv2}) in each phase. S_1-S_4 & S_2-S_3 switches of H-bridge generate (V_{pv1}+V_{pv2}) & -(V_{pv1}+− V_{pv2}) levels, while S_5-S_4 and S_6-S_2 switches generate V_{pv1} and -V_{pv1} levels. The switching states are given in Table 1. If V_{con} is five-level output, where x=a,b, and c, then KVL equation of converter and the grid with interfacing elements, is expressed as,

$$V_{conx} - V_x = i_x r_x + L_x \frac{di_x}{dt} \tag{1}$$

where, r_x and L_x are the resistance and inductance values for grid integrated solar converter.

Adopted Control Scheme

Figure 2 shows a closed-loop control strategy of the SPV converter, which ensures maximum solar power extraction from a solar panel at variable irradiances and feeds the power to the three-phase grid. The adopted control is

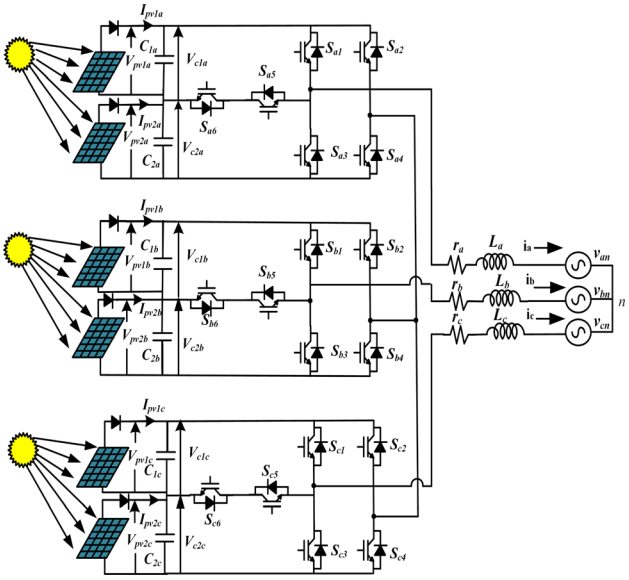


Fig. 1 Five level T-Type SPV system

Table 1 Switching states for t-type SPV converter

S_1	S_2	S_3	S_4	S_5	S_6	Voltage level
On	Off	Off	On	Off	Off	$(V_{pv1} + V_{pv2})$
Off	Off	Off	On	On	Off	V_{pv1}
On	On	Off	Off	Off	Off	0
Off	On	Off	Off	Off	On	$-V_{pv1}$
Off	On	On	Off	Off	Off	$-(V_{pv1} + V_{pv2})$

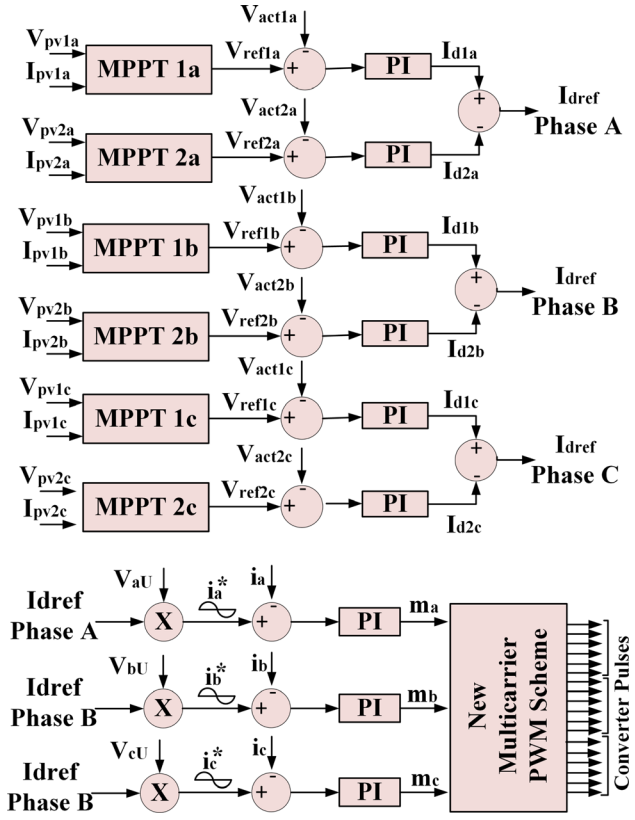


Fig. 2 Control structure for T-type solar PV system

explained in subsections here. The system specifications are shown in Table 2.

Single-Stage MPPT Control

The perturb and observe (P&O) technique is used to track the maximum power in this work. In this technique, the reference voltage (perturb) is varied, and the power variation (observe) is monitored. This technique is executed periodically and incremented or decremented the reference voltage to reach maximum power point (MPP). This MPPT control used here, is a single-stage control, which generates reference for the DC-link capacitor pair in each phase. The voltage regulator equations for voltage balancing of DC-link capacitor pair are expressed as,

$$I_{d1x}(u) = I_{d1x}(u-1) + k_p(e_{c1x}(u) - e_{c1x}(u-1)) + k_i(e_{c1x}(u)) \quad (2)$$

$$I_{d2x}(u) = I_{d2x}(u-1) + k_p(e_{c2x}(u) - e_{c2x}(u-1)) + k_i(e_{c2x}(u)) \quad (3)$$

where,

$$e_{c1x} = V_{ref1x} - V_{act1x};\ e_{c2x} = V_{ref2x} - V_{act2x} \quad (4)$$

The net reference current for each phase is a sum of each balancing current component of corresponding voltage regulators.

$$I_{drefx} = I_{d1x} + I_{d2x} \quad (5)$$

This net reference current balances the capacitor voltages and regulates the active power into the grid as per the irradiance profile of each solar panel.

Unit Template Controller

A unit template control strategy is used to control the T-type controller for grid synchronization. The unit voltage templates of grid voltages are estimated as,

$$V_{aU} = \frac{v_{an}}{v_p}, V_{bU} = \frac{v_{bn}}{v_p}, V_{cU} = \frac{v_{cn}}{v_p} \quad (6)$$

The estimated unit templates are used to generate reference grid currents as,

$$i_a^* = I_{dref,PhaseA} * V_{aU}, i_b^* = I_{dref,PhaseB} * V_{bU}, i_c^* = I_{dref,PhaseC} * V_{cU} \quad (7)$$

The reference grid currents are compared with the actual grid currents to generate modulating signals $m_{a\text{-}c}$. The modulating signals are compared with a new multicarrier scheme to generate the switching for the T-type five-level inverter.

New Multicarrier Modulation Scheme

Sinusoidal PWM (SPWM) is a well-known switching sequence that compares level-shifted high-frequency triangular carriers and reference sine wave. Many researchers have tried different alignments of the carrier wave, such as phase disposition and phase opposition disposition, to

Table 2 System specifications

Symbol	Parameter	Value
V_{pv}	Voltage of solar panel at Maximum power point	197.88 V
I_{pv}	Current of solar panel at maximum power point	31.88 Amps
$(P_{pv})_{phase}$	Power/phase of solar panel at maximum power point	12.616 kW
P_{pv}	Total Power of solar panels at maximum power point	37.850 kW
P_{grid}	Power fed to grid (at 97% overall efficiency)	36.714 kW
V_{gp}	Peak grid phase voltage	338.8 V
I_{gp}	Peak grid Phase Current	72.24 A
L	Interfacing inductor	3.5 mH
$f_{sw\ or}\ f_1$ or $1/T_s$	Switching frequency	500 Hz
K_p,k_i	Gain of Modulating signal PI Controllers	0.008,0.001
K_p, K_i	Gain of DC link controllers	4.9, 0.48

reduce harmonics. In [19], the authors have used parabolic carrier wave in place of triangular, which is generated from two different sources. Inspired by integrating two different sources to have a new SPWM, a modified triangular carrier is generated in this work, which has shown a remarkable improvement in THD. Figure 3 shows a schematic to generate a new multicarrier waveform. The triangular and new carriers are shown in Fig. 4.

Here, a triangular generator G_{tri} with frequency f_1 (500 Hz) and pulse generator G_{pulse} with frequency f_2 (2000 Hz) are integrated with scaling coefficients A, B, and C, which decide the amplitude of the final new carrier I. The A, B, and C values to have five level carrier waves between 0 and 0.5 range are 0.25, 0.25, and 4.55, respectively. The equation of newly generated carrier is formulated as,

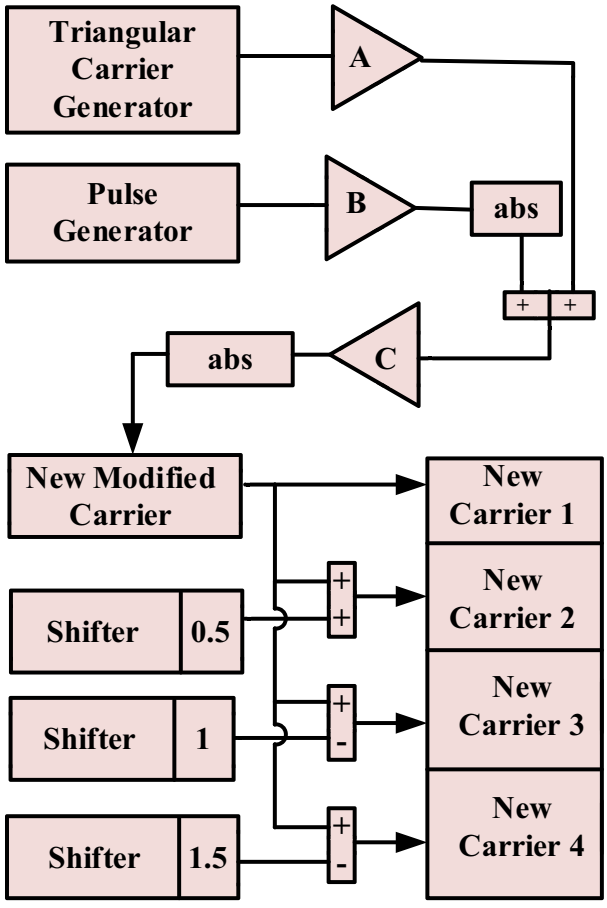


Fig. 3 Generation of new carrier wave

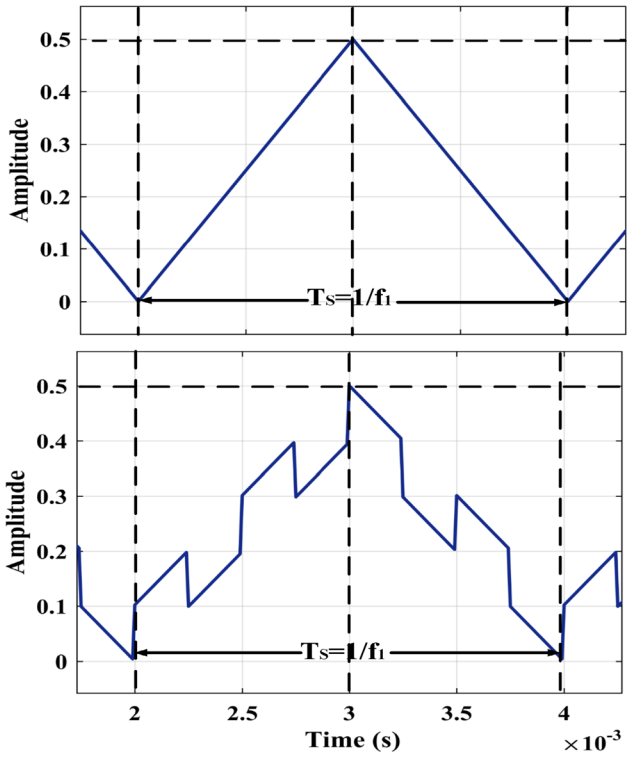


Fig. 4 Triangular and new modified carrier waveform

$$G_{new} = \left|\left(G_{tri}(f_1) * A + \left|G_{pulse}(f_2) * B\right|\right) * C\right| \quad (8)$$

The modified wave-shape generation is achieved with the integration of pulse function. However, the triangular wave frequency modulation index remains the same i.e., shape is changed, but the period T_s remains the same i.e., $1/f_1$ for the new carrier wave. The frequency modulation index is given as,

$$m_f = \frac{Carrier\ \text{Frequency}}{Modulating\ Signal\ Frequency} = \frac{f_1}{f_{\text{mod}}} = \frac{(1/Ts)}{f_{\text{mod}}} \quad (9)$$

Figure 5 validates that there is no change in the time period and the frequency of the carrier is 500 Hz with a frequency modulation index of 500/50, i.e., 10. The obtained shape is used to generate four carrier waves to have five-level T-type converter voltages. Figure 5 shows four new carriers in a level-shifted manner at m_f =10.

It is observed that modified carriers are divided into four parts within the reference sine wave. This new scheme is also applicable to higher levels. The modified shape is obtained with a suitable selection of f_1 and f_2. The change in f_2 changes the shape, while f_1 changes the time period and m_f of the carrier waveforms. Here, a low-frequency carrier wave with 500 Hz switching is utilized for harmonics mitigation. The new shape is achieved with the integration of pulse function at frequency f_2 to have improved harmonics. The presented technique is implemented for the five-level solar converter to have improved system performance.

Results and Discussion

The system is modeled in Matlab/Simulink and implemented on a real-time digital simulator OPAL-RT. The power quality is enhanced with an improved modulation strategy for a T-type five-level converter-based solar PV system. The harmonics of output voltage and grid current are realized for the new multicarrier PWM-based closed-loop control. The presented new modulation-based controller is tested with changes in solar profiles.

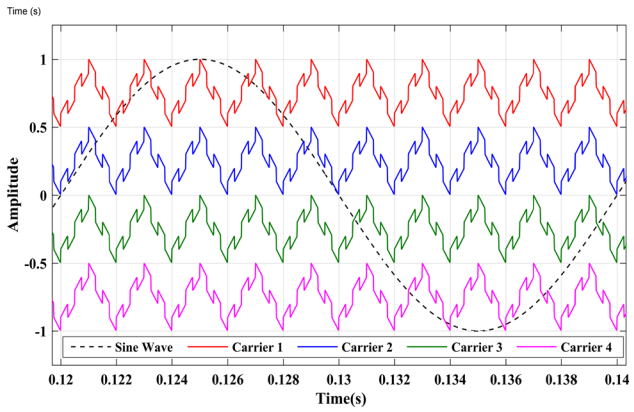


Fig. 5 New multicarrier for five-level output

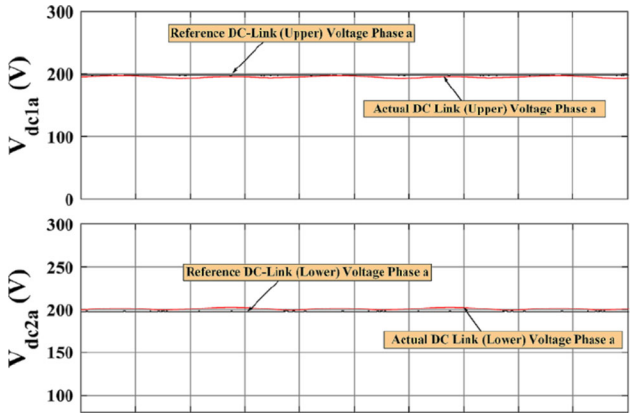


Fig. 6 Reference and actual DC link voltage

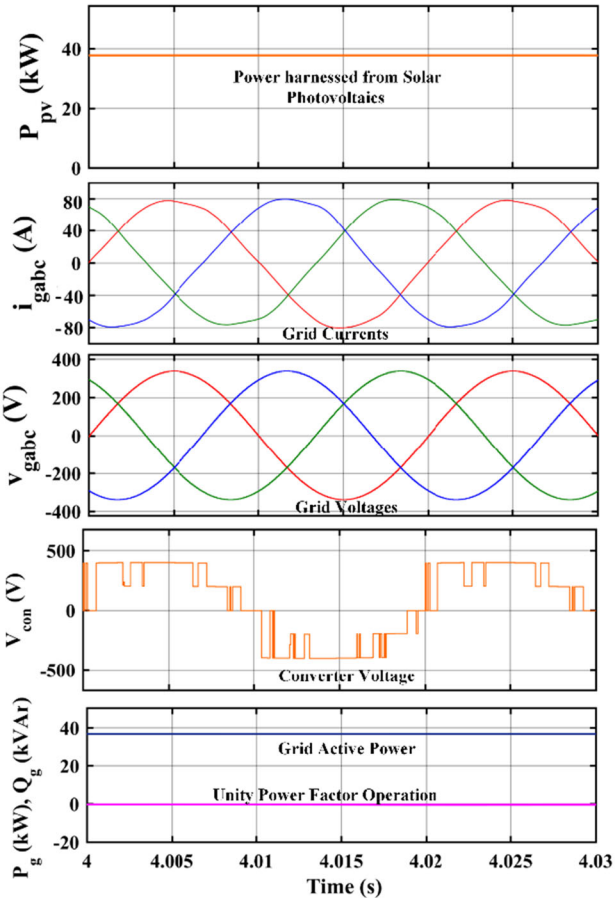


Fig. 7 Steady-state T-type SPV results

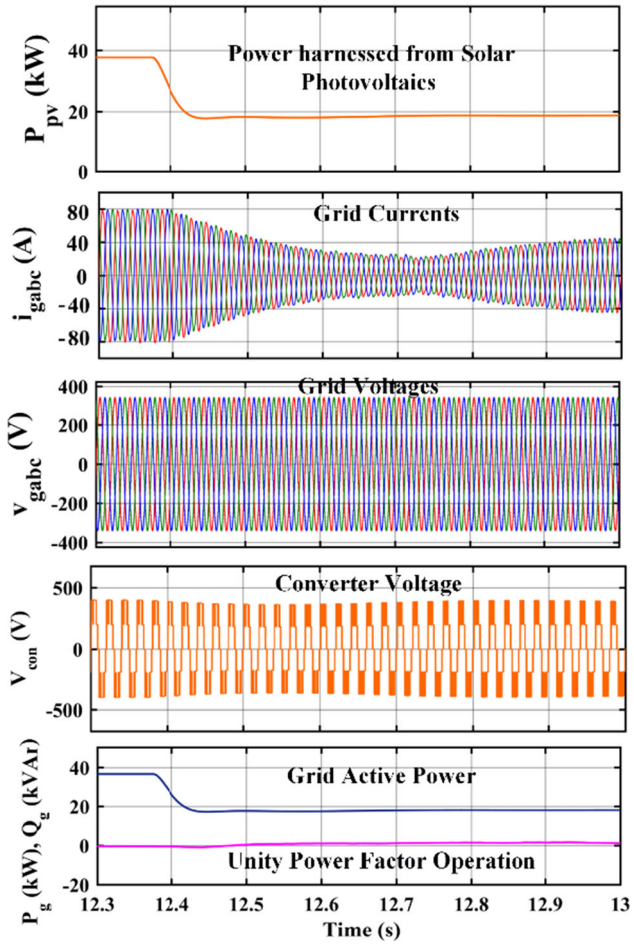


Fig. 8 Dynamic fall of irradiance in SPV system

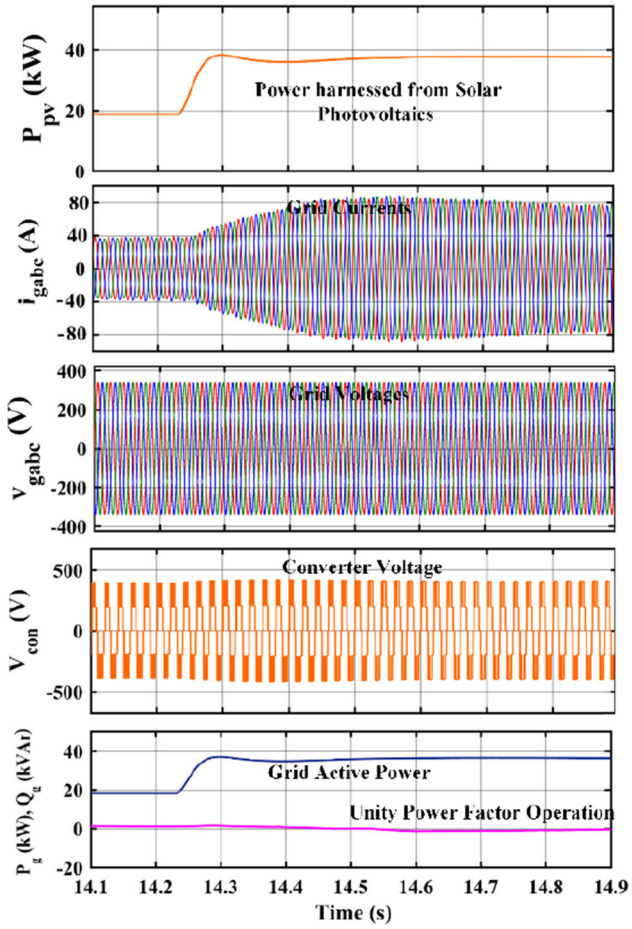


Fig. 9 Dynamic rise of irradiance in SPV system

Steady-State Performance

Figure 6 shows the balanced DC link voltages of phase A. The controllers are finely tuned to achieve a balanced DC link for the capacitors at maximum PV voltage. Figure 7 shows the system performance at an irradiance of 1000 W/m^2. The power of 37.8 kW is transferred, where each phase harnesses the power of 12.6 kW from its pair of solar arrays. The grid power is transferred at unity pf of the order of 36.7 kW with an efficiency of 97%. There is a constant injection of the grid current due to constant irradiance, and grid voltages are balanced. The active power injection is successfully achieved, and zero reactive power ensures the unity power factor operation.

Dynamic Performance

Figures 7-8 show the dynamic results of the T-type MLC based SPV system with the rise and fall of solar irradiance level. In Fig. 8, power fed in the grid is reduced to half, decreasing irradiance to half of the rated level. The grid current reduces and stabilizes at half of the rated current. The grid current injection directly depends on the amount of irradiance projected on the solar PV arrays. The five-level converter voltage is attained with an improved harmonic profile and voltage shape near to sinusoidal. In Fig. 9, harnessed power from the solar panel is increased and fed to the grid at unity pf. Each phase equally contributes to the net power injection for a three-phase system. The converter voltage levels are obtained with an excellent dynamic response. These results are obtained in closed-loop control and improved switching with the new carrier wave, indicating better power quality and performance in intermittent solar conditions.

Mitigation of Harmonics

The harmonics analysis is conducted for this new carrier modulation scheme and well-known sine PWM scheme. A comparison of THD is made for both schemes. The converter voltage THD for SPWM at 500 Hz with unity modulation index is found as 25.96%. However, the new multicarrier PWM has reduced the THD to 23.66%. The improvement in THD is observed at low switching

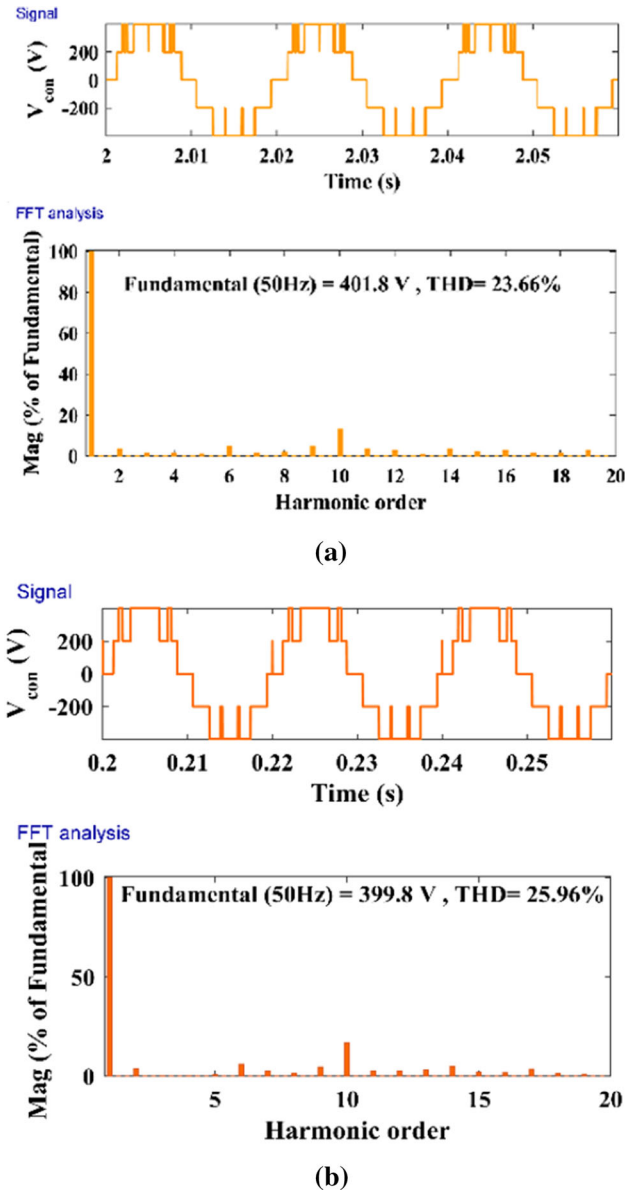


Fig. 10 Harmonic performance of multilevel voltage **a** New multicarrier modulation **b** Conventional SPWM

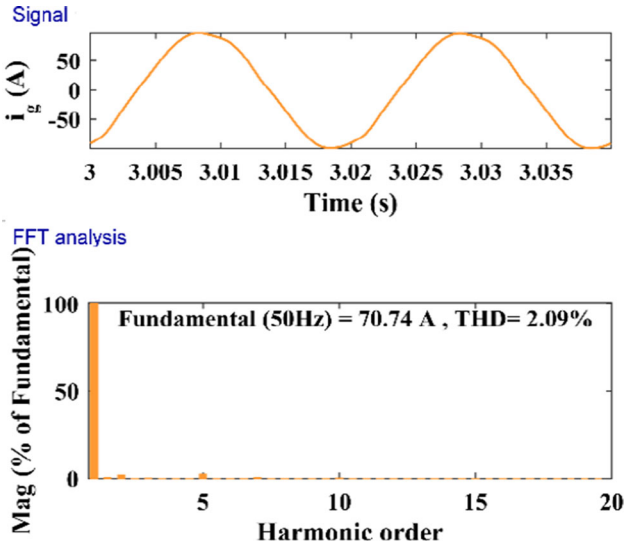


Fig. 11 Harmonic performance of grid current

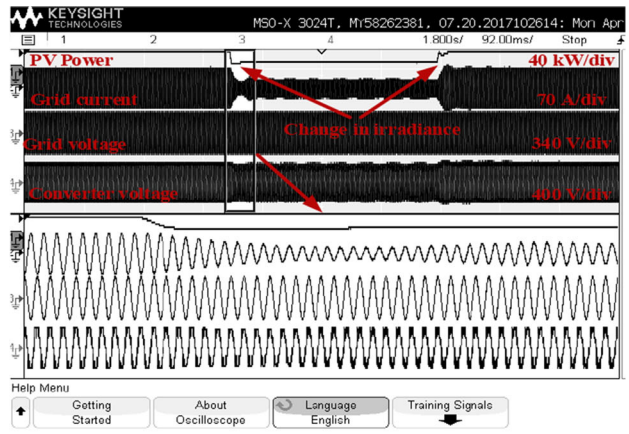
frequency with the inclusion of a new multicarrier PWM scheme. This enhances the power quality and further improves the system performance. The low switching frequency ensures lower switching losses. Figure 10a–b show that the change is significant from 25.96 to 23.66%, making it a better scheme. Moreover, this scheme provides better fundamental components and better harmonic performance. For m_f of 10 at 500 Hz, a dominant harmonic is observed at 10th harmonic order in harmonics spectra of converter voltages. Moreover, the grid current THD is also evaluated for the system. Figure 11 shows the grid current THD, which is 2.09% under the IEEE 519 standard limits. Tables 3 and 4 show the comparative assessment of the new multicarrier scheme with a level-shifted pulse-width modulation (LSPWM) having phase disposition (PD), phase opposition disposition (POD) and alternate phase opposition disposition (APOD) at various modulation indices. Table 3 covers a variation of THD with a change in

Table 3 Modulation index vs THD

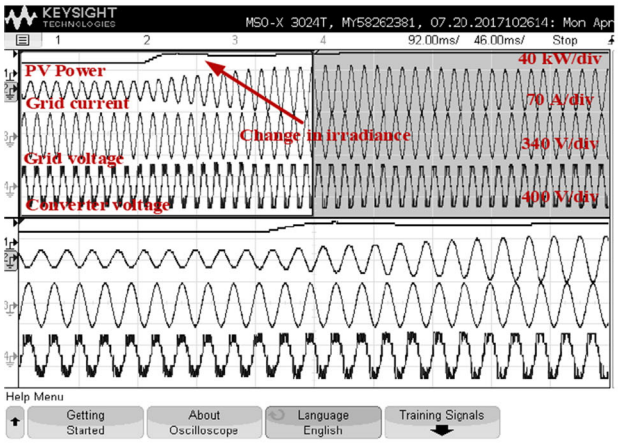
Modulation Index	THD % of Converter Voltage			
	SPWM (PD)	SPWM (POD)	SPWM (APOD)	New Multicarrier PWM Scheme
0.5	52.14	47.82	54.62	46.38
0.6	44.22	41.11	48.9	40.49
0.7	41.40	40.40	44.34	39.24
0.8	38.32	38.37	36.34	35.28
0.9	33.29	32.66	30.39	28.85
1.0	25.96	29.82	27.12	23.66

Table 4 Modulation index vs fundamental component

Modulation Index	Fundamental Voltage Component (V)			
	SPWM (PD)	SPWM (POD)	SPWM (APOD)	New Multicarrier PWM Scheme
0.5	200.2	199.9	199.7	205.9
0.6	239.7	239.1	239.0	239.9
0.7	279.6	279.1	279.3	279.7
0.8	319.3	319.1	319.7	319.8
0.9	360.0	345.8	361.3	365.6
1.0	399.8	388.7	400.1	401.8

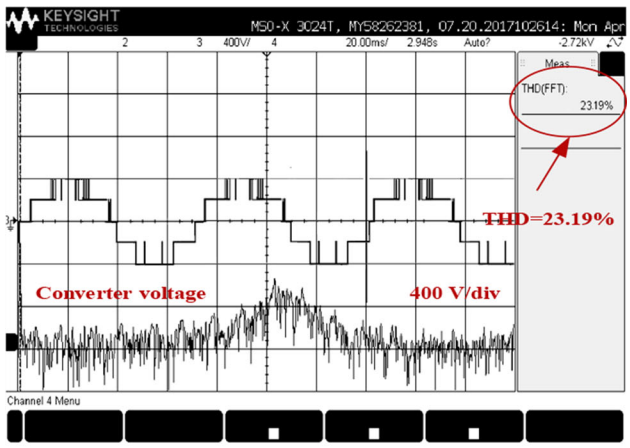


(a)

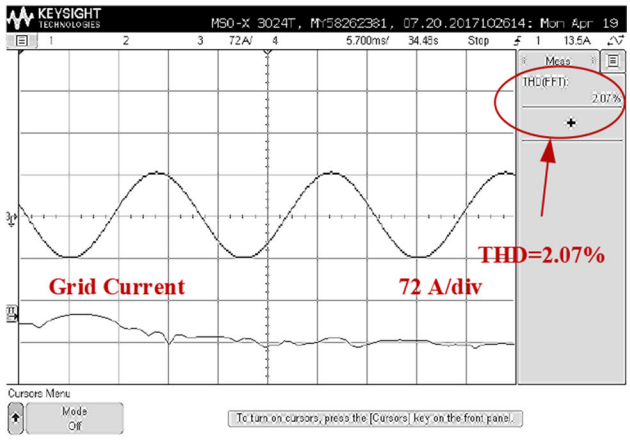


(b)

Fig. 12 Dynamic performance of T-type SPV system **a** fall in irradiance **b** rise in irradiance



(a)



(b)

Fig. 13 Real-time THD **a** Converter voltage **b**)Grid current

modulation within the range of 0.5 to 1. This change comes as solar irradiance changes in a solar PV system. It is observed that for every modulation index, the new multi-carrier-based switching scheme gives better THD and higher fundamental component in the output voltage, as shown in Table 4.

Real-Time Validation

The new scheme is tested in a real-time controller platform with an OPAL-RT real-time simulator. The discrete-time domain control is implemented to test the controller and modulation performance. The real-time core assignation and real-time laboratories (RT-LAB) compilation are done to perform the plant and controller synchronization at a sampling rate of 5e-5s. Figure 12a shows the dynamic performance of the system with a fall in irradiance. The zoomed view of the system's performance shows the effective decrease in power being fed to the grid. The grid current reduces with a fall in the solar irradiance profile. The solar power is fed with a fall in irradiance from 100% to half value. Figure 12b shows the performance of T-type with the rise in irradiance. The power is fed into the grid with a proportional increase of input solar irradiance. The five-level converter voltage is achieved in real-time and remains constant with changes in irradiance.

Figure 13a–b show the performance of harmonics in real-time of the converter voltage of T-type topology. The converter voltage performance is verified in a real-time system with a THD of 23.19%, which is less than a triangular approach for the same switching frequency. The grid current THD is below 5% in real-time at a magnitude of 2.07%. Therefore, this solar PV system is validated in real-time successfully with a new modulation strategy.

Conclusion

The harmonic performance of the T-type multilevel converter is improved with a new multicarrier scheme. This new scheme is generated from two different functions to

provide a new modified wave shape. The frequency modulation index remains the same and better fundamental output voltage is achieved. The low-frequency carrier is presented to switch the converter in closed-loop mode. The dynamic performance of the SPV system is validated in real-time digital simulator (RTDS). This work has given a new approach toward enhancing the power quality of solar PV system with a new modulation strategy. The low-frequency switching has ensured less switching stress and better conversion efficiency of this new solar PV system.

Acknowledgements The authors are grateful to the Department of Science and technology, Govt. of India to facilitate research in Electrical Department at Indian Institute of Technology Delhi under grant RP03195 (FIST Scheme). The authors also acknowledge "SERB National Science Chair Fellowship" for providing research funds.

Funding The authors are grateful to the Department of Science and Technology, Govt. of India for facilitating research in the Electrical Department at IIT, Delhi under grant RP03195 (FIST Scheme) and in part by OPAL-RT under grant RP03253. The authors also acknowledge the "SERB National Science Chair Fellowship" for providing research funds.

Declarations

Conflict of interest The Author(s) declare(s) that there is no conflict of interest.

References

1. J. Yu, G. Zhang, M. Peng, D. Song, M. Liu, Power-matching based SOC balancing method for cascaded H-bridge multilevel inverter. CPSS Trans. Power Electron. Appl. **5**(4), 352–363 (2020)
2. A. Aghazadeh, M. Davari, H. Nafisi and F. Blaabjerg, Grid integration of a dual two-level voltage-source inverter considering grid impedance and phase-locked loop. IEEE J. Emer. Select Top. Power Electron. 9(1): 401-422 (2011)
3. T. Modeer, N. Pallo, T. Foulkes, C.B. Barth, R.C.N. Pilawa-Podgurski, Design of a GaN-based interleaved nine-level flying capacitor multilevel inverter for electric aircraft applications. IEEE Trans. Power Electron. **359**(11), 12153–12165 (2020)
4. Z. Akhmetov, L. Chushan, W. Li, A. Ruderman, A Hybrid Three-Phase Seven-Level CHB Inverter with a Novel Modulation Scheme. 2020 IEEE Applied Power Electronics Conference and Exposition (APEC), New Orleans LA, USA pp. 2451–2454 (2020)
5. Y.J. Kim, S.M. Kim, K.B. Lee, Improving DC-Link Capacitor Lifetime for Three-Level Photovoltaic Hybrid Active NPC Inverters in Full Modulation Index Range. IEEE Trans. Power Electron. **36**(5), 5250–5261 (2021)
6. S.S. Lee, K. Lee, Dual-T-type seven-level boost active-neutral-point-clamped inverter. IEEE Trans. Power Electron. **34**(7), 6031–6035 (2019)
7. X. Wang, Z. Wang, M. Gu, Z. Xu, Z. Zou, W. Wang, M. Cheng, Fault-tolerant control of common electrical faults in dual three-phase pmsm drives fed by T-type three-level inverters. IEEE Trans Ind. Appl. **57**(1), 481–491 (2021)
8. S. Pal, M.G. Majumder, R. Resallayan, K. Gopakumar, U. Loganathan, D. Zielinski, A.R. Beig, A cascaded nine-level inverter topology with T-type and H-bridge with increased DC-bus utilization. IEEE Trans Power Electron. **36**(1), 285–294 (2021)
9. M. Ye, W. Ren, L. Chen, Q. Wei, G. Song, S. Li, Research on power-balance control strategy of CHB multilevel inverter based on TPWM. IEEE Access. **7**, 157226–157240 (2019)
10. V. Jammala, S. Yellasiri, A.K. Panda, Development of a new hybrid multilevel inverter using modified carrier SPWM switching strategy. IEEE Trans. Power Electron. **33**(10), 8192–8197 (2018)
11. H. Yu, B. Chen, W. Yao, Z. Lu, Hybrid seven-level converter based on T-type converter and H-bridge cascaded under SPWM and SVM. IEEE Trans. Power Electron. **339**(1), 689–702 (2018)
12. P. Kala, S. Arora, Implementation of Hybrid GSA SHE Technique in Hybrid Nine-Level Inverter Topology. IEEE J. Emerg. Sel. Top. Power Electron. 9(1),1064-1074 (2021)
13. Y. Liu, S. Liu, C. Wang, J. Jiang, J. Li, S. Ye, An improved space vector modulation with capacitor voltage control for T-type five-level nested neutral point piloted converter. IEEE Trans. Power Electron. **36**(5), 5262–5277 (2021)
14. P.P. Patel, M.A. Mulla, A single carrier-based pulsewidth modulation technique for three-to-three phase indirect matrix converter. IEEE Trans. Pow. Electron. **35**(11), 11589–11601 (2020)
15. A. Bahrami, M. Narimani, A sinusoidal pulsewidth modulation (SPWM) technique for capacitor voltage balancing of a nested T-type four-level inverter. IEEE Trans. Power Electron. **34**(2), 1008–1012 (2019)
16. H. Kim, Y. Kwon, S. Chee, S. Sul, Analysis and compensation of inverter nonlinearity for three-level T-type inverters. IEEE Trans. Power Electron. **32**(6), 4970–4980 (2017)
17. Z. Cheng, B. Wu, A novel switching sequence design for five-level NPC/H-bridge inverters with improved output voltage spectrum and minimized device switching frequency. IEEE Trans. Power Electron. **22**(6), 2138–2145 (2007)
18. S. P. Gautam, S. Gupta, S. Pattnaik, V. Singh, New topology with reduced number of switches in asymmetrical cascaded multilevel inverter, in 5th International Conference on Advances in Recent Technologies in Communication and Computing (ARTCom), Bangalore, pp. 336–344 (2013)
19. M. Sharifzadeh, K. Al-Haddad, Packed E-Cell (PEC) converter topology operation and experimental validation. IEEE Access. **7**, 93049–93061 (2019)
20. S.K. Yadav, N. Mishra, B. Singh, S. Padmanaban, F. Blaabjerg, Modified SPWM technique for improved harmonic performance of single PV array fed grid-tied five-level converter. IET Pow. Electron. **13**, 4498–4506 (2020)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Newly Realized Grounded Capacitance Multiplier Using Single CFDITA

Dheeraj Singh
Department of Electronics & Communication Engineering,
Delhi Technological University, Delhi, India
drajiitr@gmail.com

Deva Nand
Department of Electronics & Communication Engineering,
Delhi Technological University, Delhi, India
devkamboj@dce.ac.in

Atul Kumar
Department of Electronics & Communication Engineering,
Jaypee Institute of InformationTechnology,
Noida, U.P, India
atul.nit304@gmail.com

2021 7th International Conference on Signal Processing and Communication (ICSC) | 978-1-6654-2739-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICSC53193.2021.9673477

***Abstract* — This article presents the new realization of capacitance multiplier based on current follower differential input transconductance amplifier (CFDITA). Capacitance multiplier has been realized by one active block CFDITA, three resistors and one capacitor. Multiplication factor of the realized capacitance multiplier can be controlled by using external resistors as well as by bias current of CFDITA. Ideal & non-ideal analyses have been carried out for the implemented circuit. PSPICE software is used for designing & simulation of circuit in 0.18µm CMOS technology. To power the circuit, ±1.25 V is used. The simulation results are found in accordance with theoretical aspects.**

Keywords—Analog circuit, Capacitance Multiplier, CFDITA, Current Conveyor.

I. INTRODUCTION

In analog integrated circuits, several applications like filters, oscillators, sample and hold circuit require high value capacitors. With the technology advancement, active components size has been reduced to a greater extent but still passive components occupy larger silicon area which makes it difficult to realize the large value of capacitors in an integrated circuit. Capacitance multiplier offers the realization of large capacitor values by using active components which makes the system to work with larger time constants. Voltage-mode (VM) and current-mode (CM) approach using active blocks has been used to realize grounded and floating capacitance multiplier with very few numbers of active and passive components.

In the past, active blocks like operational amplifier (Op-Amp) [1,2] and operational transconductance amplifier (OTA) [3,4] have been used for implementation of capacitance multiplier. CM active blocks such as second generation current conveyor (CCII), current feedback transconductance amplifier (CFTA), current backward transconductance multiplier (CBTA), differential voltage current conveyor (DVCC) and other active blocks have also been used to realize capacitance multiplier [5-11]. In the recent times, several capacitance multipliers using CM active blocks have been presented [12-21]. Floating capacitance multiplier using one fully balanced voltage differencing buffered amplifier (FB-VDBA) has been proposed in [12] which can be varied electronically by changing the bias current. Lossy and lossless capacitance multipliers are realized using two current feedback operational amplifiers (CFOAs) in [13]. To reduce the power consumption, CCII and OTA were biased in subthreshold region to realize a low power capacitance multiplier [14]. A grounded capacitance multiplier is presented in [15] using two CFOAs and two resistors. This circuit avoids the need of matching passive component. Another CFOA based capacitance multiplier is presented using two CFOAs and two resistors in [16]. Two circuits of grounded capacitance multiplier using current differencing transconductance amplifier (CDTA) are implemented in [17] in which the low impedance input terminal serves as a virtual ground for the capacitors. Another active block, DVCC based capacitance multiplier without any resistance requirement is realized in [18]. In [19], for the realization of a grounded capacitance multiplier with electrically controllable multiplication factor, two VDBAs with a single resistor and a single capacitor were employed. Two active blocks negative second generation current conveyor (CCII-) and inverting second generation current conveyor (ICCII) are employed to realize grounded capacitance multiplier with tunable multiplier values depending on external resistor values [20]. Another active building block differential input buffered and transconductance amplifier (DBTA) using grounded resistor and capacitor implemented an electronically tuned capacitance multiplier [21].

This paper proposes a grounded capacitance multiplier using one current follower differential input transconductance amplifier (CFDITA), three resistors and one capacitor. The circuit uses only 13 MOSFETs for its realization. The capacitor can be precisely tuned via external resistors and bias current as well. This paper is organized into five sections. Section II discussed about the active block CFDITA and the equations govern the CFDITA. Section III consists of the circuit of the proposed work i.e. CFDITA based capacitance multiplier and its ideal analysis. The proposed circuit's non-ideal analysis is presented in Section IV. Section V provides the simulation of the proposed capacitance multiplier with time and frequency response analysis and section VI concludes the paper.

II. DESCRIPTION OF THE CFDITA

CFDITA is a versatile active element which has been used to realize several applications like bistable multivibrator, square and triangular waveform generator, and instrumentation amplifier [22-25]. It comprises two different units, one unit is current follower as input stage and the other unit is differential input transconductance amplifier. Symbol of the CFDITA is shown in "Fig.1", where current and voltage relations between the different terminals of CFDITA can be characterized by "(1)" as follows.

978-1-6654-2739-5/21/$31.00©2021 IEEE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:39:32 UTC from IEEE Xplore. Restrictions apply.

$$\begin{bmatrix} I_V \\ I_Z \\ I_{O-} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & -g_m & g_m & 0 \end{bmatrix} \begin{bmatrix} I_F \\ V_V \\ V_Z \\ V_{O-} \end{bmatrix} \quad (1)$$

g_m is the transconductance of CFDITA.

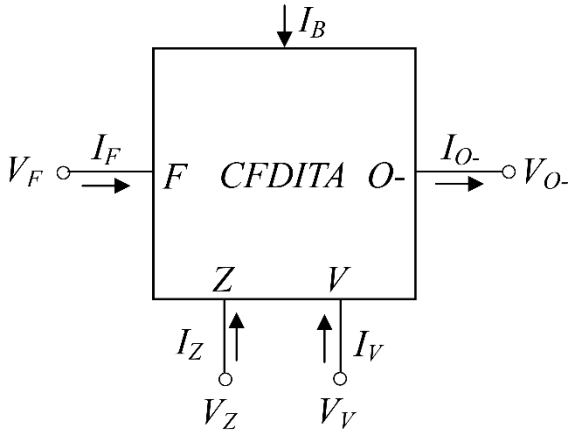


Fig. 1. Symbol of the CFDITA

III. Proposed Circuitof Capacitance Multiplier Using CFDITA

The proposed capacitance multiplier circuit comprises one active block, CFDITA and two grounded and one floating resistors along with one floating capacitor as illustrated in "Fig. 2". The following equations result from an ideal analysis of the circuit depicted in "Fig.2":

$$I_{in} = I_F - I_{O-} \quad (2)$$

$$I_F = -V_{in}sC \quad (3)$$

$$I_{O-} = g_m (V_1 - V_2) \quad (4)$$

$$V_1 = \frac{R_1(R_2+R_3)}{R_1+R_2+R_3} I_F \quad (5)$$

$$V_2 = \frac{R_1R_3}{R_1+R_2+R_3} I_F \quad (6)$$

Using equations "(2)" to "(6)", the impedance can be expressed in the following way:

$$Z_{in} = \frac{V_{in}}{I_{in}} = \frac{1}{sC(1+g_mK)} \quad (7)$$

So the equivalent capacitance of the circuit can be given as

$$C_{eq} = C(1 + g_mK) \quad (8)$$

where, K is a constant given as: $K = \frac{R_1R_2}{R_1 + R_2 + R_3}$. It is observed from "(8)" that multiplication factor is given by $(1 + g_mK)$. External resistors can be used to regulate the multiplication factor. Additionally, it is also electronically controllable via bias current.

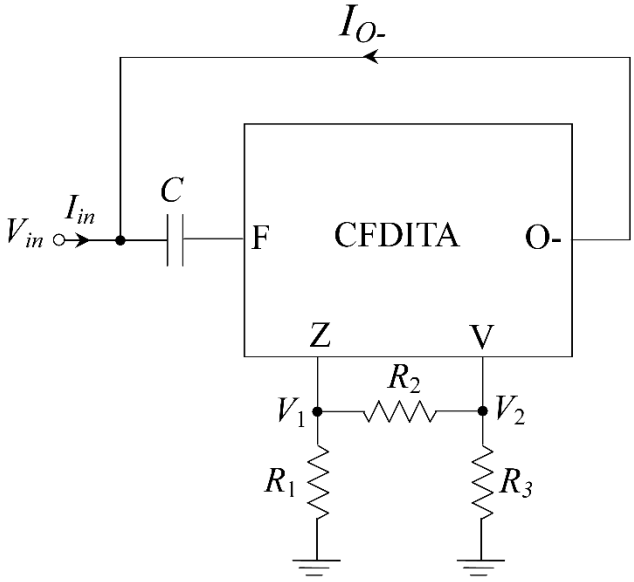


Fig. 2. The Proposed Capacitance Multiplier's Circuit

IV. Non-Ideal Analysis

"Equation (9)" can be used to describe the non-ideal conditions of CFDITA.

$$I_V = 0, I_Z = \alpha I_F \text{ and } I_{O-} = g_m(\gamma_1 V_Z - \gamma_2 V_V) \quad (9)$$

Where, α characterize non-ideal current transfer gain from F to Z terminal and imprecision in transconductance values are given by $\gamma_1$and γ_2.

Considering non-idealities characterized by equation "(9)", the equivalent capacitance can be determined using following equations:

$$I_{in} = I_F - I_{O-} \quad (10)$$

$$I_F = -V_{in}sC \quad (11)$$

$$I_{O-} = g_m (\gamma_1 V_1 - \gamma_2 V_2) \quad (12)$$

$$V_1 = \frac{R_1(R_2+R_3)}{R_1+R_2+R_3} \alpha I_F \quad (13)$$

$$V_2 = \frac{R_1R_3}{R_1+R_2+R_3} \alpha I_F \quad (14)$$

Using equations "(10)" to "(14)" the equivalent impedance can be expressed as

$$Z_{in} = \frac{V_{in}}{I_{in}} = \frac{1}{sC(1+g_mK')} \quad (15)$$

As a result, the circuit's equivalent capacitance can be written as

$$C_{eq} = C(1 + g_mK') \quad (16)$$

where, $K' = \frac{\alpha R_1}{R_1+R_2+R_3}\{\gamma_1(R_2 + R_3) - \gamma_2 R_3\}$.

In "(16)" when the values of non-idealities α, γ_1 and γ_2 approaches unity the value of K' becomes equal to K and multiplication factor is same as given in "(8)".

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:39:32 UTC from IEEE Xplore. Restrictions apply.

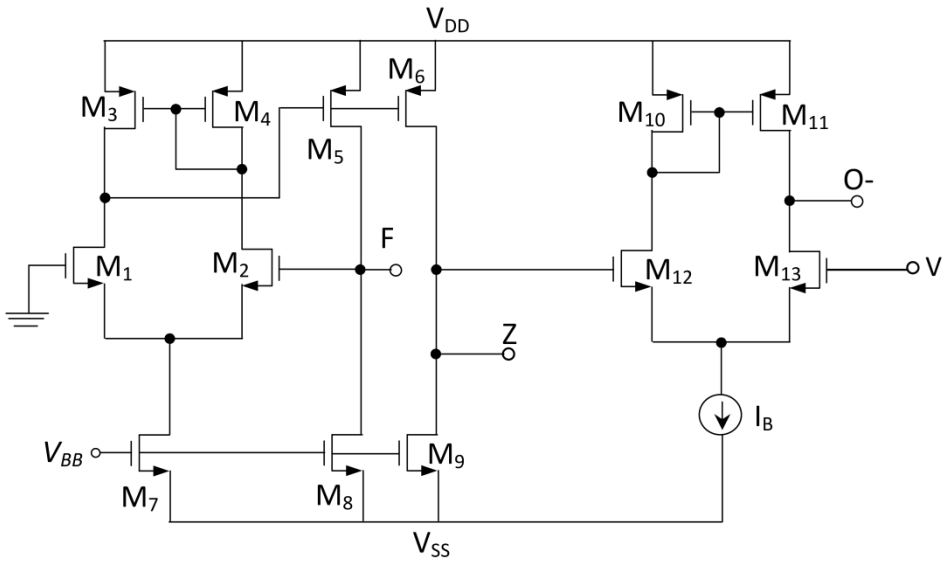


Fig. 3. CFDITA's CMOS realization [22, 23]

V. Simulation Results and Discussions

Simulation of proposed capacitance multiplier is performed on PSPICE simulator. CMOS of CFDITA (shown in "Fig. 3") is utilized to run simulations. The voltages of the power supply are ±1.25 V. In Table I, the W/L ratios are presented. The biasing current is I_B = 140µA, while the value of V_{BB} is fixed at -0.7 V. The value of transconductance for the selected bias current is 1mS.

Table I. W/L ratio of transistors

Transistors	(W/L) μm
M_1– M_2	0.14/0.18
M_3–M_4	1.44/0,18
M_5–M_6	3.6/0.18
M_7–M_9	16.2/0.18
M_{10}–M_{11}	9/0.18
M_{12}–M_{13}	1/0.18

The passive components are chosen R_1 = 2 kΩ, R_2 = 10 kΩ and R_3 = 1 kΩ to provide the multiplication factor of 2.52. The capacitor value is chosen as 100pF. Frequency response of the circuit for Zin is obtained as shown in "Fig. 4". The multiplication factor obtained from the "Fig. 4" is found to be 2.52, which is similar to the theoretical value.

Time response of the circuit has been obtained by applying a sinusoidal voltage source of 100 mV (p-p) and frequency of 1 MHz as shown in "Fig. 5". The current leads by 92° the applied voltages in "Fig. 5". Additionally, to confirm a quadrature relationship between current and applied voltage, a Lissajous pattern is shown in "Fig. 6". The quadrature relationship between current and applied voltage is confirmed by the circular pattern.

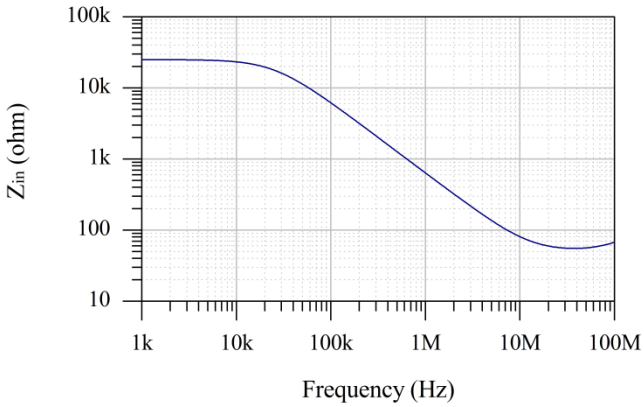


Fig. 4. Simulated AC Response for magnitude of impedance

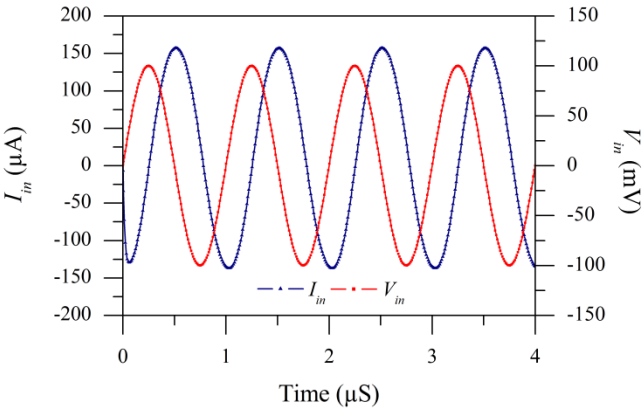


Fig. 5. Simulated Transient Response

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:39:32 UTC from IEEE Xplore. Restrictions apply.

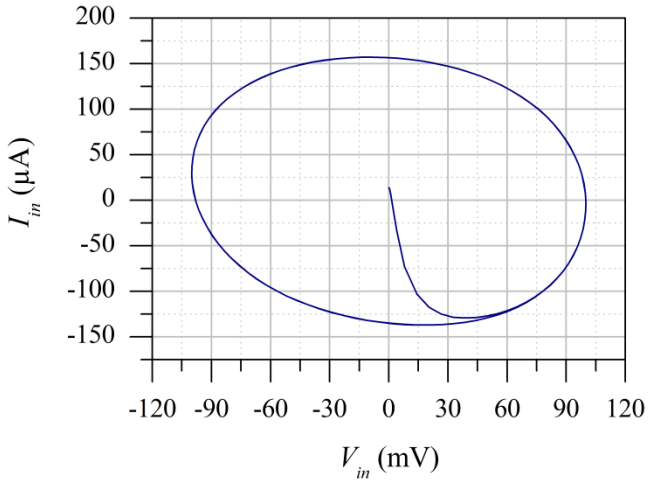


Fig.6.Lissajous pattern between I_{in} and V_{in}

VI. CONCLUSION

A new CFDITA based capacitance multiplier has been implemented and analyzed. It uses single active element which reduces the circuit's complexity. The multiplication factor can be adjusted by altering the external resistor values and alternatively by changing the bias current. Simulations were done on PSPICE 0.18μm CMOS technology to support the theoretical aspect of the proposed circuit. AC response for input impedance and transient waveforms for input voltage and input current are shown in simulations. The transient response shows that current leads the applied voltage.

REFERENCES

[1] De Marcellis, G. Ferri and V. Stornelli, "NIC-based capacitance multipliers for low-frequency integrated active filter applications," Research in Microelectronics and Electronics Conference, 2007, pp. 225-228.

[2] Y. Tang, M. Ismail, and S. Bibyk, "Adaptive Miller capacitor multiplier for compact on-chip PLL filter," Electronics Letters, vol. 39, pp. 43–45, 2003.

[3] W. Jaikla and M. Siripruchyanan, "An Electronically Controllable Capacitance Multiplier with Temperature Compensation," 2006 International Symposium on Communications and Information Technologies, 2006, pp. 356-359.

[4] I.A. Khan and M.T. Ahmed, "OTA-based integrable voltage/current-controlled ideal C multiplier," Electron Letters, vol. 22, pp. 365-366, 1986.

[5] A. Lahiri, "DO-CCII based generalised impedance convertor simulates floating inductance, capacitance multiplier and FDNR," Australian Journal of Electrical and Electronics Engineering, vol. 7, pp. 15-20, 2009.

[6] S. Minaei, E. Yuce and O. Cicekoglu, "A versatile active circuit for realising floating inductance capacitance FDNR and admittance converter," Analog Integrated Circuits and Signal Processing, vol. 47, pp. 199-202, 2006.

[7] M.T. Abuelma'atti and N.A.Tasadduq, "Electronically tunable capacitance multiplier and frequency dependent negative-resistance simulator using the current controlled current conveyor," Microelectronics Journal, vol. 30, pp. 869-873, 1999.

[8] H. Alpaslan and E. Yuce, "Bandwidth expansion methods of inductance simulator circuits and voltage-mode biquads," Journal of Circuits and Systems, vol. 20, pp. 557-572, 2011.

[9] Y. A. Li, "A series of new circuits based on CFTAs," AEU - International Journal of Electronics and Communication, vol. 66, pp. 587-592, 2012.

[10] U.E. Ayten, M. Sagbas, N. Herencsar and J. Koton, "Novel floating general element simulators using CBTA," Radioengineering, vol. 21, pp. 11-19, 2012.

[11] E.Yuce, "A novel floating simulation topology composed of only grounded passive components," International Journal of Electronics, vol. 97, pp. 249-262, 2010.

[12] W. Tangsrirat and O. Channumsin, "Tunable Floating Capacitance Multiplier Using Single Fully Balanced Voltage Differencing Buffered Amplifier," Journal of Communication Technology and Electronics, vol. 64, pp. 797-803, 2019.

[13] E. Ozer, M.E. Basak and F. Kacar, "Realizations of lossy and lossless capacitance multiplier using CFOAs," AEU - International Journal of Electronics and Communications, vol. 127, 153444, 2020.

[14] M. A. Al-Absi and A. A. Al-Khulaifi, "A New Floating and Tunable Capacitance Multiplier With Large Multiplication Factor," IEEE Access, vol. 7, pp. 120076-120081, 2019.

[15] M. Dogan and E. Yuce, "A new CFOA based grounded capacitance multiplier," AEU - International Journal of Electronics and Communications, vol. 115, pp. 153034, 2020.

[16] R. Verma, N. Pandey and R. Pandey, "Novel CFOA based capacitance multiplier and its application," AEU - International Journal of Electronics and Communication, vol. 107, pp. 192-198, 2019.

[17] D. Biolek, J. Vavra and A. Keskin, "CDTA-Based Capacitance Multipliers," Circuits Systems and Signal Processing, vol. 38, pp. 1466–1481, 2019.

[18] J. Vavra, "Floating Capacitance Multiplier Based on DVCC," in 41st International Conference on Telecommunications and Signal Processing (TSP), 2018, pp. 1-4.

[19] P. Moonmuang, T. Pukkalanun and W. Tangsrirat, "VDBA-Based Electronically Tunable Capacitance Multiplier with a Grounded Capacitor," in 15th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), 2018, pp. 485-488.

[20] A. Yesil, E. Yuce and S. Minaei, "Grounded capacitance multipliers based on active elements," AEU - International Journal of Electronics and Communications, vol. 79, pp. 243-249, 2017.

[21] J. Vavra, "A capacitance multiplier based on DBTA," in IEEE Nordic Circuits and Systems Conference (NORCAS): NORCHIP and International Symposium of System-on-Chip (SoC), 2017, pp. 1-5.

[22] A. Kumar and B. Chaturvedi, "Novel CMOS CFDITA and its application as electronically tunable bistable multivibrator," in International Conference on Signal Processing and Communication (ICSC), 2016, pp. 374-379.

[23] B. Chaturvedi and A. Kumar, "CMOS MO-CFDITA Based Fully Electronically Controlled Square/TriangularWave Generator with Adjustable Duty Cycle," IET Circuits, Devices and Systems, vol.12, pp. 817-826, 2018.

[24] B. Chaturvedi and A. Kumar, "Electronically tunable current-mode instrumentation amplifier with high CMRR and wide bandwidth," AEU-International Journal of Electronics and Communications, vol. 92, pp. 116-123, 2018.

[25] A. Kumar and B. Chaturvedi, "Fully electronically controllable Schmitt trigger circuit with dual hysteresis," Electronics Letters, vol. 53, pp. 459-461, 2017.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:39:32 UTC from IEEE Xplore. Restrictions apply.

Eur. Phys. J. Plus (2022) 137:110
https://doi.org/10.1140/epjp/s13360-021-02269-z

THE EUROPEAN PHYSICAL JOURNAL PLUS

Regular Article

Numerical simulation of analog metrics and parasitic capacitances of GaAs GS-GAA FinFET for ULSI switching applications

Bhavya Kumar[a], **Rishu Chaujar**[b]

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

Received: 11 September 2021 / Accepted: 6 December 2021
© The Author(s), under exclusive licence to Società Italiana di Fisica and Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract This paper explores the efficacy of Gallium Arsenide (GaAs) as a fin material on the analog metrics and parasitic capacitances of Gate Stack Gate-All-Around (GS-GAA) FinFET. Besides, capacitance-related parameters such as gain–bandwidth product (GBP) and transconductance frequency product (TFP) are also evaluated to analyze the device switching and DC performance. The simulated results validate that compared to Si, GaAs as a fin material increases the switching (I_{on}/I_{off}) ratio by $\sim 10^3$ times and reduces the leakage current, DIBL, and SS by $\sim$ 99%, $\sim$ 50%, and $\sim$ 26%, respectively. Moreover, for GaAs, the peak value of GBP and TFP increases by 9.93 and 11.40 times compared to Si due to the considerable decrease in the parasitic capacitances. Thus, replacing Si with GaAs as a fin material significantly improves the short-channel effects (SCEs) of the device (I_{off}, DIBL, and SS) and capacitances-related parameters (C_{gs}, C_{gd}, C_{gg}, GBP, and TFP). Besides, the impact of gate length (L_g), fin height (H_{Fin}), and fin width (W_{Fin}) of GaAs GS-GAA FinFET on analog metrics and parasitic capacitances are assessed for low power ULSI switching applications. The simulated results reveal that the SCEs and parasitic capacitances improve considerably with decreased L_g, H_{Fin}, and W_{Fin}.

1 Introduction

To meet the ULSI industry's demand, the dimensions of CMOS devices are continuously scaled down and are reaching their fundamental limits. At such small dimensions, short-channel effects (SCEs) such as subthreshold swing (SS), gate leakage current (I_{off}), drain induced barrier lowering (DIBL), etc., start dominating and degrading the device performance [1–4]. Several device structures such as multi-gate MOSFET [5], cylindrical gate MOSFET [6–8], Recessed Channel (RC) MOSFET [9–11], and FinFETs [12, 13] have been proposed to reduce the problems associated with device scaling. FinFET emerged as the most capable device to effectively overcome the SCEs while continuing to downsize CMOS devices to obey the ITRS roadmap's projection [14]. To further escalate the FinFET performance and subthreshold characteristics, gate-all-around (GAA) structure was proposed. As

[a] e-mail: bhavyakumar_phd2k18@dtu.ac.in

[b] e-mail: chaujar.rishu@dtu.ac.in (corresponding author)

Published online: 12 January 2022

the gate covers the channel from all four sides, the channel's electrostatic control shoots up for GAA FinFET [15]. Accordingly, due to the improved subthreshold characteristics, the GAA FinFET device can be scaled down further.

But, considering the scaling limits of FinFETs, further downscaling will be very demanding due to many practical limitations, like parasitic capacitance, threshold voltage roll-off, DIBL, etc. Unsurprisingly, further enhancement in transistor performance and speed while downscaling the device will be possible using new semiconductor materials. Given future logic applications, III-V compound semiconductor materials are the encouraging contender among the new materials [16]. Compared to silicon, GaAs demonstrates many high-caliber electrical properties, for instance, large energy bandgap, high electron mobility, high device on-currents at the low power supply, and a simple hetero-structure approach in microelectronic devices [17, 18]. But, there are some technical challenges associated with using these new wide bandgap semiconductor materials. It includes the need to exploit these materials to their full potential, compete directly against existing silicon power devices, and further development and commercialization. The major hurdle in GaAs-bases devices is (1) infeasible monolithic integration with silicon digital circuitries, (2) no consistent scaling technology driven by a well-planned roadmap [19], and (3) the unavailability of thermodynamically stable, high-quality insulators on GaAs, which can complement the device standards as SiO_2 on Si. However, after years of endeavors, it has been feasible to form a high-quality dielectric on III-V semiconductors with atomic layer deposition (ALD) [20] and molecular beam epitaxy (MBE) [21, 22].

The gate leakage current in a device can be reduced significantly by using a larger energy bandgap gate material which provides a higher potential barrier than other materials of the same thickness. Aluminum oxide (Al_2O_3) is a highly preferable gate dielectric because of its high thermal stability, good interface quality on GaAs, large energy bandgap ~ 9 eV, and it remains amorphous under normal processing conditions [20, 23]. However, the dielectric constant of Al_2O_3 is small and insufficient for aggressive effective oxide thickness (EOT) scaling [24]. To further scale down the EOT, Gate Stack (GS) configuration [25] of high-k dielectric HfO_2 with dielectric constant $\kappa = 25$ [26] and Al_2O_3 $\kappa = 9$ [27] has been formed on GaAs by ALD [28]. Thus, we put forward a Gate Stack Gate-All-Around (GS-GAA) FinFET considering all the considerations.

As the device dimension scales down, the parasitic capacitance becomes more prominent, affecting the performance of the device and making it incompatible for use in low power and switching applications. Thus, there is a necessity to investigate the C-V (capacitance–voltage) measurements besides I-V (current–voltage) measurements to analyze the impact of parasitic capacitances on the performance [29, 30]. In this article, a comparative assessment of parasitic capacitances and analog metrics of GS-GAA FinFET has been performed by undertaking Si and GaAs as fin materials. It is found that on using GaAs as fin material, SCEs and parasitic capacitances are reduced significantly.

The rest of the paper is organized as follows: Section 2 delineates the 3D structure of the device. Section 3 focuses on the simulation framework along with the calibration of experimental and simulation data. Section 4 analyzes the impact of GaAs on device performance in terms of parasitic capacitances and short-channel effects. Further, the impact of gate length, fin height, and fin width variation on parasitic capacitances and analog metrics of GS-GAA FinFET is investigated. Section 5 authenticates the paper's originality with concluding remarks.

2 Device structure

The 3-dimensional systematic structure of the proposed GS-GAA FinFET is displayed in Fig. 1a. The horizontal and vertical 2-dimensional view of GaAs GS-GAA FinFET with parasitic capacitances is shown in Fig. 1b and c, respectively. The channel region is uniformly doped to 1×10^{16} cm^{-3} (n-type), whereas drain and source regions are uniformly heavily doped to 5×10^{18} cm^{-3} (n-type) to reduce the parasitic capacitances [31]. The oxide thickness and gate length are maintained at 1 nm and 7 nm, respectively. Al_2O_3 and HfO_2 are considered in equal proportions to stack the gate oxide, and all the default device parameters are mentioned in Table 1. During the GaAs simulations, width quantization property is followed by maintaining W_{Fin} at a fixed dimension in multiple of H_{Fin} [4, 32]. In the fin region, GaAs and silicon material are incorporated to evaluate GaAs impact on the SCEs and parasitic capacitances. Gate–source voltage (V_{gs}) and drain–source voltage (V_{ds}) are varied from 0 to 1 V and 0 to 0.5 V, respectively. TiN metal gate is taken into account in place of the polysilicon gate because of its compatibility with CMOS processing, low resistivity, high purity, and thermal stability [33–35].

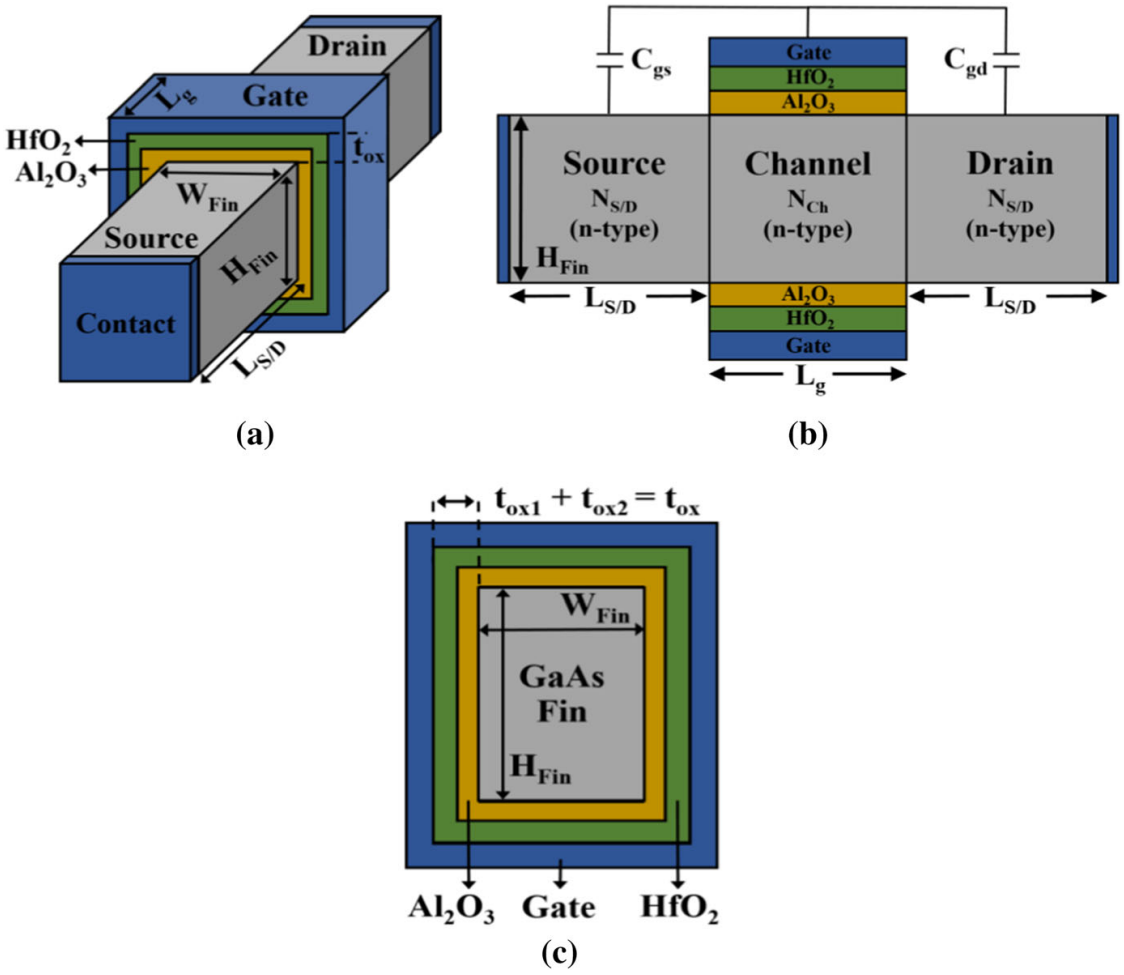


Fig. 1 **a** The systematic simulated 3-D structure **b** horizontal **c** vertical 2-D view of GaAs GS-GAA FinFET with parasitic capacitances

Table 1 Default device parameters of GaAs GS-GAA FinFET

Device parameters	GaAs GS-GAA FinFET
Gate–source voltage (V_{gs})	1.0 V
Drain–source voltage (V_{ds})	0.5 V
Source/drain doping ($N_{S/D}$)	5×10^{18} cm^{-3} (n-type)
Channel doping (N_{Ch})	1×10^{16} cm^{-3} (n-type)
Oxide thickness (t_{ox})	1 nm
Gate length (L_g)	7 nm
Fin height (H_{Fin})	10 nm
Fin width (W_{Fin})	5 nm
Source/drain regions length ($L_{S/D}$)	10 nm
Work function (ϕ_m)	4.65 eV (TiN)

3 Model validation

3.1 Simulation framework

The proposed device GS-GAA FinFET has been simulated by the SILVACO Atlas 3D simulator [36]. The general framework for device simulation is provided by Continuity and Poisson equations. However, secondary equations are also required to obtain more accurate and realistic results. In the aggressively scaled devices, quantum confinement effects become very prominent and cannot be neglected. Thus, Bohm Quantum Potential (BQP) model with gamma = 1.4 and alpha = 0.3 is employed to incorporate the quantum confinement effects [37, 38]. To include the generation and recombination effects, the Shockley–Read–Hall recombination model is implemented with a 1×10^{-7} s fixed carrier lifetime [36]. Fermi–Dirac statistics are included to account for the properties of highly doped materials. To consider both direct and indirect tunneling of electrons between conduction and valence band, the band-to-band Klaassen tunneling model is employed. The concentration-dependent mobility model is enabled to correlate the low-field carrier mobility at 300 K to the impurity concentration. Crowell-Sze impact ionization and bandgap narrowing models are also applied. Further, we have implemented the block and newton's methods to perform all the mathematical carrier transport equations [36].

3.2 Experimental calibration and fabrication feasibility

As the emphasis of this paper is on the effectiveness of GaAs as a fin material, we have extracted the published results of Ye et al. [39] without changing the device parameters to authenticate the simulations. The experimental and simulated output characteristics of Al_2O_3/GaAs MOSFET at $V_{gs} = -1.0$ V and $V_{gs} = -0.5$ V are plotted in Fig. 2a. Besides, the All-Around-Gate (AAG) Si FinFET is calibrated with the experimental data extracted from Hyunjin Lee et al. [40] to validate the above-discussed physical models. The transfer and output characteristics of the same device are portrayed in Fig. 2b and c, respectively. The selection of simulation models is validated due to the close agreement between the experimental and simulated data sets.

Figure 2d exhibits the fabrication feasibility of the GaAs GS-GAA FinFET using a step-by-step fabrication process flowchart. The first step is the GaAs film thinning, and then fin patterning is performed by the self-aligned quadruple patterning (SAQP) technique [41].

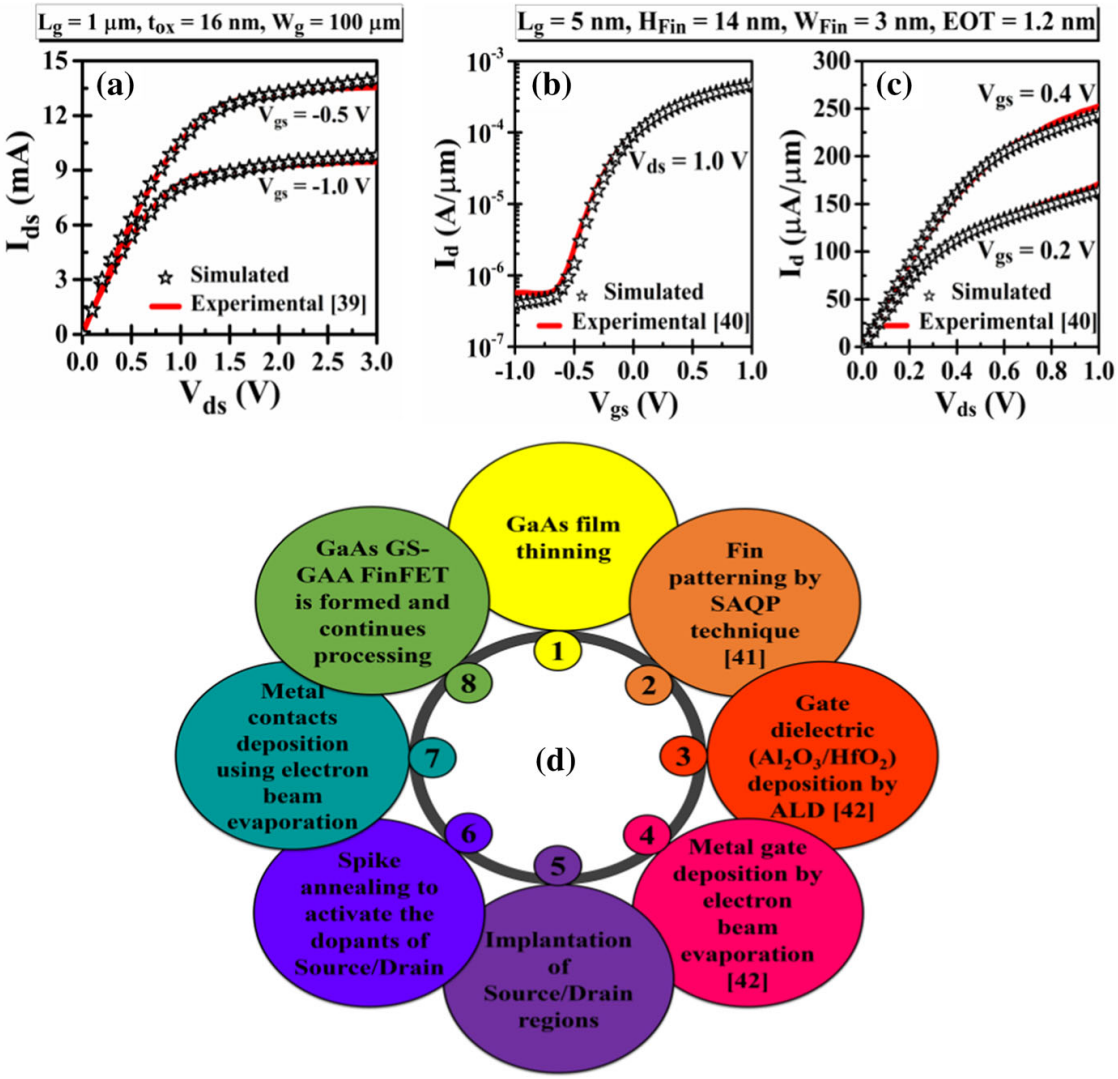


Fig. 2 **a** Experimental and simulated output characteristics of 1 μm gate length Al_2O_3/GaAs MOSFET [39] **b** transfer **c** output characteristics of calibrated AAG Si FinFET [40] **d** fabrication process flowchart of proposed GaAs GS-GAA FinFET

Using the atomic layer deposition (ALD), the gate dielectric (Al_2O_3/HfO_2) deposition is executed on the GaAs interfacial facial [42]. Afterward, a metal gate (TiN) is grown at room temperature on the top of the gate insulator by electron beam evaporation [42]. The drain and source regions are implanted, and the dopants of these regions are activated using spike annealing. By electron beam evaporation, drain/source metal contacts are deposited, followed by a lift-off process. This completes the process of manufacturing the proposed device.

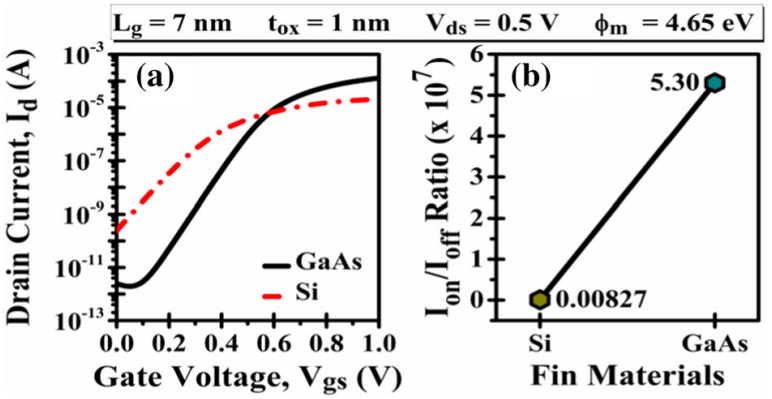


Fig. 3 Variation in (**a**) transfer characteristics **b** switching ratio of Si and GaAs GS-GAA FinFET

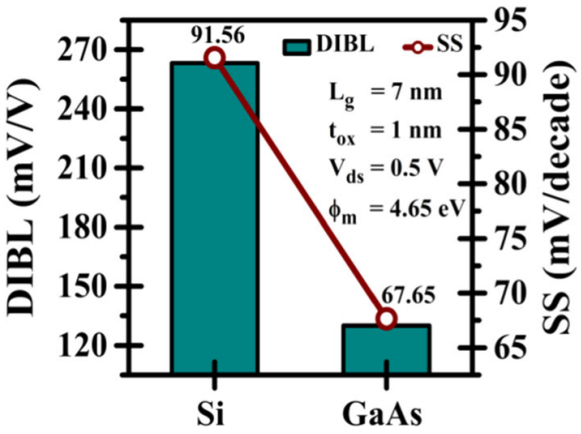


Fig. 4 Variation of DIBL and SS for Si and GaAs GS-GAA FinFET

4 Results and discussion

4.1 Impact of GaAs on analog metrics and parasitic capacitances

Figure 3a and b represents the comparison of $I_d - V_{gs}$ characteristics in log scale and switching (I_{on}/I_{off}) ratio for Si and GaAs GS-GAA FinFET, respectively. The drain current (I_d) enhances and leakage current (I_{off}) reduces appreciably, due to which I_{on}/I_{off} ratio increases by $\sim 10^3$ times for GaAs. The I_{off} is found to be in the range of 10^{-12} and 10^{-10} for GaAs and Si, respectively. GaAs large energy bandgap and high electron mobility is the primary reason for this significant reduction in the leakage current. The study of DIBL and SS becomes essential for aggressively scaled devices. DIBL extracted using Eq. (1) signifies the control of drain bias on the channel region's potential barrier, whereas SS provides perception about the leakage currents related to device characteristics. The lower the value of DIBL and SS, the better will be the SCEs. Figure 4 exhibits the collaborated plot of DIBL and SS for both Si and GaAs. DIBL reduces by ~50% from 263.19 mV/V to 130.04 mV/V, while a ~26% reduction is observed in SS for GaAs.

$$\text{DIBL} = \frac{\left|(V_{th})_{V_{ds}=1.0V} - (V_{th})_{V_{ds}=0.1V}\right|}{(1.0 - 0.1)} \tag{1}$$

Figure 5a reflects the variation of parasitic capacitances (C_{gs}, C_{gd} and C_{gg}) against gate—source voltage. It is observed that the parasitic capacitances increase with the increase in V_{gs} due to the aggregation of charge carriers close to the gate. The collaborated plot of C_{gs}, C_{gd}, and C_{gg} against drain–source voltage is shown in Fig. 5b. The charge carriers accumu-

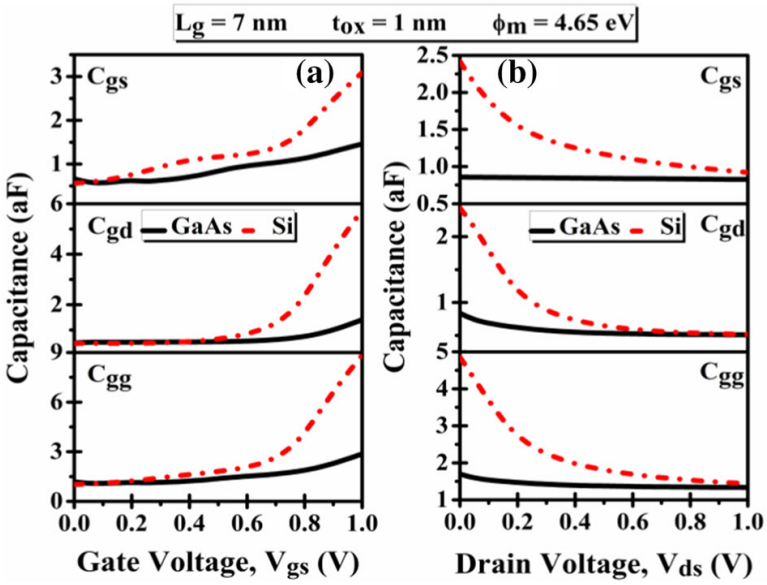


Fig. 5 Change in parasitic capacitances of Si and GaAs GS-GAA FinFET against **a** gate–source voltage **b** drain–source voltage

late near the source side with the increase in V_{ds}, and to conserve charge neutrality, drain capacitance decreases, and consequently, gate capacitance decreases. Moreover, the value of parasitic capacitances decreases considerably for GaAs than Si due to its higher energy bandgap and electron mobility. Thus, higher switching speed is acquired, and consequently, delay time decreases, making GaAs GS-GAA FinFET a suitable device for ULSI switching applications.

Mathematically expressed in Eq. (2) [43], gain–bandwidth product (GBP) is the product of the bandwidth and gain of an amplifier. The peak value of GBP denotes the frequency at which the device acquires maximum gain [30]. Figure 6a outlines the variation of GBP against gate bias for both GaAs and Si, and it is noticed that GaAs exhibits a higher peak value of GBP (7.35 THz) compared to Si (0.74 THz). It is due to the significant improvement in transconductance and drain capacitance.

$$\text{GBP} = \left(\frac{g_m}{2\pi \times 10 \times C_{gd}}\right) \tag{2}$$

$$\text{TFP} = \left(\frac{g_m}{I_d}\right) \times \left(\frac{g_m}{2\pi(C_{gs} + C_{gd})}\right) \tag{3}$$

Transconductance frequency product (TFP) is an essential capacitance-dependent FOM as defined in Eq. (3) [44]. Principally, it is the product of cutoff frequency and device efficiency. TFP exhibits an agreement between bandwidth and power and is mainly utilized in high-speed designs [45, 46]. Figure 6b portrays the change in TFP against gate bias for both GaAs and Si. It can be seen that TFP increases with the increase in V_{gs} in the subthreshold region and then achieve a maximum value in the inversion region. It starts declining due to an increment in C_{gg} with further enhancement in V_{gs}, i.e., in the deep inversion region. Besides, the peak value of TFP improves from 21.84 THz/V to 249.18 THz/V for GaAs due to the parasitic capacitance reduction and transconductance increment.

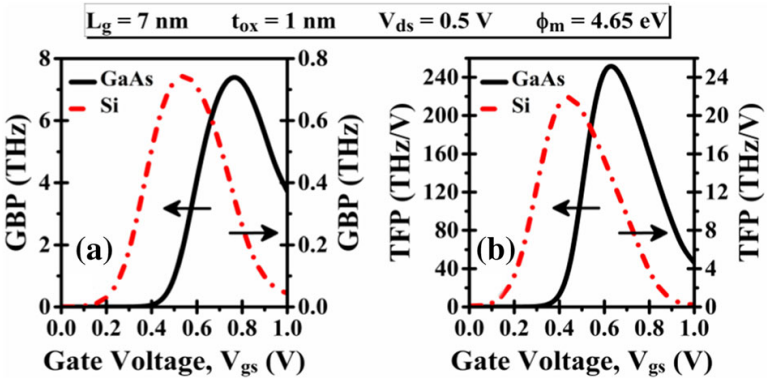


Fig. 6 Plot of **a** GBP **b** TFP against V_{gs} for Si and GaAs GS-GAA FinFET

Thus, as portrayed in Table 2, replacing Si with GaAs significantly improves the device analog metrics (I_{on}/I_{off} ratio, I_{off}, DIBL, and SS) and capacitance-related parameters (C_{gs}, C_{gd}, C_{gg}, TFP, and GBP). Thereby, on account of these advantages, the impact of gate length, fin height, and fin width of GaAs GS-GAA FinFET on analog metrics and parasitic capacitances has been explored further in the paper.

4.2 Impact of gate length (L_g) on analog metrics and parasitic capacitances

To investigate the analog metrics and parasitic capacitances characteristics concerning gate length, the GaAs GS-GAA FinFET device has been simulated for 5, 7, and 10 nm gate lengths. Figure 7a and b shows the device transfer characteristics in the log scale and switching ratio for mentioned gate lengths. The I_d is almost the same, while the I_{off} increases with the gate length reduction. I_{off} current increases by 150% when gate length reduces from 10 to 7 nm, while a further 301.22% increase in I_{off} is observed when gate length changes from 7 to 5 nm. This increase in the leakage current with gate length reduction results in the decrement of the switching ratio. The I_{on}/I_{off} ratio reduces by 89.24% with L_g reduction from 10 to 5 nm. Figure 8a represents the variation of DIBL against different gate lengths. It has been observed from the simulated results that, like I_{off}, DIBL also increases with the gate length reduction. When we reduce the L_g from 10 to 7 nm, DIBL enhances by 36.75%, and for 7 to 5 nm L_g reduction further 23.95% increase is noticed in DIBL. The SS plot concerning mentioned gate lengths is portrayed in Fig. 8b, and a similar trend for SS also follows, although the increase is not significant. A 6.11% increase in SS is obtained when L_g changes from 10 to 7 nm, while a further 2 nm reduction in L_g results in an 8.81% enhancement of SS. Thus, as represented in Table 3, a decrease in gate length increases the SCEs appreciably.

Figure 9a depicts the combined plot of C_{gs}, C_{gd}, and C_{gg} against gate–source voltage for different gate lengths. It is evident that parasitic capacitances decrease with the decrease in the gate length. The same movement is observed in the previously published results as well [47, 48]. C_{gs}, C_{gd}, and C_{gg} reduce by 18.07%, 12.50%, and 15.38%, respectively, with the shrinking of gate length from 10 to 7 nm. Besides, when L_g changes from 7 to 5 nm, C_{gs} decreases by 13.10%, C_{gd} by 10.71%, and C_{gg} by 12.23%. The change in parasitic capacitances against drain–source voltage for mentioned gate lengths is shown in Fig. 9b. In V_{ds} variation for different L_g, a similar trend is observed with minimal decrement. When we decrease L_g from 10 to 7 nm, C_{gs} reduces by 2.89%, C_{gd} by 0.58%, and C_{gg} by 2.02%. Also, for a further 2 nm reduction in L_g (7 to 5 nm), a 3.04%, 1.35%, and 2.39% decrease in C_{gs}, C_{gd}, and C_{gg} are obtained. Thus, parasitic capacitances reduce in GaAs GS-GAA

Table 2 Summary of different analog metrics and parasitic capacitances for Si and GaAs GS-GAA FinFET

Device	Parameters								
	I_{off} (A)	I_{on}/I_{off} ratio	DIBL (mV/V)	SS (mV/dec)	C_{gs} (aF)	C_{gd} (aF)	C_{gg} (aF)	GBP (THz)	TFP (THz/V)
Si GS-GAA FinFET	2.48×10^{-10}	8.27×10^{4}	263.19	91.56	3.08	5.74	8.82	0.74	21.84
GaAs GS-GAA FinFET	2.45×10^{-12}	5.30×10^{7}	130.04	67.65	1.45	1.40	2.86	7.35	249.18

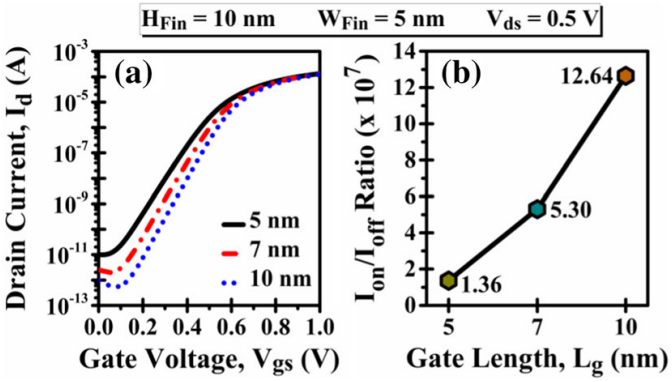


Fig. 7 Variation of **a** transfer characteristics **b** switching ratio against different gate lengths

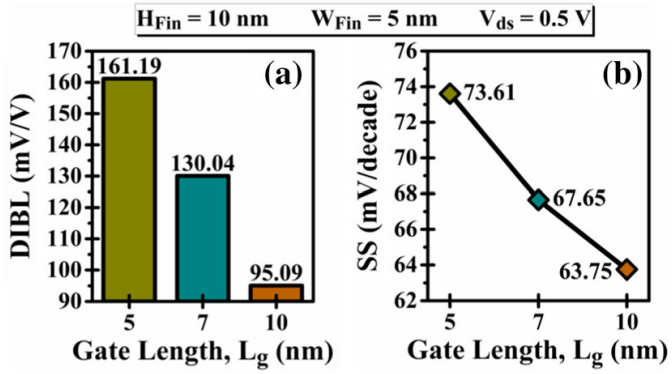


Fig. 8 Plot of **a** DIBL **b** SS against various gate lengths

FinFET with gate length shrinkage. Consequently, the proposed device exhibits enhanced circuit density in IC fabrication. Figure 10a and b reflects the change in GBP and TFP with gate bias for different gate lengths, and an opposite behavior is observed in GBP and TFP with the decrease in the L_g. The drain capacitance and transconductance reduce significantly with L_g reduction, leading to an increment in GBP and a decrement in TFP.

4.3 Impact of fin height (H_{Fin}) on analog metrics and parasitic capacitances

In this section, the GaAs GS-GAA FinFET device has been simulated for 5, 10, and 15 nm fin heights to explore its impact on analog metrics and parasitic capacitances characteristics. The change in the $I_d - V_{gs}$ characteristics in the log scale and switching ratio of the device for mentioned fin heights is reflected in Fig. 11a and b. The device leakage current improves appreciably with the decrease in the fin height from 15 to 5 nm, but a slight reduction in drain current is observed. I_{off} current decreases by 45.91% when H_{Fin} changes from 15 to 10 nm, while a further 5 nm decrease in H_{Fin} (10 to 5 nm) results in an 82.45% improvement in I_{off}. This improvement in I_{off} current leads to a significant increase in the I_{on}/I_{off} ratio. It is increased by 253.93% when H_{Fin} changes from 15 to 5 nm. The plot of DIBL against different fin heights is displayed in Fig. 12a, and a similar trend also follows for DIBL. DIBL improves by 15.68% when we reduce the H_{Fin} from 15 to 10 nm. An additional 15.54% improvement in DIBL is observed for change in H_{Fin} from 10 to 5 nm. Figure 12b shows the change in SS with the change in fin height, and it has been observed from the simulated results that, like I_{off} and DIBL, SS also decreases with the fin height reduction. When we reduce the H_{Fin} from 15 to 10 nm, SS changes from 68.56 mV/dec to 67.65 mV/dec, and for 10 to 5 nm H_{Fin} reduction, SS reduces to 65.49 mV/dec, which is closest to the ideal value of SS (60 mV/dec). Thus, reducing the fin height of the device significantly improves the SCEs characteristics, as displayed in Table 4.

Table 3 Summary of different analog metrics and parasitic capacitances of GaAs GS-GAA FinFET for various gate lengths

Gate lengths	Parameters								
	I_{off} (A)	I_{on}/I_{off} ratio	DIBL (mV/V)	SS (mV/dec)	C_{gs} (aF)	C_{gd} (aF)	C_{gg} (aF)	GBP (THz)	TFP (THz/V)
5 nm	9.83×10^{-12}	1.36×10^{7}	161.19	73.61	1.26	1.25	2.51	7.40	232.25
7 nm	2.45×10^{-12}	5.30×10^{7}	130.04	67.65	1.45	1.40	2.86	7.35	249.18
10 nm	0.98×10^{-12}	12.6×10^{7}	95.09	63.75	1.77	1.60	3.38	7.19	253.30

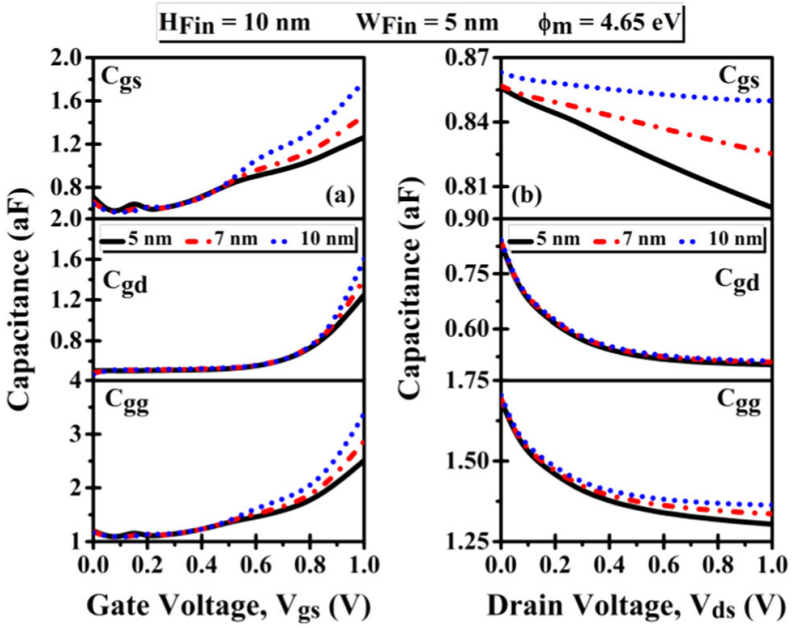


Fig. 9 Change in parasitic capacitances against **a** gate–source voltage **b** drain–source voltage for mentioned gate lengths

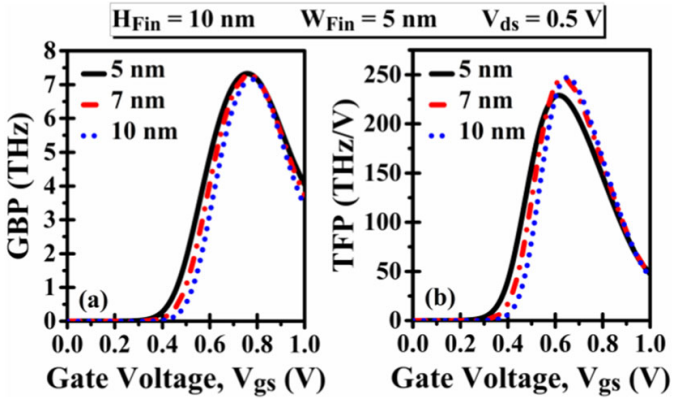


Fig. 10 Variation of **a** GBP **b** TFP against different gate lengths

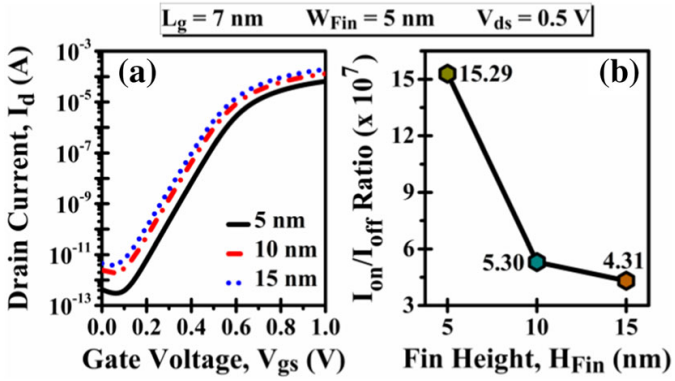


Fig. 11 Variation of **a** transfer characteristics **b** switching ratio against different fin heights

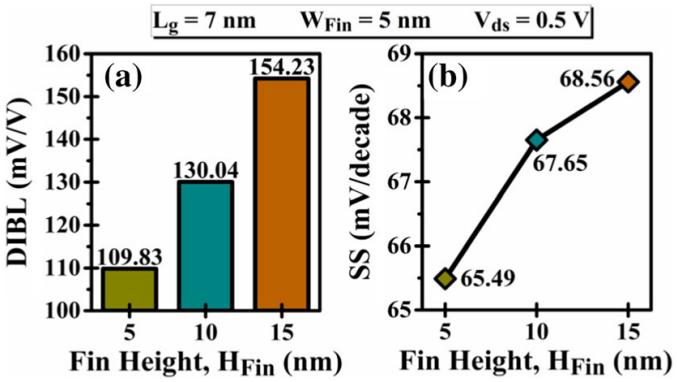


Fig. 12 Plot of **a** DIBL **b** SS against various fin heights

The collaborated plot of parasitic capacitances (C_{gs}, C_{gd}, and C_{gg}) against V_{gs} for different fin heights are represented in Fig. 13a. It is visible that parasitic capacitances decrease notably with the decrease in fin height. When we decrease the H_{Fin} from 15 to 10 nm, C_{gs} reduces by 32.24%, C_{gd} by 33.01%, and C_{gg} by 32.54%. Also, for a further 5 nm reduction in H_{Fin} (10 to 5 nm), a 46.89%, 47.85%, and 47.55% decrease in C_{gs}, C_{gd}, and C_{gg} are obtained. Figure 13b shows the alteration in the parasitic capacitance with change in V_{ds} for mentioned fin heights. C_{gs}, C_{gd}, and C_{gg} reduce by 32.6%, 32.58%, and 32.59%, respectively, with the shrinking of H_{Fin} from 15 to 10 nm. When H_{Fin} changes from 10 to 5 nm, C_{gs} decreases by 48.84%, C_{gd} by 48.61%, and C_{gg} by 48.75%. Thus, with shrinkage in fin height from 15 to 5 nm, parasitic capacitances reduce in the proposed device. Figure 14a and b exhibits the variation in GBP and TFP with V_{gs} for different fin heights, and the peak value of both the parameters increases with the reduction in the H_{Fin}. When we reduce the H_{Fin} from 15 to 5 nm, the GBP peak value obtained at $V_{gs} = 0.75$ V shifts from 7.34 THz to 7.58 THz, and the peak value of TFP obtained at $V_{gs} = 0.65$ V changes from 241.72 THz/V to 272.70 THz/V. This increase in GBP and TFP is due to the enhanced g_m and reduced parasitic capacitances with fin height.

4.4 Impact of fin width (W_{Fin}) on analog metrics and parasitic capacitances

The GaAs GS-GAA FinFET has been evaluated for different fin widths (3, 5, and 7 nm) to evaluate the impact of fin width on analog metrics and parasitic capacitances characteristics. Figure 15a and b represents the variation in $I_d - V_{gs}$ characteristics in the log scale and switching ratio for mentioned fin widths. It has been observed from the simulated results that the drain current reduces slightly, and the leakage current improves significantly with the fin width reduction. When W_{Fin} changes from 7 to 5 nm and from 5 to 3 nm, the I_{off} current decreases by 84.66% and 93.46%, respectively. This significant improvement in I_{off} current leads to a massive enhancement in the I_{on}/I_{off} ratio. Figure 16a portrays the variation in DIBL for different fin widths. It has been observed that DIBL decreases with the decrease in fin width from 7 to 3 nm. When we reduce the W_{Fin} from 7 to 5 nm, DIBL improves from 152.05 mV/V to 130.04 mV/V, and when W_{Fin} reaches 3 nm, DIBL further reduces to 108.97 mV/V. The SS plot against mentioned fin widths is depicted in Fig. 16b, and like I_{off}, DIBL, SS also decreases with the decrease in the W_{Fin}. A 4.39% decrease in SS is obtained when W_{Fin} changes from 7 to 5 nm, and an additional 4.22% improvement in SS is observed when W_{Fin} changes from 5 to 3 nm. Thus, as presented in Table 5, SCEs characteristics of the GaAs GS-GAA FinFET improve appreciably with a decrease in fin width.

Table 4 Summary of different analog metrics and parasitic capacitances of GaAs GS-GAA FinFET for various fin heights

Fin heights	Parameters								
	I_{off} (A)	I_{on}/I_{off} ratio	DIBL (mV/V)	SS (mV/dec)	C_{gs} (aF)	C_{gd} (aF)	C_{gg} (aF)	GBP (THz)	TFP (THz/V)
5 nm	0.43×10^{-12}	15.3×10^{7}	109.83	65.49	0.77	0.73	1.50	7.58	272.70
10 nm	2.45×10^{-12}	5.30×10^{7}	130.04	67.65	1.45	1.40	2.86	7.35	249.18
15 nm	4.53×10^{-12}	4.31×10^{7}	154.23	68.56	2.14	2.09	4.24	7.34	241.72

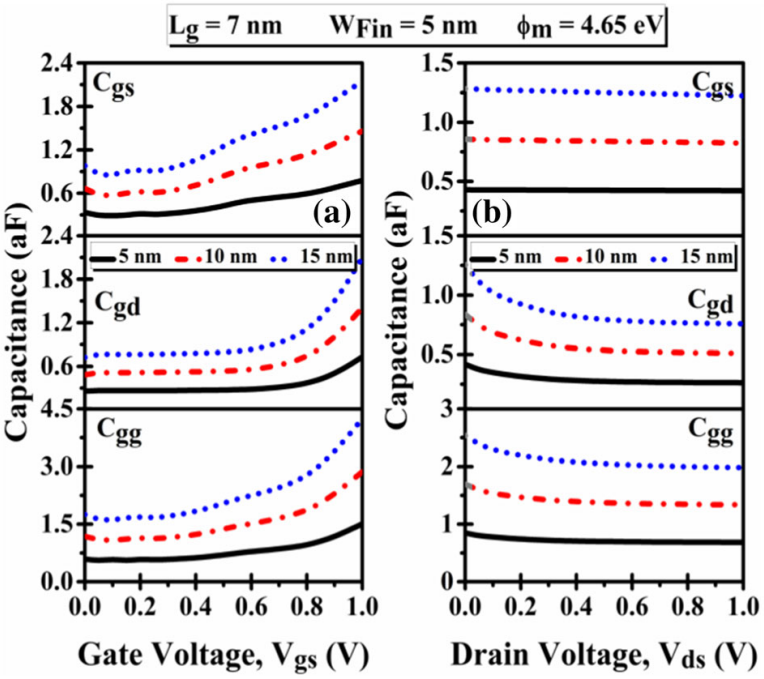


Fig. 13 Change in parasitic capacitances against **a** gate–source voltage **b** drain–source voltage for mentioned fin heights

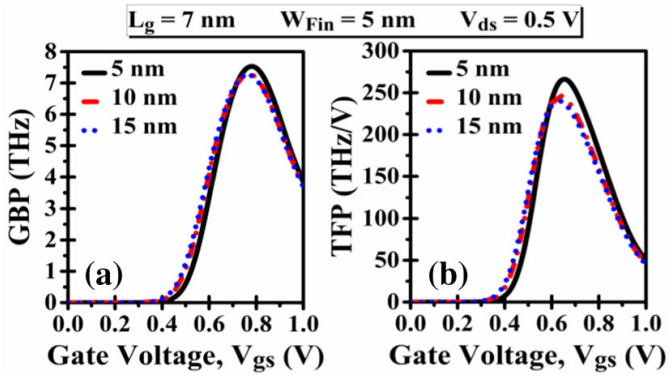


Fig. 14 Variation of **a** GBP **b** TFP against different fin heights

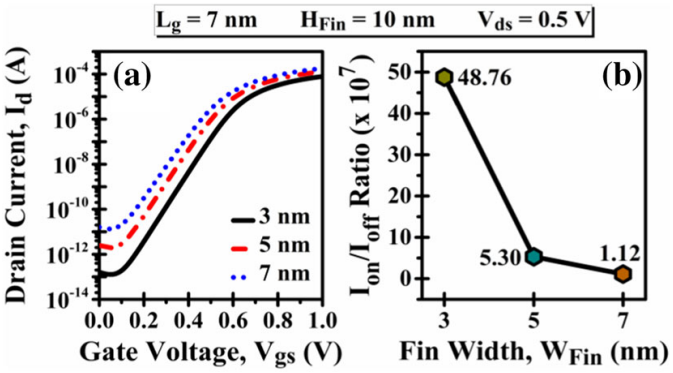


Fig. 15 Variation of **a** transfer characteristics **b** switching ratio against different fin widths

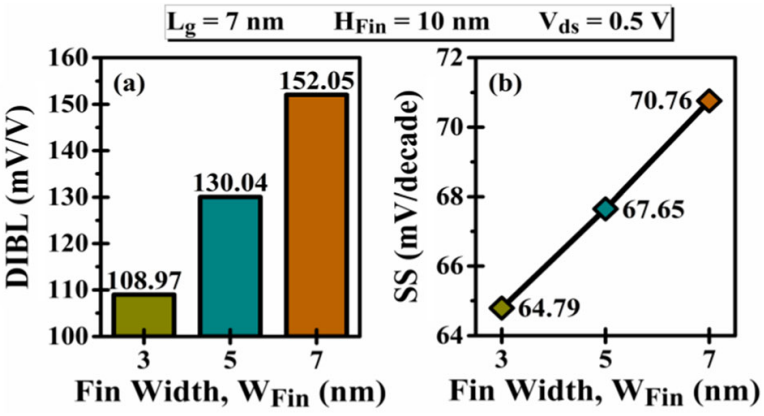


Fig. 16 Plot of **a** DIBL **b** SS against various fin widths

Figure 17a outlines the variation of parasitic capacitances against gate–source voltage for different fin widths. It is observed that parasitic capacitances decrease significantly with the decrease in the fin width. The shrinking of W_{Fin} from 7 to 5 nm leads to a decrease of 23.28%, 24.32%, and 23.52% in C_{gs}, C_{gd}, and C_{gg}, respectively. Besides, when W_{Fin} changes from 5 to 3 nm, C_{gs} decreases by 33.79%, C_{gd} by 36.42%, and C_{gg} by 35.31%. The variation of parasitic capacitances as a function of drain–source voltage and fin widths is shown in Fig. 17b. C_{gs}, C_{gd}, and C_{gg} reduce by 26.19%, 26.54%, and 26.32%, respectively, with the reduction of W_{Fin} from 7 to 5 nm. C_{gs} decreases by 38.22%, C_{gd} by 37.82%, and C_{gg} by 38.07% when W_{Fin} changes from 5 to 3 nm. Thus, a decrease in fin width significantly improves the parasitic capacitances characteristics of the GaAs GS-GAA FinFET. The change in GBP and TFP with gate bias for different fin widths is shown in Fig. 18a and b. The peak value of both the parameters increases and shifts toward the higher gate bias with a reduction in the fin width. GBP and TFP peak value increases because of the reduction in parasitic capacitances and increment in transconductance. The peak value of GBP shifts from 7.15 THz to 7.81 THz when we reduce the W_{Fin} from 7 to 3 nm. Similarly, the peak value of TFP increases from 229.27 THz/V to 286.31 THz/V. In addition, we have compared the analog metrics and parasitic capacitances obtained in this work with several different device structures at fixed 10 nm gate length, as shown in Table 6. It can be seen that the results obtained in this work stand out from the recently published work. The reduction in SCEs, parasitic capacitances, and increment in GBP and TFP is very significant compared to other papers.

5 Conclusion

This paper explores the analog metrics and parasitic capacitances characteristics of Gate Stack Gate-All-Around (GS-GAA) FinFET with GaAs as a fin material. We have also analyzed certain capacitance-related FOMs like GBP and TFP for switching applications. It is noticed from the simulated results that the use of GaAs as a fin material significantly reduces the

Table 5 Summary of different analog metrics and parasitic capacitances of GaAs GS-GAA FinFET for various fin widths

Fin widths	Parameters								
	I_{off} (A)	I_{on}/I_{off} ratio	DIBL (mV/V)	SS (mV/dec)	C_{gs} (aF)	C_{gd} (aF)	C_{gg} (aF)	GBP (THz)	TFP (THz/V)
3 nm	0.16×10^{-12}	48.7×10^{7}	108.97	64.79	0.96	0.89	1.85	7.81	286.31
5 nm	2.45×10^{-12}	5.30×10^{7}	130.04	67.65	1.45	1.40	2.86	7.35	249.18
7 nm	15.9×10^{-12}	1.12×10^{7}	152.05	70.76	1.89	1.85	3.74	7.15	229.27

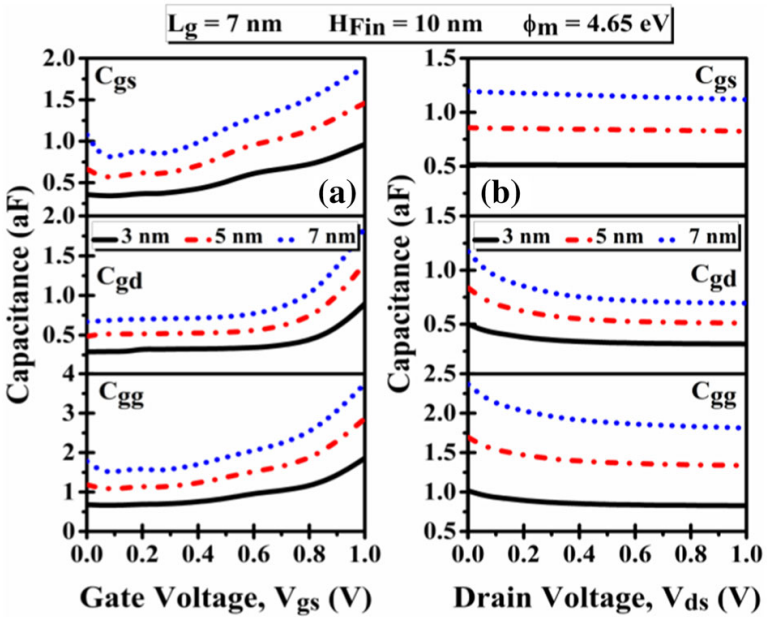


Fig. 17 Change in parasitic capacitances against **a** gate–source voltage **b** drain–source voltage for mentioned fin widths

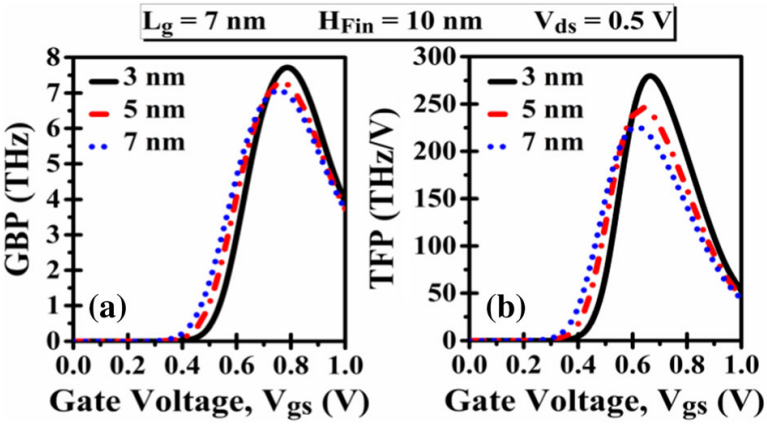


Fig. 18 Variation of **a** GBP **b** TFP against different fin widths

Table 6 Comparison of various analog metrics and parasitic capacitances of different devices at a fixed gate length, $L_g = 10$ nm

References	Year	DIBL (mV/V)	SS (mV/dec)	C_{gs} (fF)	C_{gd} (fF)	C_{gg} (fF)	GBP (THz)	TFP (THz/V)
[49]	2017	N.A	79.8	0.38	3.62	3.99	0.00039	N.A
[50]	2018	99.79	114.60	N.A	N.A	0.97	N.A	3.78
[51]	2019	N.A	N.A	5.36	1.01	6.37	0.37	2.22
[52]	2020	42.95	74.63	0.07	0.07	0.14	N.A	2.18
[43]	2020	N.A	N.A	0.0043	0.0019	0.0059	1.61	45.70
[53]	2021	81.27	71.87	0.04	0.04	0.08	0.05	0.01
[54]	2021	98.48	65.6	0.27	0.24	0.52	0.46	9.48
Present work	–	95.09	63.75	0.0017	0.0016	0.0033	7.19	253.30

The data which is not available is mentioned as N.A

SCEs and parasitics capacitances of the device. The I_{off} current approximately reduces by 100 times, DIBL becomes half, and SS decreases by ~26% for GaAs compared to Si. The parasitic capacitances follow the same trend, with C_{gs} reduced by 52.92%, C_{gd} by 75.60%, and C_{gg} by 67.57%. Due to this considerable decrease in the parasitic capacitances, the peak value of GBP and TFP increases by 893.24% and 1040.93%, respectively. Moreover, parameters like L_g, H_{Fin}, and W_{Fin} are varied to enhance the device performance further. It is found that with the decrease in the L_g, H_{Fin}, and W_{Fin}, the SCEs and parasitic capacitances decrease appreciably. Thus, GaAs GS-GAA FinFET architecture with $L_g = 5$ nm, $H_{Fin} = 5$ nm, and $W_{Fin} = 3$ nm can be considered a suitable candidate for low power ULSI switching applications.

Acknowledgements For providing all the facilities to carry out this research work, the authors would like to thank the Microelectronics Research Laboratory, Delhi Technological University.

Authors' contribution All authors contributed to the study's conception and design.

Funding This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data Availability Statement This manuscript has associated data in a data repository. [Authors' comment: The authors mentioned above have all the relevant data associated with this research work and will be dedicated to share that if asked to do so in the future.].

Declarations

Conflict of interests The authors declare that they have no known conflict of interests or personal relationships that could have appeared to influence the work reported in this paper.

Ethical approval The authors have seen all the Ethical Standards and will suppose to follow them in the future.

Consent for participate Since the concerned research paper is based on the 'non-life science journal.' So, 'Not Applicable' here. However, the authors have gone through all journal policies and consented to the authorities for further processing.

Consent for publication Since the concerned research paper is based on the 'non-life science journal.' So, 'Not Applicable' here. However, the authors have gone through all journal policies and consented to the authorities for further processing.

References

1. A. Chaudhary, M.J. Kumar, Controlling short-channel effects in deep-submicron SOI MOSFETs for improved reliability: a review. IEEE Trans. Dev. Mater. Reliab. **4**(1), 99–109 (2004). https://doi.org/10.1109/TDMR.2004.824359
2. A. Kumar, N. Gupta, R. Chaujar, TCAD RF performance investigation of transparent gate recessed channel MOSFET. Microelectron. J. **49**, 36–42 (2016). https://doi.org/10.1016/j.mejo.2015.12.007
3. H. Iwai, Roadmap for 22 nm and beyond. Microelectron. Eng. **86**(7–9), 1520–1528 (2009). https://doi.org/10.1016/j.mee.2009.03.129
4. B. Kumar, A. Kumar, R. Chaujar, The Effect of Gate Stack and High-K Spacer on Device Performance of a Junctionless GAA FinFET, in *IEEE VLSI Device, Circuit and System Conference (VLSI-DCS)* (2020), pp. 159–163. https://doi.org/10.1109/VLSIDCS47293.2020.9179855
5. R.M. Barsan, Analysis and modeling of dual-gate MOSFET's. IEEE Trans. Electron Dev. **28**(5), 523–534 (1981). https://doi.org/10.1109/T-ED.1981.20377

6. C.P. Auth, J.D. Plummer, Scaling theory for cylindrical, fully-depleted, surrounding-gate MOSFET's. IEEE Electron Dev. Lett. **18**(2), 74–76 (1997). https://doi.org/10.1109/55.553049
7. A. Sarkar, S. De, A. Dey, C.K. Sarkar, Analog and RF performance investigation of cylindrical surrounding-gate MOSFET with an analytical pseudo-2D model. J. Comput. Electron. **11**(2), 182–195 (2012). https://doi.org/10.1007/s10825-012-0396-9
8. D. Madadi, A.A. Orouji, Investigation of 4H-SiC gate-all-around cylindrical nanowire junctionless MOSFET including negative capacitance and quantum confinements. Eur. Phys. J. Plus **136**, 785 (2021). https://doi.org/10.1140/epjp/s13360-021-01787-0
9. R. Chaujar, R. Kaur, M. Saxena, M. Gupta, R.S. Gupta, TCAD assessment of gate electrode workfunction engineered recessed channel (GEWE-RC) MOSFET and its multi-layered gate architecture, Part II: analog and large signal performance evaluation. Superlattices Microstruct. **46**(4), 645–655 (2009). https://doi.org/10.1016/j.spmi.2009.07.027
10. A. Kumar, M.M. Tripathi, R. Chaujar, Reliability issues of In_2O_5Sn gate electrode recessed channel MOSFET: impact of interface trap charges and temperature. IEEE Trans. Electron Dev. **65**(3), 860–866 (2018). https://doi.org/10.1109/TED.2018.2793853
11. A. Kumar, M.M. Tripathi, R. Chaujar, Comprehensive analysis of sub-20 nm black phosphorus based junctionless-recessed channel MOSFET for analog/RF applications. Superlattices Microstruct. **116**, 171–180 (2018). https://doi.org/10.1016/j.spmi.2018.02.018
12. C. Y. Chang, C. H. Chang, C. H. Hou, K. L. Lin, K. Y. Lee, X. F. Yu, C. O. Chui, Semiconductor devices, Finfet devices and methods of forming the same, US Patent App 15/876,223 (2019).
13. B. Kumar, R. Chaujar, Numerical study of JAM-GS-GAA FinFET: a Fin aspect ratio optimization for upgraded analog and intermodulation distortion performance. SILICON (2021). https://doi.org/10.1007/s12633-021-01395-8
14. The International Technology Roadmap for Semiconductors 2.0: 2015, Itrpv, 2015 [Online]. Available: http://www.itrs2.net/.
15. Y.C. Huang, M.H. Chiang, S.J. Wang, J.G. Fossum, GAAFET versus pragmatic FinFET at the 5 nm Si-based CMOS technology node. IEEE J. Electron Devices Soc. **5**(3), 164–169 (2017). https://doi.org/10.1109/JEDS.2017.2689738
16. L. Zhou, B. Bo, X. Yan, C. Wang, Y. Chi, X. Yang, Brief review of surface passivation on III-V semiconductor. Curr. Comput.-Aided Drug Des. **8**(5), 1–14 (2018). https://doi.org/10.3390/cryst8050226
17. S.H. Chen, W.S. Liao, H.C. Yang, S.J. Wang, Y.G. Liaw, H. Wang, H. Gu, M.C. Wang, High-performance III-V MOSFET with nano-stacked high-k gate dielectric and 3D fin-shaped structure. Nanoscale Res. Lett. **7**, 1–5 (2012). https://doi.org/10.1186/1556-276X-7-431
18. T.A. Bhat, M. Mustafa, M.R. Beigh, Study of short channel effects in n-FinFET structure for Si, GaAs, GaSb and GaN channel materials. J. Nano- Electron. Phys. **7**(3), 1–5 (2015)
19. T.P. Chow, I. Omura, M. Higashiwaki, H. Kawarada, V. Pala, Smart power devices and ICs using GaAs and wide and extreme bandgap semiconductors. IEEE Trans. Electron Dev. **64**(3), 856–873 (2017). https://doi.org/10.1109/TED.2017.2653759
20. P.D. Ye, G.D. Wilk, B. Yang, J. Kwo, S.N.G. Chu, S. Nakahara, H.-J.L. Gossmann, J.P. Mannaerts, M. Hong, K.K. Ng, J. Bude, GaAs metal-oxide-semiconductor field-effect transistor with nanometer-thin dielectric grown by atomic layer deposition. Appl. Phys. Lett. **83**(1), 180–182 (2003). https://doi.org/10.1063/1.1590743
21. B. Yang, P.D. Ye, J. Kwo, M.R. Frei, H.-J.L. Gossmann, J.P. Mannaerts, M. Sergent, M. Hong, K. Ng, J. Bude, Impact of metal/oxide interface on DC and RF performance of depletion-mode GaAs MOSFET employing MBE grown $Ga_2O_3(Gd_2O_3)$ as gate dielectric. J. Cryst. Growth **251**(1–4), 837–842 (2003). https://doi.org/10.1016/S0022-0248(02)02273-X
22. C.P. Chen, Y.J. Lee, Y.C. Chang, Z.K. Yang, M. Honga, J. Kwo, H.Y. Lee, T.S. Lay, Structural and electrical characteristics of Ga_2O_3 (Gd_2O_3) GaAs under high temperature annealing. J. Appl. Phys. **100**(10), 1–5 (2006). https://doi.org/10.1063/1.2386946
23. H.C. Lin, P.D. Ye, G.D. Wilk, Leakage current and breakdown electric-field studies on ultrathin atomic-layer-deposited Al_2O_3 on GaAs. Appl. Phys. Lett. **87**(18), 1–3 (2005). https://doi.org/10.1063/1.2120904
24. T.E. Taouririt, A. Meftah, N. Sengouga, Effect of the interfacial (low-k SiO_2 vs high-k Al_2O_3) dielectrics on the electrical performance of a-ITZO TFT. Appl. Nanosci. **8**(8), 1865–1875 (2018). https://doi.org/10.1007/s13204-018-0866-x
25. N. Gupta, R. Chaujar, Optimization of high-k and gate metal work function for improved analog and intermodulation performance of Gate Stack (GS)-GEWE-SiNW MOSFET. Superlattices Microstruct. **97**, 630–641 (2016). https://doi.org/10.1016/j.spmi.2016.07.021
26. B. Kumar, R. Chaujar, Fin aspect ratio optimization of novel junctionless Gate Stack Gate All Around (GS-GAA) FinFET for Analog/RF applications. Microelectron. Circuits Syst. **755**, 59–67 (2021). https://doi.org/10.1007/978-981-16-1570-2_6

27. J. Robertson, High dielectric constant oxides. Eur. Phys. J. Appl. Phys. **28**, 265–291 (2004). https://doi.org/10.1051/epjap:2004206
28. M.M. Frank, G.D. Wilk, D. Starodub, T. Gustafsson, E. Garfunkel, Y.J. Chabal, J. Grazul, D.A. Muller, HfO_2 and Al_2O_3 gate dielectrics on GaAs grown by atomic layer deposition. Appl. Phys. Lett. **86**(15), 1–3 (2005). https://doi.org/10.1063/1.1899745
29. D.R. Steinke, J. Piccirillo, S.C. Gausepohl, S. Vivekand, M.P. Rodgers, J.U. Lee, Parasitic capacitance removal of sub-100 nm p-MOSFETs using capacitance-voltage measurements. Solid. State. Electron. **68**, 51–55 (2012). https://doi.org/10.1016/j.sse.2011.09.014
30. A. Kumar, M.M. Tripathi, R. Chaujar, Investigation of parasitic capacitances of In_2O_5Sn gate electrode recessed channel MOSFET for ULSI switching applications. Microsyst. Technol. **23**(12), 5867–5874 (2017). https://doi.org/10.1007/s00542-017-3348-2
31. K. Biswas, C. K. Sarkar, Optimizing Fin aspect ratio of junctionless bulk FinFET for application in Analog/RF circuit, in *2018 IEEE Electron Devices Kolkata Conference (EDKCON)* (2018), pp. 591–595. https://doi.org/10.1109/EDKCON.2018.8770515
32. A. Razavieh, P. Zeitzoff, E.J. Nowak, Challenges and limitations of CMOS scaling for FinFET and beyond architectures. IEEE Trans. Nanotechnol. **18**, 999–1004 (2019). https://doi.org/10.1109/TNANO.2019.2942456
33. C. Zhao, J. Xiang, Atomic layer deposition (ALD) of metal gates for CMOS. Appl. Sci. **9**(11), 2039 (2019). https://doi.org/10.3390/app9112388
34. Y. Liu, S. Kijima, E. Sugimata, M. Masahara, K. Endo, T. Matasukawa, K. Ishii, K. Sakamoto, T. Sekigawa, H. Yamauchi, Y. Takanashi, E. Suzuki, Investigation of the TiN gate electrode with tunable work function and its application for FinFET fabrication. IEEE Trans. Nanotechnol. **5**(6), 723–728 (2006). https://doi.org/10.1109/TNANO.2006.885035
35. S.A. Vitale, J. Kedzierski, P. Healey, P.W. Wyatt, C.L. Keast, Work-function-tuned TiN metal gate FDSOI transistors for subthreshold operation. IEEE Trans. Electron Dev. **58**(2), 419–426 (2011). https://doi.org/10.1109/TED.2010.2092779
36. ATLAS User's Manual, SILVACO International, Santa Clara, CA, USA (2016).
37. B. Kumar, R. Chaujar, TCAD temperature analysis of gate stack gate all around (GS-GAA) FinFET for improved RF and wireless performance. SILICON **13**, 3741–3753 (2021). https://doi.org/10.1007/s12633-021-01040-4
38. N. Gupta, A. Kumar, R. Chaujar, Impact of device parameter variation on RF performance of gate electrode workfunction engineered (GEWE)-silicon nanowire (SiNW) MOSFET. J. Comput. Electron. **14**(3), 798–810 (2015). https://doi.org/10.1007/s10825-015-0715-z
39. P.D. Ye, G.D. Wilk, J. Kwo, B. Yang, H.-J.L. Gossmann, M. Frei, S.N.G. Chu, J.P. Mannaerts, M. Sergent, M. Hong, K.K. Ng, J. Bude, GaAs MOSFET with oxide gate dielectric grown by atomic layer deposition. IEEE Electron Dev. Lett. **24**(4), 209–211 (2003). https://doi.org/10.1109/LED.2003.812144
40. H. Lee, L. E. Yu, S. W. Ryu, J. W. Han, K. Jeon, D. Y. Jang, K. H. Kim, J. Lee, J. H. Kim, S. C. Jeon, J. S. Oh, Y. C. Park, W. H. Bae, H. M. Lee, J. M. Yang, J. J. Yoo, S. I. Kim, Y. K. Choi, Sub-5nm all-around gate FinFET for ultimate scaling, in *Digest of Technical Papers-Symposium on VLSI Technology* (2006), pp. 58–59, https://doi.org/10.1109/VLSIT.2006.1705215
41. C. Auth, A. Aliyarukunju, M. Asoro, D. Bergstrom et al., A 10nm high performance and low-power CMOS technology featuring 3rd generation FinFET transistors, Self-Aligned Quad Patterning, contact over active gate and cobalt local interconnects. IEEE Int. Electron Dev. Meet. **2**, 673–676 (2017). https://doi.org/10.1109/IEDM.2017.8268472
42. S.N. Choi, S.E. Moon, S.M. Yoon, Impact of oxide gate electrode for ferroelectric field-effect transistors with metal-ferroelectric-metal-insulator-semiconductor gate stack using undoped HfO_2 thin films prepared by atomic layer deposition. Nanotechnology **32**, 085709 (2021). https://doi.org/10.1088/1361-6528/abc98c
43. N. Gupta, A. Kumar, R. Chaujar, Design considerations and capacitance dependent parametric assessment of gate metal engineered SiNW MOSFET for ULSI switching applications. SILICON **12**(6), 1501–1510 (2020). https://doi.org/10.1007/s12633-019-00246-x
44. K.P. Pradhan, S.K. Mohapatra, P.K. Sahu, D.K. Behera, Impact of high-k gate dielectric on analog and RF performance of nanoscale DG-MOSFET. Microelectronics J. **45**(2), 144–151 (2014). https://doi.org/10.1016/j.mejo.2013.11.016
45. B. Kumar, R. Chaujar, Analog and RF performance evaluation of junctionless accumulation mode (JAM) gate stack gate all around (GS-GAA) FinFET. SILICON **13**, 919–927 (2021). https://doi.org/10.1007/s12633-020-00910-7
46. S.K. Mohapatra, K.P. Pradhan, L. Artola, P.K. Sahu, Estimation of analog/RF figures-of-merit using device design engineering in gate stack double gate MOSFET. Mater. Sci. Semicond. Process. **31**, 455–462 (2015). https://doi.org/10.1016/j.mssp.2014.12.026

47. S. Cho, J.S. Lee, K.R. Kim, B.G. Park, J.S. Harris, I.M. Kang, Analyses on small-signal parameters and radio-frequency modeling of gate-all-around tunneling field-effect transistors. IEEE Trans. Electron Devices **58**(12), 4164–4171 (2011). https://doi.org/10.1109/TED.2011.2167335
48. S. Ghosh, K. Koley, C.K. Sarkar, Impact of the lateral straggle on the analog and RF performance of TFET. Microelectron. Reliab. **55**(2), 326–331 (2015). https://doi.org/10.1016/j.microrel.2014.10.008
49. Q. Wang, S. Wang, H. Liu, W. Li, S. Chen, Analog/RF performance of L- and U-shaped channel tunneling field-effect transistors and their application as digital inverters. Jpn. J. Appl. Phys. **56**, 064102 (2017). https://doi.org/10.7567/JJAP.56.064102
50. N. Vadthiya, S. Tripathi, R.B.S. Naik, A Two-dimensional (2D) analytical modeling and improved short channel performance of graded-channel gate-stack (GCGS) dual-material double-gate (DMDG) MOSFET. SILICON **10**(6), 2399–2407 (2018). https://doi.org/10.1007/s12633-017-9683-1
51. K.E. Kaharudin, F. Salehuddin, A.S.M. Zain, A.F. Roslan, Effect of channel length variation on analog and rf performance of junctionless double gate vertical mosfet. J. Eng. Sci. Technol **14**(4), 2410–2430 (2019)
52. Y.P. Pundir, R. Saha, P.K. Pal, Effect of gate length on performance of 5nm node N-channel nano-sheet transistors for analog circuits. Semicond. Sci. Technol. **36**, 015010 (2020). https://doi.org/10.1088/1361-6641/abc51e
53. V.B. Sreenivasulu, V. Narendar, Junctionless gate-all-around nanowire FET with asymmetric spacer for continued scaling. SILICON (2021). https://doi.org/10.1007/s12633-021-01471-z
54. V.B. Sreenivasulu, V. Narendar, Design and deep insights into sub-10 nm spacer engineered junctionless FinFET for nanoscale applications. ECS J. Solid State Sci. Technol. **10**(1), 013008 (2021). https://doi.org/10.1149/2162-8777/abddd4

Materials Today: Proceedings xxx (xxxx) xxx

Contents lists available at ScienceDirect

Materials Today: Proceedings

journal homepage: www.elsevier.com/locate/matpr

Numerical study of passive mixing in a 3-dimensional helical micromixer with two inlets at offset

Akash Sinha *, M. Zunaid

Department of Mechanical Engineering, Delhi Technological University, Delhi 110042, India

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
Micromixer
CFD
Helical channel
Water
Passive mixing
Simple T-micromixer (STM)
3D dimensional helical micromixer with two inlets at offset (TDHM-TO)

ABSTRACT

The mixing of Newtonian fluid water is numerically investigated in two different micromixers. The present study tackles the problem of time-consumption in mixing at the micro-level by introducing a micromixer named 3D helical micromixer having two inlets at offset (TDHM-TO). It is concluded that this micromixer requires comparatively less mixing channel length and provide much better efficiency than a Simple T-micromixer (STM). A detailed work has been done to examine the mixing process and performance, various streamlines, effective Reynolds numbers for both the micromixers, and pressure drop by solving Navier-Stokes, species transport, and continuity equations for Reynolds numbers (8–400). For all ranges of Reynolds numbers, helical micromixer performed better than STM. TDHM-TO provides maximum efficiency of 97.32% at Re = 310 and minimum efficiency of 79.95% at Re = 8. STM shows comparatively poorer results with a mixing efficiency of just 43.23% at Re = 410. This proposed micromixer can be utilized with different microfluidic systems for various applications which include lab-on-chip and micro-total analysis systems (μ-TAS) for chemical, biochemical and biomedical applications.
Copyright © 2022 Elsevier Ltd. All rights reserved.
Selection and peer-review under responsibility of the scientific committee of the First International Conference on Design and Materials (ICDM)-2021.

1. Introduction

Mixing fluids at the micro-scale level is a difficult task to do. Micro-scaled devices range from 10 to 100 μm and have several applications in biochemical industries, biological industries and other domains. They have several advantages such as less use of reagents, time-efficient, low cost, easily portable etc. The flow is laminar in micromixers due to the low Reynolds number hence it is tough to get the required level of mixing. In some cases, it can even demand a longer span of the micromixer to do the required level of mixing. Molecular diffusion (transfer and movement of molecules through a fluid) mainly pilots the mixing inside the micromixer which is a time-consuming process. According to the mode of operations, micromixers can be divided into two types as active micromixers and passive micromixers. An active micromixer is a micromixer that requires require external energy sources, such as Magnetic force, Ultrasonic, Dielectric-phoretic (phenomenon in which when a non-uniform electric field is exerted on a dielectric particle then it experiences force), etc. to disturb the flow [1,2]. These micromixers are difficult to fabricate and are expensive. Whereas, a passive micromixer is a micromixer that does not require any external energy input. These micromixers only require fluid pumps and the pressure drop to drive the flow which is achieved by making changes in geometrical shapes or by providing obstacles in the path of fluid flow [3,4].

T micromixers have the most basic design of a micromixer and several studies have been done on it describing several flow regimes [5–11]. There have been several studies where modifications in T micromixer were done to enhance the mixing of the fluids. Kurnia et al. [10] proposed and studied a T micromixer with a twisted tape insert (a twisted strip). Their study evaluated several parameters such as Reynolds numbers, twisted tape width, nature of the twisted tape and permeability of the porous tape. It was concluded that using twisting tape can increase the mixing efficiency by two times that of a T micromixer. Dundi et al. [11] evaluated the effects of swirl velocities at the inlets of a T micromixer and compared the result with a T micromixer without having swirl components. The results indicated a noticeable increase in mixing efficiency for all values of Reynolds numbers.

The Interfacial area (area of contact between the fluids) between the fluids also plays a major role in increasing mixing effi-

* Corresponding author.
E-mail address: akashsinha_2k20the03@dtu.ac.in (A. Sinha).

https://doi.org/10.1016/j.matpr.2021.12.437
2214-7853/Copyright © 2022 Elsevier Ltd. All rights reserved.
Selection and peer-review under responsibility of the scientific committee of the First International Conference on Design and Materials (ICDM)-2021.

Please cite this article as: A. Sinha and M. Zunaid, Numerical study of passive mixing in a 3-dimensional helical micromixer with two inlets at offset, Materials Today: Proceedings, https://doi.org/10.1016/j.matpr.2021.12.437

ciency which can be increased by introducing helical and curved channels and shapes. This type of geometrical shape introduces a swirl effect on the fluid flow resulting in an improved diffusion process. Chen et al. [12] studied and investigated an E-shape micromixer. It was concluded that a greater mixing efficiency can be reached by chaotic advection and splitting-recombination due to the design of the mixing channel. Highest achieved mixing efficiency was 94% at Re = 80. Chen et al. [13] in their another study, numerically investigated the performance of serpentine microchannel. A total of 6 micromixers with three different geometries (square-wave, multi-wave and zigzag) were studied with blue and yellow inked fluids. It was concluded that zigzag and square-wave micromixers were more efficient than multi-wave micromixers. Maximum efficiency of 95% was seen when Reynolds number was increased beyond 100.

Placing obstacles in the path of fluid flow is also an efficient way to improve the mixing efficiency. It helps in creating chaotic advection inside the microchannel. Chaotic advection can be defined as the phenomenon where a fluid particle attains complex behaviour and flow traces develop into complex fractals. Since the fluids now have to travel a longer distance, hence, it gets more time to diffuse. Md. Readul Mahmud [14] numerically investigated a modified T micromixer which contained circular and hexagonal shape obstacles with water as the fluid to mix. The study was done on a wide values of Reynolds numbers. The result showed that modified T micromixer with obstacles had much better performance than modified T micromixer without obstacles. Li et al. [15] studied three different micromixers on a wide values of Reynolds numbers (0.01–100). All three micromixers had obstacles placed at different plans A, B, and C named MSMA, MSMB, and MSMC respectively. It was concluded that the mixing efficiency of MSMA and MSMC was much better with mixing efficiency going beyond 90% for all the Reynolds numbers. Similar studies can be found in the references [16,17].

The fabrication of micromixer devices is a challenging issue when common techniques are used. It requires the use of unconventional manufacturing methods. Mondal et al. [18] elaborated a speedy process for the fabrication of micromixers with poly-dimethyl-siloxane (PDMS- silicon based organic polymer). Two types of micromixers were fabricated- one raccoon type and the other serpentine type micromixer. The complete fabrication process took two steps, one is mould preparation and then fabrication. Mould was produced using Co_2 assisted lase machining on poly-methyl-meth-acrylate (PMMA), and then micromixers were fabricated by the soft lithography technique. Similar fabrication techniques can be found in the references [19,20]. The fabrication of micromixers creates several structural problems which need to be looked at as well. There are several techniques that can be used to analyse such structural problems in the micromixer such as EFGM and MLPG which are meshless methods and are used to solve 1D elastostatics problems. Manish et al. [21] analysed various meshless methods and concluded that EFGM and MLPG both can be used to replace FEM and are much better alternatives. Also, EFGM is found to be more accurate than MLPG.

As can be deduced from the brief literature review stated above, it can be said that numerous works exist for micromixers with various types of geometry but very less work for helical micromixers. Although few in-plane spiral micromixers have been studied which are mentioned in reference [22–25], none of them is three-dimensional. In the present study, a numerical study has been done to study and compare the mixing index of both the micromixers i.e.-Simple T micromixer [STM] and a 3D Helical Micromixer with rectangular cross-sections and two inlets which are at an offset to each other [TDHM-TO]. For whole analysis, water is considered as the working fluid. A relative study of mixing efficiency and pressure drop between both the micromixers have been numerically studied using CFD simulations.

2. Micromixer model and specifications

Two micromixers are selected for analysis. The simple T micromixer, and TDHM-TO having two inlets at offset with rectangular cross-sections. The inlet channels of STM and TDHM-TO micromixers have a length and breath as 100 × 100 μm. The outlet has a length and breath as 200 × 100 μm for both the geometries.

The axial length of the straight channel of STM is 3000 μm (L_a). The peripheral length of TDHM-TO (L_m) is equal to the length of the Simple T micromixer (STM) plus length equivalent to one extra spiral which is 1411.71 μm. Fig. 1(c) shows all the parameters required to define TDHM-TO. Pitch (P_h) and mean diameter (d_m) of the helix curve and other dimensions have been mentioned in Table 1. For Reynolds number, hydraulic diameter and velocity of the microchannel are considered.

2.1. Stress concentration in micromixers

The stress concentration because of the curved surface of TDHM-TO cannot be ignored. This stress is generally developed during the fabrication process and several studies have been done on stress concentration due to round or filleted corners. From Fig. 1 (a, b), it can be seen that stress concentration would be high in TDHM-TO compared to STM. This is due to the curved shape of TDHM-TO. Reference [26] can be referred for a detailed study on stress concentration in micromixers.

3. Mathematical modelling, numerical validation and grid independence

Numerical analysis using computer simulation is employed to examine the behaviour of flow and to calculate the mixing index. Incompressible and Newtonian flow is considered in the microchannel, hence continuity and Navier-Stokes equations have been numerically solved as they are the governing equations for the same. These are:

$$(\text{Continuity Equation}) \qquad \nabla.\vec{V} = 0 \tag{1}$$

Where 'V' is the velocity vector of the fluid.

$$(\text{Navier} - \text{Stokes equation}) \qquad V.\nabla V = -\frac{1}{\rho}\nabla P + \nu\nabla^2 V \tag{2}$$

'ρ' is the density of the fluid, 'P' is pressure and 'ν' is kinematics viscosity. ∇ represents the Del operator and is given by

$$\nabla = \frac{\partial}{\partial x}\hat{i} + \frac{\partial}{\partial x}\hat{j} + \frac{\partial}{\partial x}\hat{k} \tag{3}$$

The degree of mixing is obtained by calculating the parameter Mixing index (MI) in the microchannel at any particular location. MI is evaluated by following equation (3),

$$Mi = \left(1 - \sqrt{\frac{\sigma^2}{\sigma^2_{max}}}\right) \times 100 \tag{4}$$

$$\sigma^2 = \frac{1}{n}\sum_{j=1}^{n}\left(c_i - c_{avg}\right)^2 \tag{5}$$

Where σ^2 is the sum of variance in the species concentration of fluid at a particular location, whereas σ^2_{max} denotes maximum variance. 'c_i' is the mass fraction at sampling point 'i', and 'c_{avg}' is the average value of mass fraction at that particular location. The total number of points taken on the plane is 'n' which is equal to 500. If

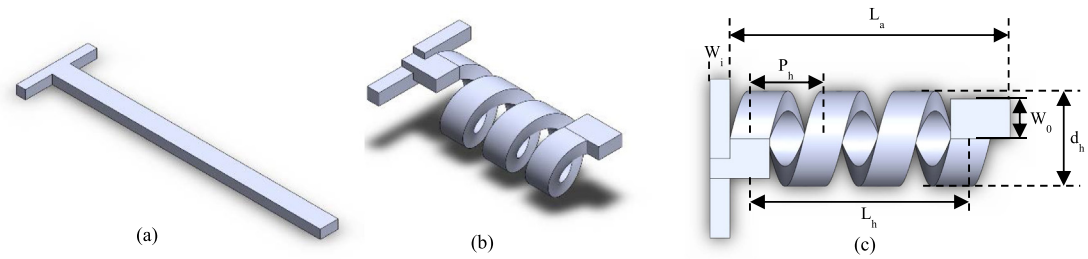


Fig. 1. Model description of (a) Simple T micromixer (STM); (b) 3D Helical Micromixer with two inlets at offset (TDHM-TO); (c) Parameters of the geometry.

Table 1
Various dimensions of micromixer geometries.

Geometric parameter	Values (μm)
Width of the inlet of STM and TDHM-TO (W_i)	100
Width of the outlet of the channel (W_0)	200
Pitch of the curve (P_h)	370.57
The total axial length of the helical curve (L_a)	1411.71
Mean diameter of the curve (d)	380

'MI' is '0', it means that no mixing is taking place whereas, if 'MI' is '1', it denoted that complete mixing is taking place. Residuals for all the equations are set to 10^{-14}. The study was done on a broad range of Reynolds numbers (8–400). To validate the results of the present study, the computation method was firstly compared with an existing work for STM only. The results obtained from STM with water (refer Table 2 for the properties of water) were compared with that of Dundi et al. [11] for Reynolds number 266. The compared result is depicted in Fig. 2. An average error of 3.39% is observed which is a satisfactory error. Hence, the computational method is validated. The grid independence test was done to get an optimum element size. This size is needed to have a better and accurate analysis of the geometries. As shown in Fig. 3, for TDHM-TO, the values of mixing index got saturated after 3.5 million mesh elements. After 3.5 million mesh, there is no difference in the mixing index.

This variation of mixing index is calculated for Reynolds number 230. Hence, 3.5 million mesh elements selected for the analysis. Similar test was done for STM as well prior to the actual simulations.

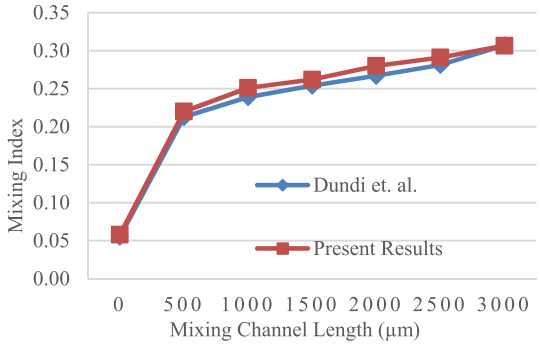


Fig. 2. Validation of current study with the existing literature [11].

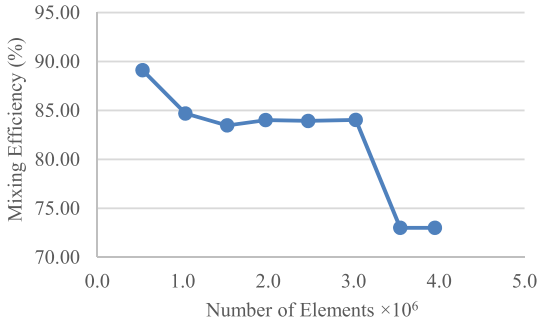


Fig. 3. Grid sensitivity study.

4. Results and discussions

As explained in previous sections, this study deals with the study flow characteristics, mixing index and pressure drop inside STM, and TDHM-TO. These Two micromixers are then compared based on mixing index/efficiency and pressure drop. The study is done on an extensive range of Reynolds numbers (8–400). This section is divided into two parts on the basis of geometry.

4.1. STM analysis with water

Numerical simulation using Ansys Fluent is being performed to evaluate the mixing index and pressure drop of STM for an extensive range of Reynolds numbers (8–400). As described in several previously done studies, there are mainly three types of flow regimes formed in a T-micromixer with water as fluid– straight laminar flow regime, vortex flow regime and the engulfment flow regime. The present study found similar behaviour inside the STM with water as fluid. For low Reynolds number, i.e., for Re $\leq$ 75, the streamlines appeared to be parallel where both the specie were flowing separately as can be visualized in Fig. 4(a) for Re = 8. Here, vortex flow doesn't occur and is hence called a straight laminar flow regime. In this regime, molecular diffusion is solely responsible for mass transfer between the fluid species. With the increase in Reynolds number, for Re = 125, a small weak can be seen in Fig. 4(b). As can be seen from the Figure, the vortices are symmetrical about the vertical plane. This regime is called the vortex flow regime. But as can be seen from Fig. 6, which shows the graph between mixing index and Reynolds number, the achieved mixing obtained from this range of Reynolds number doesn't show significant improvement. This is because, with the increase in Reynolds number, the residence time (the time spent by a fluid particle in the given volume) of fluid species decreases. STM showed this

Table 2
Property of water used in the simulations.

Fluid	Viscosity (Pa s)	Density (Kg/m^3)	Diffusivity (m^2/s)	Temperature (℃)
Water	0.001	998.2	0.09	20

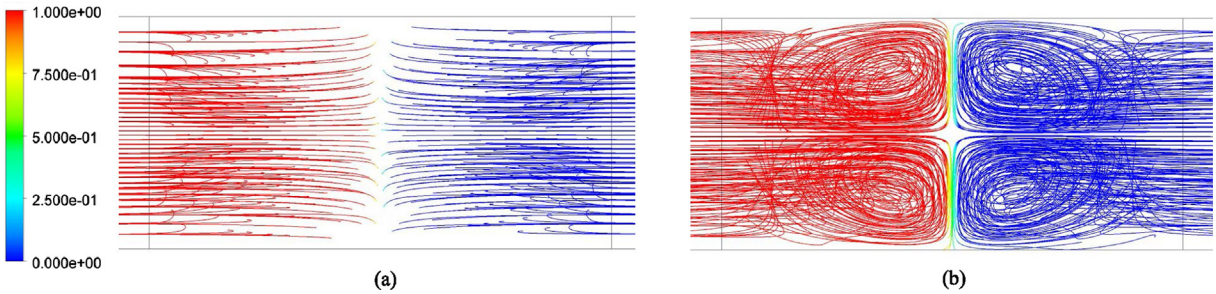


Fig. 4. Species concentration-based streamlines in STM at (a) Re = 8 (b) Re = 125.

same behaviour till Re = 170, after that, the fluid created an engulfment flow regime. In this regime, the symmetry got broken up and strong mixing was observed due to chaotic advection and now the streams are no longer parallel. The mixing index has been calculated at various cross-sections in the mixing channel using equations (4) and (5). The mixing index over the range of Reynolds numbers has been depicted in Fig. 6.

4.2. TDHM-TO analysis and comparison with STM with water

Similarly, numerical simulation has been performed to evaluate the mixing index and pressure drop of TDHM-TO for a broad range of Reynolds numbers (8–400). Fig. 6 presents a comparison of the mixing index of both the micromixers- STM and TDHM-TO with Reynolds numbers. Here, the mixing indexes have been calculated at the outlet of the mixing channels. As previously described in section 4.1, the mixing efficiency is very poor in STM for low Reynolds number. This is due to the flow being laminar and straight and hence results in a low mass diffusion phenomenon. As can be seen from Fig. 6, the mixing index is further reduced to a position known as 'critical Reynolds number'. After this point, vortex increased with the increase in Reynolds number and this zone is called the engulfment zone. In this zone, the mixing of the fluid species increased due to the chaotic advection and after critical Reynolds number, the strength of secondary flow increased.

As can be seen from Fig. 6, TDHM-TO provides better mixing efficiency for all values of Reynolds numbers than STM. TDHM-TO prove to be a better option than STM. For low Reynolds number (Re $\leq$ 75), TDHM-TO provides more interfacial area between the fluid layers and hence stretches the fluid which increases the mass diffusion and hence the mixing index. In this range of low Reynolds

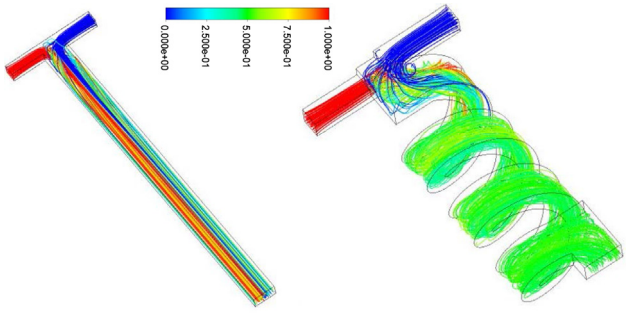


Fig. 5. Mixing of water for (a) STM; (b) TDHM-TO at Re = 230.

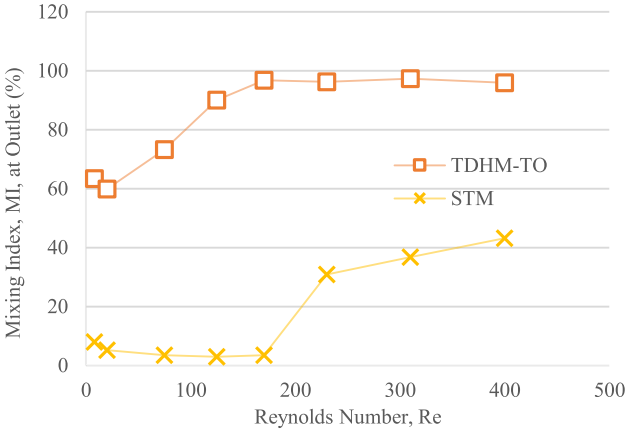


Fig. 6. Variation of mixing index (MI) with Reynolds number.

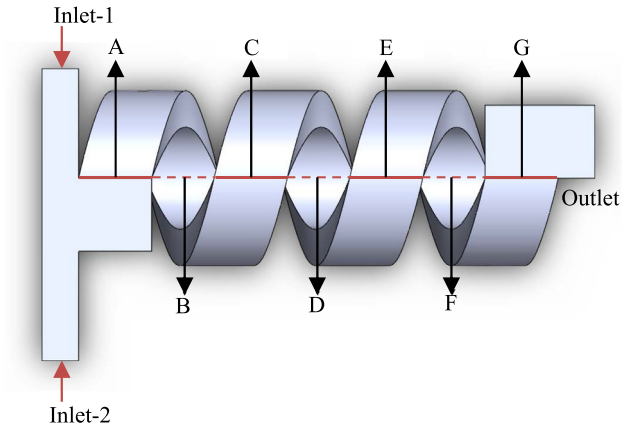


Fig. 7. Location of planes at various section (Top View).

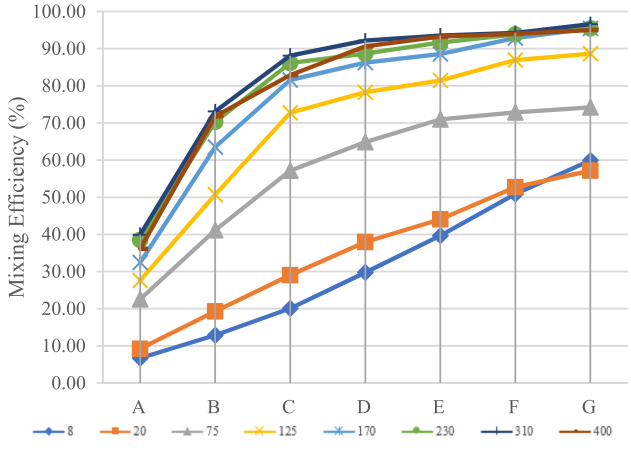


Fig. 8. Variation of Mixing index at seven different locations in TDHM-TO.

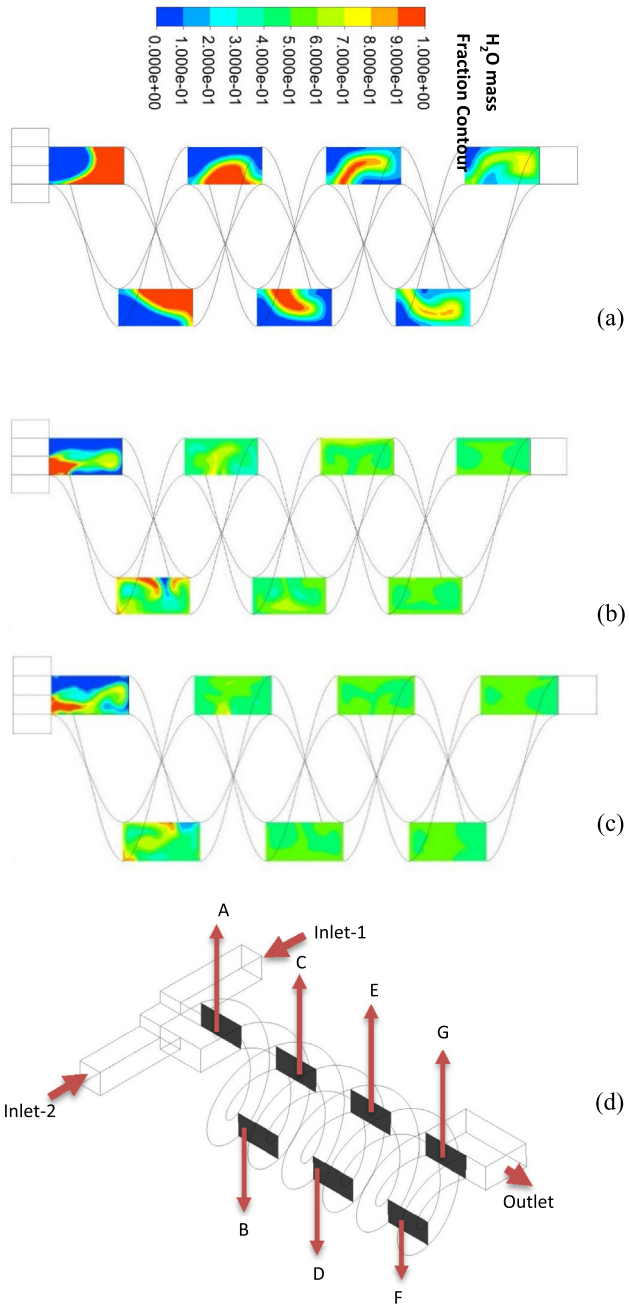


Fig. 9. Species concentration at different sections A, B, C, D, E, F & G in TDHM-TO for Re = (a) 8; (b) 170; (c) 310; (d) Location of planes at various sections (isometric view).

number (Re $\leq$ 75), the highest mixing index was found to be 73.28% for Re = 75. This value is larger than STM by 69.83%. For a higher Reynolds number (Re = 170), the mixing index was found to be 96.76% which is higher than STM by 93.27%. This is because TDHM-TO provides vortex at the inlet due to its design and also because secondary flow is dominating now that has generated due to the generation of centrifugal forces because of the three-dimensional curvature of TDHM-TO. Secondary flow usually develops around curved three dimensional surfaces. Secondary flows are perpendicular to the main direction of fluid flow. This introduces chaotic advection which increased the mixing index. Fig. 5 shows the streamlines for STM and TDHM-TO at Re = 230 which shows the above-mentioned phenomenon.

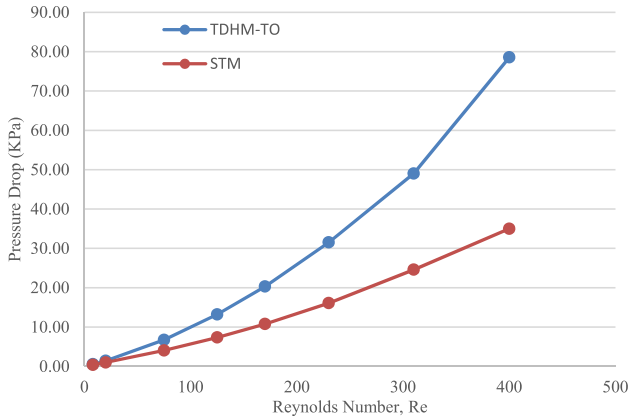


Fig. 10. Pressure drop variation in STM and TDHM-TO with Reynolds number.

An in-depth study is being done to calculate the mixing index at various locations inside the micromixer. These planes have been taken along the mixing channel length at specified distances which can be seen in the Fig. 7 and Fig. 8(d) which show the location of all seven planes. Fig. 8 shows the mixing index of TDHM-TO at those planes for different values of Reynolds number. For Re = 8 and 20, the curve of the mixing index seems to be linear which, however, becomes non-linear as the Reynolds number increases. This is due to the rapid mixing of fluids. At Re = 230, the mixing efficiency increased from 38.49% at first plane A to 91.69% at plane E which is a sharp increase. This shows that a short length TDHM-TO can also be used for some specific problems. Fig. 9 represents species concentration at various planes A, B, C, D, E, F, & G for Re = 8, 170 & 310. The mixing area appears to be small, linear and slim on initial plane A and plane B, but when the flow goes through the mixing channel, this interface between the fluid surface grows and mixing increases. This is due to the inertia of fluids which generates rotation in the fluid leading to the stretching and vortex is formed which enhances the mixing of the fluids.

Pressure drop is a necessary parameter to calculate. This is because enhanced mixing needs increased pumping power to drive the flow of fluid species inside microchannels. Pressure drop provides the idea of required input energy and especially in passive micromixers. Fig. 10. represents a comparison of pressure drop between STM and TDHM-TO micromixers with increasing Reynolds number. As can be seen from the figure, no noticeable difference was found. for the low Reynolds number below, i.e., for Re = 8, but after that TDHM-TO micromixer deviated from that of STM. For all considered Reynolds numbers, TDHM-TO showed the highest pressure drop between the two.

5. Conclusions

In this paper mixing efficiency of two micromixers named Simple T-micromixer (STM), and TDHM-TO which has both inlets at offset are being studied for Newtonian fluid water. A CFD simulation tool, ANSYS 20.1, has been used to perform the simulations. Following are the conclusions that can be made based on the simulated results. TDHM-TO provided better efficiency for all values Reynolds number when compared with STM. This is due to the presence of a large interfacial area and the vortex of flow that is being created at the junction of two inlets which are placed at an offset to each other. A Significant improvement of 93.27% was observed at Re = 170. This is also due to the presence of chaotic advection in the case of a high Reynolds number (Re $\geq$ 125). For higher Reynolds number (Re = 400), the value of the mixing index got increased from 36.12% at plane A to 90.63% at plane D which

happened in a comparatively smaller length. Hence, it can be concluded that a small length micromixer can also be used for certain applications in this particular range of Reynolds numbers. These three-dimensionally curved micromixers require less mixing length (1411.71 μm) when compared with STM (3000 μm) and still produces higher efficiency. Such highly efficient devices are very useful when high accuracy is needed such as in development of vaccines and antibodies, in chemical engineering field to study chemical reactions. They use less reagents, are time-efficient, cheap and easily portable. Although the pressure drop is comparatively higher in the case of TDHM- TO when compared with STM.

CRediT authorship contribution statement

Akash Sinha: Data curation, Writing- Original draft preparation, Visualization, Investigation, Software, Editing. **M. Zunaid:** Conceptualization, Methodology, Software, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We gratefully thank Delhi Technological University for their support.

References

[1] D. Nouri, A. Zabihi-Hesari, M. Passandideh-Fard, Rapid mixing in micromixers using magnetic field, Sens. Actuators, A 255 (2017) 79–86.
[2] M. Annabestani, S. Azizmohseni, P. Esmaeili-Dokht, N. Bagheri, A. Aghassizadeh, M. Fardmanesh, Multiphysics analysis and practical implementation of a soft μ-actuator- based microfluidic micromixer, J. Microelectromech. Syst. 29 (2) (2020) 268–276.
[3] W. Raza, S. Hossain, K.-Y. Kim, A review of passive micromixers with a comparative analysis, Micromachines 11 (5) (2020) 455.
[4] M.B. Okuducu, M.M. Aral, Performance analysis and numerical evaluation of mixing in 3-D T-shape passive micromixers, Micromachines 9 (5) (2018) 210.
[5] A. Mariotti, C. Galletti, M.V. Salvetti, E. Brunazzi, Unsteady flow regimes in a T-shaped micromixer: mixing and characteristic frequencies, Ind. Eng. Chem. Res 58 (29) (2019) 13340–13356.
[6] C.P. Fonte, D.F. Fletcher, P. Guichardon, J. Aubin, Simulation of micromixing in a T-mixer under laminar flow conditions, Chem. Eng. Sci. 222 (2020) 1–10.
[7] X. Luo, Y. Cheng, W. Zhang, K.e. Li, P. Wang, W. Zhao, Mixing performance analysis of the novel passive micromixer designed by applying fuzzy grey relational analysis, Int. J. Heat Mass Transfer 178 (2021) 121638, https://doi.org/10.1016/j.ijheatmasstransfer.2021.121638.
[8] M. Zunaid, S. Tokas, Computational analysis of passive mixing in T-Micromixer with Non- Newtonian, Blood 10 (3) (2020) 9889–9898.
[9] R. Prakash, M. Zunaid, Samsher, Simulation analysis of mixing quality in T-junction micromixer with bend mixing channel, Mater. Today:. Proc. 47 (13) (2021) 3833–3838.
[10] J. Kurnia, A.P. Sasmito, Performance evaluation of liquid mixing in T-junction passive micromixer with twisted tape insert, Ind. Eng. Chem. Res. 59 (9) (2019) 3904–3915.
[11] T.M. Dundi, V.R.K. Raju, V.P. Chandramohan, Numerical evaluation of swirl effect on liquid mixing in a passive T-micromixer, Aust. J. Mech. Eng. 00 (00) (2019) 1–15.
[12] X. Chen, J. Shen, Numerical and experimental investigation on splitting-and-recombination micromixer with E-shape mixing units, Microsyst. Technol. 23 (10) (2017) 4671–4677.
[13] X. Chen, T. Li, H. Zeng, Z. Hu, B. Fu, Numerical and experimental investigation on micromixers with serpentine microchannels, Int. J. Heat Mass Transf. 98 (2016) 131–140.
[14] M.d. Mahmud, Readul, numerical investigation of liquid-liquid mixing in modified T Mixer with 3D Obstacles, J. Eng. Adv. 2 (02) (2021) 87–94.
[15] T. Li, X. Chen, Numerical investigation of 3D novel chaotic micromixers with obstacles, Int. J. Heat Mass Transf. 115 (2017) 278–282.
[16] X. Shi, L. Wang, S. Huang, F. Li, A novel passive micromixer with array of Koch fractal obstacles in microchannel, J. Dispersion Sci. Technol. 42 (2) (2021) 236–247.
[17] Z. Wu, X. Chen, A novel design for 3D passive micromixer based on Cantor fractal structure, Microsyst. Technol. 25 (1) (2019) 225–236.
[18] B. Mondal, S. Pati, P.K. Patowari, Fabrication of wavy micromixer using soft lithography technique, Mater. Today:. Proc. 26 (2) (2020) 1271–1278.
[19] C. Chung, Y.-J. Chen, P.-C. Chen, C.-Y. Chen, Fabrication of PDMS passive micromixer by lost-wax casting, Int. J. Precis. Eng. Manuf 16 (9) (2015) 2033–2039.
[20] R.A. Taheri, V. Goodarzi, A. Allahverdi, Mixing performance of a cost-effective split-and- recombine 3D micromixer fabricated by Xurographic method, Micromachines 10 (11) (2021) 786.
[21] M. Kumar, J. Rajiyan, P. Gupta, A computational approach for solving elasto-statics problems, Mater. Today:. Proc. 46 (15) (2021) 6876–6879.
[22] P. Vatankhah, A. Shamloo, Parametric study on mixing process in an in-plane spiral micromixer utilizing chaotic advection, Anal. Chim. Acta 1022 (2018) 96–105.
[23] E. Tripathi, P.K. Patowari, S. Pati, Numerical investigation of mixing performance in spiral micromixers based on Dean flows and chaotic advection, Chem. Eng. Processing - Process Intensification 169 (2021) 108609, https://doi.org/10.1016/j.cep.2021.108609.
[24] S. Tokas, M. Zunaid, M.A. Ansari, Numerical investigation of the performance of 3D-helical passive micromixer with Newtonian fluid and non-Newtonian fluid blood, Asia-Pac. J. Chem. Eng. 16 (1) (2021), https://doi.org/10.1002/apj.v16.110.1002/apj.2570.
[25] S. Tokas, M. Zunaid, M. Ansari, Non-Newtonian fluid mixing in a Three-Dimensional spiral passive micromixer, Mater. Today:. Proc. 47 (13) (2021) 3947–3952.
[26] R. Hatwar, P. S. Gandhi, M. D. Atre, C. Patel, V. Giramkar, S. Joseph, G. J. Phatak, Analysis, fabrication and testing of ultra high pressure microchannels using LTCC technology, 1st International Symposium on Physics and Technology of Sensors (ISPTS-1), 2012, pp. 224-227.

Journal of the Brazilian Society of Mechanical Sciences and Engineering (2022) 44:56
https://doi.org/10.1007/s40430-022-03362-9

TECHNICAL PAPER

Obstacle avoidance for a swarm of unmanned aerial vehicles operating on particle swarm optimization: a swarm intelligence approach for search and rescue missions

Girish Kumar[1] · Arham Anwar[1] · Abhinav Dikshit[1] · Abhirup Poddar[1] · Umang Soni[2] · Weon Keun Song[3]

Received: 18 April 2021 / Accepted: 3 January 2022
© The Author(s), under exclusive licence to The Brazilian Society of Mechanical Sciences and Engineering 2022

Abstract
An approach, based on a multi-plane system, is conceptualized in this work to solve the problem of collision avoidance for a swarm of unmanned aerial vehicles, being used for search and rescue to minimize affecting the searching algorithm. Relevant chronological advancements in the last two decades of the parent algorithm, particle swarm optimization, are summarized. As each optimization algorithm for search and rescue has its own niche area of application, various well-established algorithms such as particle swarm optimization and novel algorithms like layered search and rescue, spiral search and fish-inspired task allocation are compared with each other qualitatively. Simulations with 100 different cases were used to compare the original particle swarm optimization with the additional novel collision avoidance algorithm. The statistical z test was run based on which it was found that the proposed algorithm significantly reduces the number of collisions and does not put a toll on the iterations to convergence. Standardized residuals of all cases indicate minimal error difference in the optimum average fitness value calculated by the particle swarm optimization, with and without the conceptualized anti-collision algorithm.

Keywords Swarm robotics · Search and rescue missions · Unmanned aerial vehicles · Particle swarm optimization · Obstacle avoidance

1 Introduction

A search and rescue mission poses critical challenges as every passing second is important when life is at stake. The survival of affected people is dependent on the success of search and rescue missions. Such missions are often employed when unforeseen events such as natural calamities and other disasters occur. According to the Centre of Research on the Epidemiology of Disasters and United Nations Office for Disaster Risk Reduction, 2021, in the span of the last two decades from 2000 to 2019, a total of 7348 natural hazard-related disasters were reported, leading to 1.23 million deaths, 4.03 billion people affected and a loss of 2.97 trillion US dollars [1]. Search and rescue missions take place to save as many lives as possible, wherein the key objective is to find as many targets as possible in the least amount of time. In many challenging cases, on-ground search becomes infeasible and an aerial approach is opted. Nuclear decay zones, regions prone to natural calamities and military hot zones can particularly be infeasible to penetrate via direct human intervention. The use of unmanned aerial vehicles (UAVs) in these cases serves as a plausible solution.

Technical Editor: Rogério Sales Gonçalves.

✉ Girish Kumar
girish.kumar154@gmail.com

Arham Anwar
arhamanwar_2k17me59@dtu.ac.in

Abhinav Dikshit
abhinavdikshit_2k17me08@dtu.ac.in

Abhirup Poddar
abhiruppoddar_2k17me09@dtu.ac.in

Umang Soni
umangsoni.1@gmail.com

Weon Keun Song
bauman98@naver.com

[1] Delhi Technological University, Delhi, India

[2] Netaji Subhas University of Technology, Delhi, India

[3] School of Engineering, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand

Published online: 17 January 2022

Search and rescue (SAR) missions have seen a burgeoned potential of UAVs, wherein the limitations of human search groups are bridged by autonomous search and rescue robot units. However, the scope of single UAV search operations is severely limited to the drone's layered search and rescue (LSAR) algorithm follows the key limitations. A single UAV search and rescue (UAV–SAR) operation has a limited range of coverage for a fixed amount of time. In other words, the scope of single UAV searching is restricted to the path it follows. However, in a group of UAVs employed for search and rescue missions the scope is extended to the combined paths of all UAVs. Moreover, sending a single drone to a hot spot will drastically reduce the success rate of the operation. These limitations are overcome by using multiple UAVs instead of one, operating coherently as a system to perform search and rescue missions.

Many researchers, recently, have proposed SAR algorithms where a swarm of unmanned aerial vehicles is used to optimize the performance of search and rescue operations. Layered search and rescue (LSAR) algorithm for a swarm of SAR-UAVs, for the case where the centre of disaster is accurately known, was proposed by Alotaibi et al. [2]. LSAR algorithm follows the key idea that in natural disasters there is a centre where most of the survivors are located and the algorithm focuses more attention on the known centre and gradually less attention with increasing distance from the known centre of the disaster. A spiral search algorithm for the case of immediate search after big natural disasters, in which UAVs would spiral around each identified point, was suggested by Arnold et al. [3]. Each algorithm has its niche area wherein it works the best.

One popular root algorithm is particle swarm optimization (PSO), which has been modified numerous times to improve its efficiency for SAR missions using a swarm of UAVs, wherein each UAV acts as a PSO swarm particle to search for targets. PSO is popular because of its advantages, which are listed by Lee et al. [4], as relatively low computational complexity, easy to code, less sensitivity to the nature of the objective function. PSO also has a limited number of parameters including only the inertia weight factor and two acceleration coefficients in comparison with other competing heuristic optimization methods, and the impact of parameters on the solutions is considered to be less sensitive compared to other heuristic algorithms. Many PSO implementations, however, have either taken the assumption of no collisions and obstacles or have integrated an obstacle avoidance algorithm which significantly affects the searching algorithm's function by changing the trajectories of swarm particles in directions in which fitness value is modified. In this document, an obstacle and collision avoidance algorithm is proposed which is integrated with particle swarm optimization without theoretically affecting the original algorithm's functioning.

The paper presents an exhaustive literature review in Sect. 2 which summarizes the classification of UAV-SAR cases and lists a few solutions for each case followed by a summary of relevant PSO advancements and modifications as a UAV-SAR algorithm and the research gaps. Section 3 explains the methodology wherein the collision and obstacle avoidance algorithms are explained in detail. Section 4 explains the simulation of the integration of collision and obstacle avoidance algorithm with PSO. Section 5 discusses the results and implications of the model followed by Sect. 6 as a conclusion which contains limitations of the work, future directions and application extensions.

2 Literature review

The literature review is divided into two sections. Initially, Sect. 2.1 explicates the scope of search and rescue problems using unmanned aerial vehicles, provides a summary of relevant search and rescue unmanned aerial vehicle solutions and explains the classification of search and rescue problems. Section 2.2 enlists related works, explains the capacity of particle swarm optimization as an algorithm for search and rescue missions using UAVs, and highlights the necessary adjustment that needs to be addressed for working of PSO in a physical search space using unmanned aerial vehicles as particles of PSO.

2.1 Search and rescue problems

The domain of search and rescue problem (SAR) envelopes the activities of searching for people in distress or imminent danger and the provision of aid to them. Based on the search area location, environment and size, search and rescue missions are classified as: maritime search and rescue (MSAR), combat search and rescue (CSAR), urban search and rescue (USAR) and wilderness search and rescue (WiSAR) [5]. Figure 1 shows the general search and rescue operation workflow. This workflow addresses the factors corresponding to the classification, which are search area location, environment and size, and further decides which category of SAR the case belongs to. If the target position is known, rescue plans are made. If the targets are not known, search execution takes place via the most suitable searching algorithm for the case.

2.1.1 Maritime search and rescue

Maritime search and rescue (MSAR) involve cases where objectives are lost in natural water bodies such as the sea and ocean. When it comes to tackling maritime search and rescue missions, three major constraints limit the usage of unmanned aerial vehicles [6]. First of all, the flight time

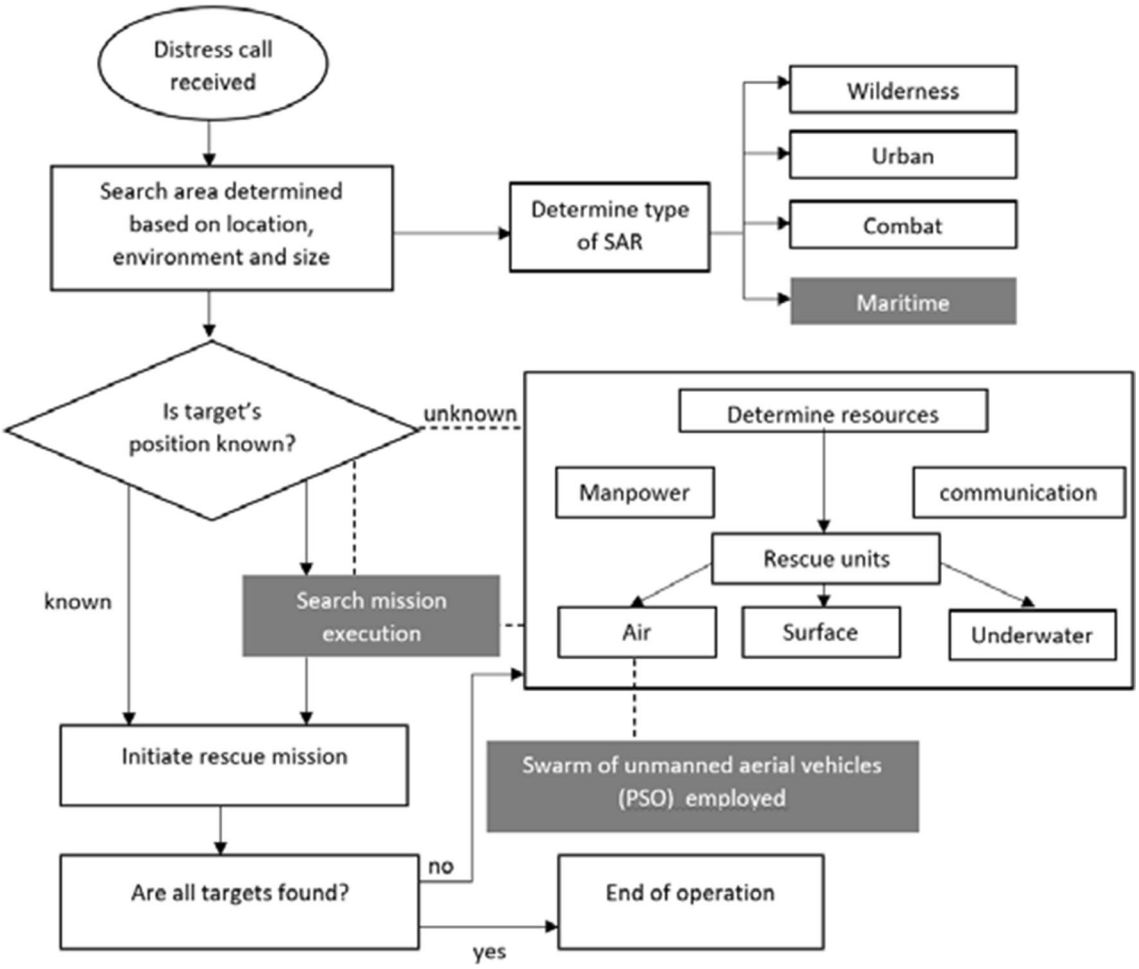


Fig. 1 General operation workflow of search and rescue missions

of commercial unmanned aerial vehicles is limited due to battery or capacity. Moreover, the dynamicity and uncertainty of survivor positions and the need for automated control of fuel service stations make the case furthermore challenging. These challenges were addressed by Lee et al. [6], using the mixed-integer linear programming (MILP) efficiently. Using the cumulative knowledge of last seen locations, cooperative game theory to enable cognitive multi-unmanned aerial vehicles, in an attempt to achieve optimization based on unmanned aerial vehicles with on-board computers, was applied by Rahmes et al. [7]. While the model seems robust for cases where last seen locations are accurately known, the model fails in general cases where this information is not known. A combination of Gaussian mixture models and Fourier transformations for detection of survivors was employed by Dinnbier et al. [8]. However, this algorithm heavily suffers around coastal regions since the algorithm only works for the case where there is no land in the field of vision of UAVs.

2.1.2 Urban search and rescue

Human teams start looking for survivors manually in times of disasters such as an earthquake or terrorist attack. Robots are integrated with human teams since they can efficiently squeeze into tight spaces and reach dangerous areas. Urban search and rescue (USAR) envelopes location, extrication and initial medical stabilization of victims trapped in an urban area. These missions can further be classified as accidental and deliberate, based on the cause. If the scenario is caused by say a terror group, the case is referred to as deliberate USAR which demands a different set of skills from UAVs as compared to simple to an accidental USAR. All other cases correspond to accidental USAR. Technical aspects of autonomous USAR robots were presented by Kohlbrecher et al. [9]. The solution focused on two aspects which were: survivor search and detection, and navigation towards the goal. The first aspect can be solved if the system is as close to the survivors as possible even

if the sensors cover only a small portion of the survivor. This proved to be a deterrent to most survivor locations. For navigation towards the goal, the robot observes the survivor it is directly facing. Therefore, the robot has to turn around in each direction for detection. Cui et al. [10] proposed a solution to mitigate the dynamicity which involves real-time image stitching, indoor navigation, vision-based and poses estimation with multiple cooperative MAVs. The implementation methodology had two phases—first, MAVs explore and search the area for the algorithm to aerially photograph and stitch a map together, and second, they search and identify each individual point of interest. The proposed solution, however, is not fully autonomous because a human operator is required to manually stitch the map together for an obstacle-free route for the MAVs. Moreover, in the next phase, an operator is required to identify the victim's house based on the feedback received from the MAV's camera.

2.1.3 Wilderness search and rescue

The search process in which people are lost or are in distress in the wilderness is known as wilderness SAR (WiSAR). Subjected to potentially toxic environments, medical UAVs can be employed in the operations concerning locating targets, emergency medical care and swift transfer of patients. Grissom et al. [11] show how unmanned aerial vehicles also provide information about the area of operation which can be used for future missions. A camera-based position detection system for SAR operations was developed and integrated into UAVs by Sun et al. [12]. The system identified real-time targets, post-targets and also photographed the disaster areas for subsequent operations. Al-Kaff et al. [13] proposed a model WiSAR where operations were handled by mainly protecting human life in risky and unsafe environments using UAVs. The authors presented a human body detection and tracking algorithm. This was accomplished with an on-board sensor on the UAV that captured colour and depth data followed by validation of their work in both real and simulated environments along with an ability to detect multiple survivors.

Nattero et al. [14] compared the real-time performance of multi-chopper systems, developed an autonomous multi-drone SAR strategy and categorized the algorithms based on energy consumed and time-to-completion. Arnold et al. [3] proposed a spiral search method for search and rescue via UAVs after a big natural disaster, which has very high area coverage rates with respect to time.

2.1.4 Combat search and rescue

SAR operations carried out during war are called combat SAR. An auction-based approach, developed by Wie et al. [15], was used to solve SAR task priority problems where tasks were assigned to each robot based on techniques for deciding winner tasks. When a robot was not in a condition to operate, the task was assigned to idle robots. The completion time and required steps were measured along with the energy consumed. The disadvantage of this algorithm is that it is assumed that the robots have knowledge of the previous tasks, environment and the initial location of the survivors. However, this contrasts with one of the most frequent conditions in search and rescue operations, that is survivor locations are neither known nor static. Table 1 contains a brief list of solutions for unmanned aerial vehicle search and rescue missions, classified according to the type of search and rescue, that is, MSAR, USAR, WiSAR and CSAR.

2.2 Particle swarm optimization for unmanned aerial vehicles

A basic variant of the particle swarm optimization (PSO) algorithm works by having a swarm of candidate solutions or particles. Many alterations have been made in the past for a plethora of use cases and environments. Many new metaheuristic algorithms with unique features have been proposed in the recent past, and the proposed multi-plane approach of collision and obstacle avoidance is the extension of these. Some examples of new metaheuristic algorithms include the African vulture optimization [16], which features a unique exploration and exploitation phase split for searching, and the African gorilla optimization [17], which specializes in global optima problems of higher dimensions. Fish-inspired algorithm for multiple-UAV task allocation, or FIAM [18], features low computational task allocation aiming to minimize search time. However, PSO is a robust and widely tested metaheuristic algorithm is used as the base algorithm in this paper. Some of the relevant UAV adaptations of PSO modifications are enlisted in Table 2, while Table 3 enlists some other popular algorithms developed for the particular use case of search and rescue.

2.3 Advantages of the algorithm proposed

There are several studies done in the past on modifications and limitations of the PSO algorithm. However, in light of search and rescue missions with UAVs as PSO particles, collisions avoidance is absent in many proposals. dPSO-U for exploring disaster scenario area with UAVs as PSO particles was simulated and compared to the lawn mower algorithm, which sweeps the entire scenario area, by Garcia et al [19]. Lawn mower algorithm was found to be slower than dPSO-U. These particles can be considered immune to collision if it is assumed that each particle is an abstract entity that can pass through another without hindering the motion of either particle. However, UAVs are neither infinitesimally small nor are abstract entities; thus, in systems where multiple drones

Table 1 Summary of UAV solutions for each classification of SAR

Reference	Type	Algorithm used	Remarks	Advantages
[6]	MSAR	Mixed-integer linear programming	developed a method to efficiently plan the tasks for an unmanned aircraft system (UAS)	Features fuel station solution for battery capacity
[8]	MSAR	Gaussian mixture models and Fourier transforms	Efficient detection of survivors with exceptionally low false positives, albeit reliant on strict assumptions	A robust survivor detection model with very low false positive for the given assumptions
[10]	USAR	Real-time image stitching	Accuracy relies on image stitching and manual detection of survivors	A two-phase searching algorithm that stitches images of MAVs to create a map for an obstacle-free search space—mitigating dynamicity of targets
[3]	USAR	Spiral method	The modular algorithm has potential to be fused with other algorithms to optimize searching; Ideal for congregated collective behavioural positions	Can be used a modular algorithm to fast track convergence
[12]	WiSAR	All-in-one camera-based target detection and positioning algorithm	Camera-based target identification algorithm proposed relies on the assumption of linear scales of the camera and world coordinates	UAVs realize searching and geo-information acquisition in a single flight—features accurate position detection
[13]	WiSAR	Multi-object tracking algorithm and semi-autonomous reactive control	Specially designed to cope with aerial sequences, where the appearance, position and orientation of the objects are continuously changing, and the use of the traditional trajectory-based tracking is not viable	Provides information about the survivors in unfriendly environments and assists the pilot to perform smooth and safe approaching manoeuvres to the detected points
[15]	CSAR	Auction-based novel algorithm	There is a trade-off between fuel consumption and minimum search time for different application domains when using the algorithm	An auction-based approach and a novel prediction approach to dynamic task allocation
[2]	Known centre of disaster	Layered search and rescue	Efficiency of the algorithm relies on whether the centre of disaster is known accurately	A novel task distribution technique for SAR scenarios involving multiple autonomous UAVs

Table 2 Summary of chronological and relevant PSO modifications

Reference	Algorithm	Key results	Chronological advancement	Swarm size	Remarks
Hereford et al. [27]	dPSO	Improved average search time with sensor and position corrections	PSO with basic obstacles and one target	3	Finding the target (the brightest point of light in the search space) in the presence of obstacles
Peng et al. [30]	PSO + dynamic weights	15.5% less Voyage as compared to pure PSO	Dynamic weights for PSO algorithm parameters	20	Model predictive control with partially known environment information
Geng et al. [31]	PSO + contraction factor	Fuel cost lower from 49 to 40 units in the modified algorithm	Developed contraction factor—increased convergence rate and convergence speed	30	Used contraction factor to mitigate convergence into local minima
Loscri et al. [32]	PSO-S	For a varying coverage threshold, the algorithms show a threshold effect around 0.7	Dynamic Search Space	30–80	Integration wtih virtual forces algorithm
Ghamry et al. [21]	PSO-CTPD	Zero collisions for forest fire trajectory planning	Minimized travelling distance with obstacle avoidance	6	Integration with control parameterization and control descretization
Zhang et al. [23]	Per-drone iterated PSO	Coverage by PSO with 800 UAVs was achieved with 200 UAVs using the proposed algorithm	Analysis of drone–cell user coverage and feasible working zone by leveraging drone-to-user and drone-to-base pathloss models	10	maximizing the user coverage while maintaining D2B link qualities, for a given number of drone cells being deployed
Garcia et al. [19]	dPSO	Discovers 25%, 50% and 75% of targets faster than lawn mower algorithm	Tailored for disaster scenario area, UAVs follow a DTN networking approach	6	Finding victim clusters post-disasters in minimum time
Abdollahzadeh et al. [16]	PSO-GA, PSO-HSA	7% less distance traversed compared to PSO	Hybridization with GA and HAS, respectively	100	Reduced traversal time and fuel consumed

Table 3 Some other popular algorithms used for SAR via a swarm of UAVs

Reference	Algorithm	Swarm size	Key deliverables	Robot employed	# targets	Objective
[33]	Genetic algorithm	3–6	Proposed Simultaneous Inform and Connect strategy which offers shorter mission time compared to inform-first and connect-first strategies	UAV	–	To minimize the mission completion time, which includes area coverage and network connectivity
[3]	Spiral	1–20	With the development of spiral search algorithm, an area can be searched at very fast rates which increase rapidly with an increase in deployed UAVs	UAV	–	To gather maximum situational awareness data during the first few hours after a major natural disaster UAVs With spiral search, 98.8%
[2]	LSAR	2–128	Developed LSAR algorithm and compared the same with other algorithms in terms of rescued survivors and rescue and execution times with max-sum, auction-based and locust-inspired approaches for multi-UAV task allocation	UAV	2–4096	Minimizing the SAR total time while saving the maximum number of people
[18]	FIAM	4–128	Novel low computational algorithm to address the challenges of task allocation problem	UAV	2–4096	Minimizing mean rescue time

operate in close proximity, chances of collisions are high. This brings forth the need for a collision avoidance system that is incorporated in the original PSO algorithm. The proposed algorithm features this advancement of integration of collision avoidance without compromising on the searching efficiency of the original PSO. This modification allows for smooth and faster convergence of UAV particles with significantly reduced collisions.

Many strategies that enforce a minimum distance between UAVs add additional constraints (of minimum distance) to the optimization problem that complicate the original problem [20]. A strategy that incorporates obstacle and collision avoidance without changing the trajectory of the particles and the optimality conditions would provide to be more fruitful [21]. A real-time path planning strategy for UAV based on improved particle swarm optimization was proposed by Cheng et al. [22].

Based on these research gaps, in this paper, integration of a novel obstacle and collision avoidance mechanism into the searching algorithm is proposed in this paper, which theoretically does not affect the searching algorithm significantly and makes the employment of UAVs for search and rescue missions a step closer to practical use.

3 Methodology

Section 3.1 provides a brief overview of how the PSO algorithm for searching via unmanned aerial vehicles. Section 3.2 explains the novel collision avoidance algorithm introduced, and Sect. 3.3 explains the implementation of the same. Section 3.4 follows the special cases, and Sect. 3.5 explains the algorithm in steps.

3.1 Particle swarm optimization overview

The PSO is a population-based stochastic technique inspired by bird flocking, fish schooling and swarming theory in particular. It employs a swarm of particles that move through the search space while "adjusting" their position according to their fitness value and that of the other particles. This "adjustment" of position is rendered by recording the best personal position of the particle and that of the swarm called the best global position. These positions, personal and global best, correspond to the location of optima found by each particle and that by the swarm. By using Eq. (1) at every time step, these positions of optima and current positions are utilized to calculate the velocity of the particles mathematically. This velocity, in turn, determines the position of the particle in the next time step using Eq. (2).

$$\overrightarrow{v_i^{t+1}} = m\overrightarrow{v_i^t} + \mathrm{cr}_1^t\left[\overrightarrow{P_{\mathrm{best},i}^t} - \overrightarrow{x_i^t}\right] + sr_2^t\left[\overrightarrow{G_{\mathrm{best},i}^t} - \overrightarrow{x_i^t}\right] \quad (1)$$

$$\overrightarrow{x_i^{t+1}} = \overrightarrow{x_i^t} + \overrightarrow{v_i^{t+1}} \tag{2}$$

where v_i^t represents the velocity of drone i in at time t, x_i^t represents the position of drone i in at time t, P_{best}^t is the best personal position of drone i in from the initialization through time t. G_{best}^t is the best global position of particle i in found from initialization through time t; c and s are the constants that are used to level the contribution of cognitive and social component. r_1^t and r_2^t are the random numbers generated between 0 and 1 generated at time t. m is the inertial weight.

Using the above equations, particles will keep updating positions until a global optimum solution is reached. However, being a metaheuristic algorithm, PSO does not guarantee that an optimal solution will ever be reached. Though the above algorithm is attributed to particles, it can also be used for drones operating in a large search space as they can simply be abstracted as particles in the swarm. A modified version of PSO having with adaptive PSO constants, developed by Shi et al. [23], is used in this work.

3.2 Collision avoidance algorithm for UAV particle swarm optimization application

The idea of the unmanned aerial vehicles coherently operating in unique planes is exploited in this conceptualization. If each UAV operates in a plane whose vertical distance from other horizontal planes (in which other drones are operating) is sufficiently large, it will not collide with other UAVs. Such ideas have been presented in the literature [24]. However, their motive was not of proposing a collision avoidance model. This research attempts to make a collision avoidance algorithm based on a multiple-plane approach which offers the benefits of no collision while changing the altitude of the drones and no significant change in the original PSO algorithm by changing the trajectories of the drones for collision avoidance. The following assumptions were made while proposing the collision avoidance algorithm:

3.2.1 Assumptions for the surroundings

- Search is being performed in a relatively open area with a low density of obstacles per unit area like a mine quarry, marine environment, etc. Even if the area has a significant amount of obstacles, their height should be limited so that UAVs can hover over the obstacles.
- The environment is static, i.e. the conditions of the environment are not changing and the obstacle is stationary.
- Dimensions of the search space is known even though anything about the search space is not known, the number of obstacles, location of victims, etc. However, a rough qualitative idea of the density of obstacles can be known from the area of application like a forest is supposed to have more obstacles than a desert.

3.2.2 Unmanned aerial vehicle-related assumptions

- UAVs have short-range communication capabilities which make communication among the UAVs possible. Some UAVs among the swarm have long-range communication capability or at least have a beacon nearby to relay information about the ongoing SAR to a central station, where first responders are.
- UAVs are of moving rotor type. Fixed-wing UAVs are unsuitable for this algorithm as they lack hovering capability.
- UAVs are equipped with an obstacle detection system to detect obstacles at short ranges. UAVs can use LiDAR, SONAR, computer vision techniques, etc. to gauge the dimensions of the obstacles.

3.3 Implementation of multi-plane system

If the UAVs travel in different altitudes, it is almost certain that they cannot collide, but if there is some obstacle in the plane of operation of the drones, it would have to change its plane of operation and shift to a new plane. As soon as the obstacle is detected the PSO algorithm is stopped and the swarm switches to the anti-collision algorithm.

UAVs initially hover at a particular location; after encountering an obstacle, drones move to the nearest free plane and use it to cross the obstacle from above. All other UAVs, which are not facing any obstacle, remain stationary to avoid collision while some of the drones switch planes. After the drone crosses the obstacle and returns to its original plane, swarm switches to PSO to continue the search.

The number of the free planes is dependent on the number of the obstacles in the surroundings and the number of the UAVs operating in the swarm. A general formula is given in Eq. (3) for number of planes P in the system,

$$P = n + [n/a] \tag{3}$$

$$F = [n/a] \tag{4}$$

where n = number of UAVs in the swarm, a is a constant decided on the expected number of obstacles in the surroundings, [.] is the smallest integer function, F = number of free planes.

The difference in the altitude of the planes is such that the thrust exerted by the wind blown by the rotors is not affecting the motion of the other UAVs below them. The drones are distributed equally amidst the occupied in and equally distributed such that the number of occupied planes between any two free planes is same, as shown in

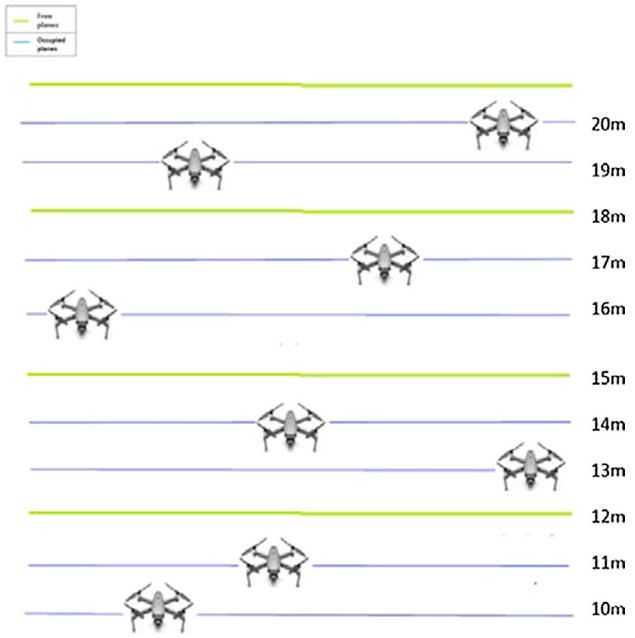


Fig. 2 An illustration of the distribution of free planes among occupied planes

Fig. 2. This is done to minimize the time of travel of each drone to the free plane and also increase the accessibility of the free planes. The number of occupied planes between each free plane is given by:

$$N_{\mathrm{O}} = \left[\frac{P}{[n/a]}\right]$$

The range at which the UAV detects the obstacle highly depends on the obstacle detection system installed, but the distance at which it has to switch to the anti-collision algorithm depends on the size and manoeuvrability of the UAV, as the anti-collision algorithm will be using this information to be able to plan the trajectory.

3.4 Special problems and solutions

The conceptualization presented in the above sections fail in certain scenarios or at least becomes inefficient. Some additional features are accommodated in the algorithm for these special conditions to make the system more robust. These conditions and their solutions are mentioned in Table 4:

3.5 Mathematical expression of the proposed algorithm

The UAVs in the modified particle swarm optimization algorithm constantly switch planes in order to avoid collisions and obstacles, the behaviour of which is decided by the following equations. In order to eliminate zero error for varying heights, z values are adjusted using the following equations

$$\overrightarrow{v_i^{t+1}} = m\overrightarrow{v_i^t} + \mathrm{cr}_1^t\left[\overrightarrow{P_{\mathrm{best},i}^t} - \overrightarrow{x_i^t}\right] + sr_2^t\left[\overrightarrow{G_{\mathrm{best},i}^t} - \overrightarrow{x_i^t}\right] \tag{5}$$

$$\overrightarrow{x_i^{t+1}} = \overrightarrow{x_i^t} + \overrightarrow{v_i^{t+1}} \tag{6}$$

$$F = [n/a] \tag{7}$$

$$P = n + [n/a] \tag{8}$$

$$Z_i^* = Z_i - \left(P_i * h\right) \tag{9}$$

$$S_i = \left\{\begin{array}{l} 0 \text{ if} \left(O_z - Z_i\right)^2 + \left(O_z - Y_i\right)^2 + \left(O_z - X_i\right)^2 < d^2 \\ 1 \text{ if} \left(O_z - Z_i\right)^2 + \left(O_z - Y_i\right)^2 + \left(O_z - X_i\right)^2 > d^2 \end{array}\right\} \tag{10}$$

where v_i^t, represents the velocity of ith drone at tth iteration, x_i^t represents the position of ith drone at tth iteration, P_{best}^t, is the best personal position of drone i till tth iteration for the objective function, G_{best}^t is the best global position of the swarm till tth iteration for the objective function t; c and s are the constants that are used to level the contribution of cognitive and social component, r_1^t and r_2^t are the random numbers generated between 0 and 1 generated at time t, m is the inertial weight, Z_i represents actual vertical position of

Table 4 Additional novel features of proposed algorithm for accommodating exceptional cases

Condition/situation	Remedy
If the lowest plane encounter many obstacles or if the terrain crosses a plane	The whole system of the planes shift in altitude by the height of the difference in altitude of two consecutive planes
If an obstacle is very tall (crosses more than three planes)	The whole system of the planes shift in altitude by the thrice height of the difference in altitude of two consecutive planes
In the event in which if 2 or more drones encounter same obstacle	The drones above will shift their position in their plane and let the drones in the lower planes cross over the obstacle

the ith UAV particle at tth iteration, Z_i^* represents zero error adjusted vertical position of the ith UAV at tth iteration, n = number of UAVs in the swarm, a is a constant decided on the expected number of obstacles in the surroundings, [.] is the smallest integer function, F = number of free planes, P_i represents plane position of ith drone, d represents the constant threshold value of Euclidean distance between the particle and obstacle that counts as a collision, O_z represents z position of the obstacle, O_y represents y position of the obstacle, O_x represents x position of the obstacle, h represents the constant height between 2 consecutive planes, S_i represents the switch plane function for the ith drone; if the value is 0, then the plane is not switched; if the value is 1, then drone shifts the plane to the nearest free plane.

Arrays are created to store plane information in the following syntax to achieve this task

$$\text{Total_Planes} = \left[\text{True/False, drones_in_plane, Plane_Index}\right]$$

wherein the first variable can hold two types of values—True and False. True represents that the plane is not free and that there are drones in the plane. False represents that there are no drones in the plane and the plane is free for drones to switch to. Second variable stores the ID number of drones in plane. Plane Index represents the position of the plane. For example, if the 2nd and 27th drones are in 7th plane, then the 7th element of the array would look like [True, [3, 7], 7].

3.6 Significance of threshold value *d*

Threshold value d is the maximum Euclidean distance between a UAV particle and its obstacle that counts as a collision. This threshold value depends on the physical attributes of the UAV, search area and obstacle size. For instance, a UAV that requires a safe spherical radius of operation as twice its diagonal length x will have the d value as 3×.

3.7 Algorithm steps

Based on the earlier discussion and equation used in section three, the algorithm with steps is presented in this subsection:

R1 Decide the number of drones and a factor for the environment.
R2 Calculate the number of total and free planes according to the equations using Eqs. (3) and (4) in the and velocity of the drones according to Eqs. (1) and (2) for initialization
R3 Allocate planes to drones and decide the base height (altitude of the bottom plane above ground) of the multiplane system. Also store the altitude of each of the drone in an array according to the plane of the respective drone.
R4 Evaluate the fitness value for each of the UAV for population in R3.
R5 Update velocity and positions for all the Drones using Eq. (1) and (2) and start motion of UAVs
R6 Continuously check for obstacles via obstacle detection during motion for each drone while in motion and
R7 If obstacle is encountered, then go to R8; otherwise, go to R22.
R8 Stop the motion of all the drones.
R9 Set variable m = number of drones detecting obstacles in their path.
R10 If m is not 0, then go to R11; otherwise, go to R19.
R11 Check whether groups of UAVs are facing the same obstacle or not; if yes, then go to R12; otherwise, go to R15
R12 For each group of shift planes for drone with higher altitude until its path is clear. Cross the obstacle and return to the original plane and put q=0 and put variable p=0
R13 Repeat R12 for next highest altitude UAV until the stack of UAVs is clear
R14 Assign number of drones that faced same obstacle and got their path cleared to p and set counter, $c = p$.
R15 Set $m = m - q - p$ and go to R10.
R16 For each one of the UAVs not facing the same obstacle, send them to a free plane with higher altitude until the path is clear. Cross the obstacle and return to the original plane and put variable p and $q = 0$
R17 Set q = number of UAVs facing different obstacle whose path is cleared
R18 Set $m = m - q - p$ and go to R10.
R19 If $c > 3$, then go to R20; otherwise, go to R21.
R20 Go to R3.
R21 Resume the motion of all Drones and go to R6.
R22 Evaluate the fitness function for each UAV and calculate personal and global best
R23 Is the convergence criteria reached? If yes, then go to R24; otherwise, go to R5.
R24 End.

3.8 Flowchart

Figure 3 represents a flowchart of the novel algorithm proposed and discussed in Sect. 3.5.

4 Simulation

This section focuses on the description of the simulation, developed using MATLAB to visualize a set of drones using the PSO with the proposed collision avoidance algorithm. The first of the three subdivisions discusses the kind of objective functions used and the rationale behind using those

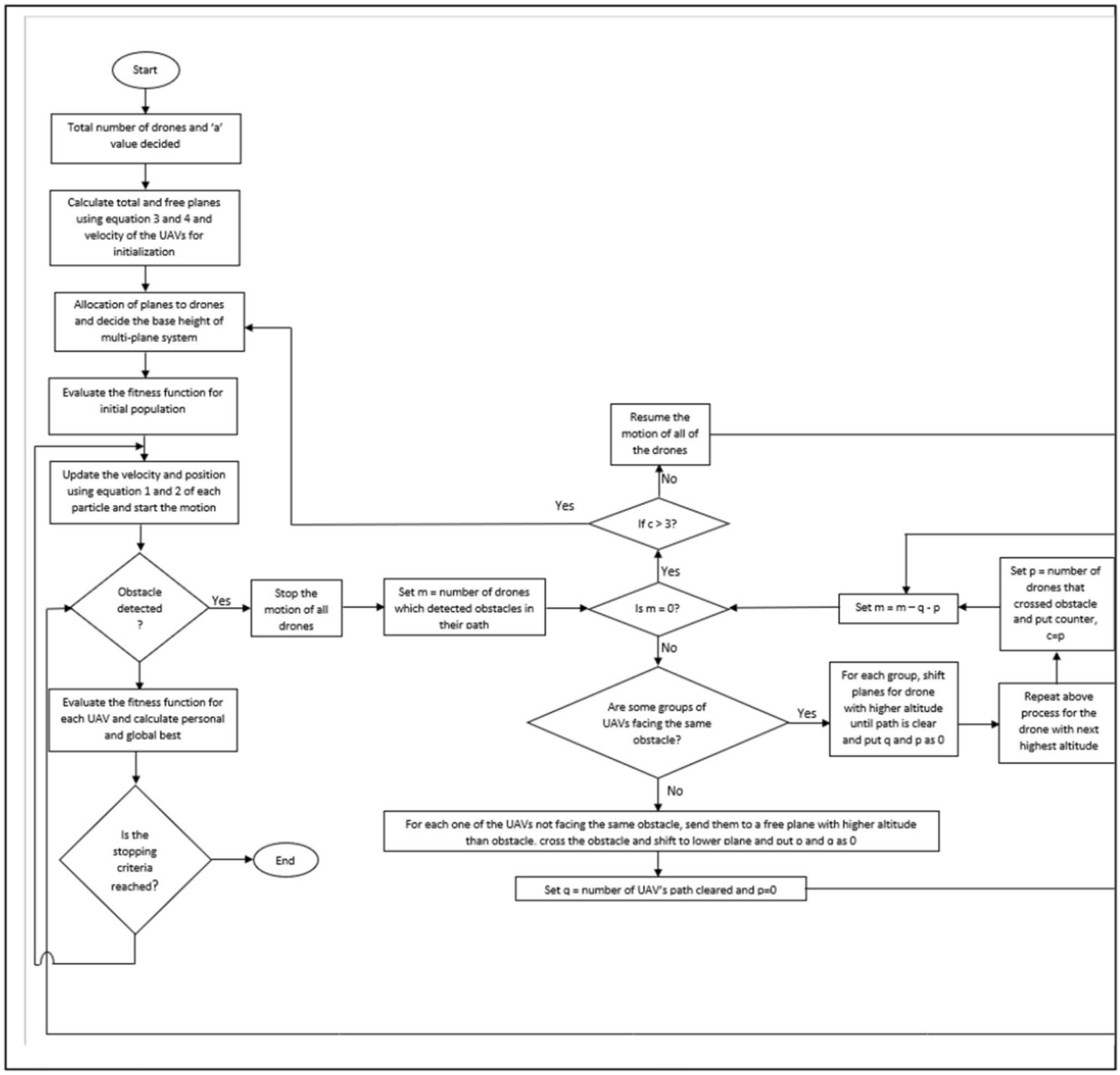


Fig. 3 Flowchart of the novel algorithm

functions, the second part mentions the specifications of the search domain and its bounds, and the third part discusses various considerations taken into account regarding UAVs and the anti-collision algorithm. The fourth part discusses how collisions were calculated in the simulations.

4.1 Objective functions for benchmarking

Finding a general objective function that can mimic the actual search and rescue mission in a simulation was a difficult task as each kind of search and rescue may use different types of inputs as fitness values. In the real world, various kinds of wireless technologies, which may aid in finding the victims, are available at disposal [25]. A few of these technologies are IEEE 802.11 standards [26]. Each of the drones can use, for example, strength of the distress signal being emitted by the survivors or use image processing for gauging the likeliness of finding the victims [27]. In either of the aforementioned ways (though not limited to the two), each drone is getting an input fitness value which can be used by the swarm to develop the strategy. Thus, a real-life search and rescue problem can also be called as an optimization problem in which the global optimum is the position of the victim.

Based on this argument, it was decided to use mathematical benchmarking functions, commonly used by the researchers as the objective functions for the simulations in addition to the typical constrained optimization mathematical functions.

4.2 Search area and bound

For this research, it was decided to use Rastrigin and Schwefel functions as the main objective functions as they are commonly used by researchers for benchmarking [28]. Moreover, randomized constrained optimization functions are used with different parameters to add for an exhaustive dataset and to imitate contour as shown in Fig. 6. The constants are randomized to generate random contours.

For the simulation, the contour of the function can be considered as the search space and the bounds as the limits of the search area where the drones move from point to point in search of the optimum position. For the sake of visualization, the three-dimensional surfaces of the functions can also be considered as the terrain over which UAVs fly, checking the fitness value at every location as they travel. The drones calculate the value of the objective function at each coordinate to get the fitness value. The distribution of the drones in the various planes in the beginning and their convergence over the minima at $x=0$, in the end, is illustrated in Figs. 4 and 5.

In Fig. 6: $Minf(x, y) = (1-x)^2 + 100\left(y - x^2\right)^2$

subjected to: $(x-1)^2 - y + 1 \leq 0$

$x + y - 1 \leq 0$

$-1.5 < x < 1.5$

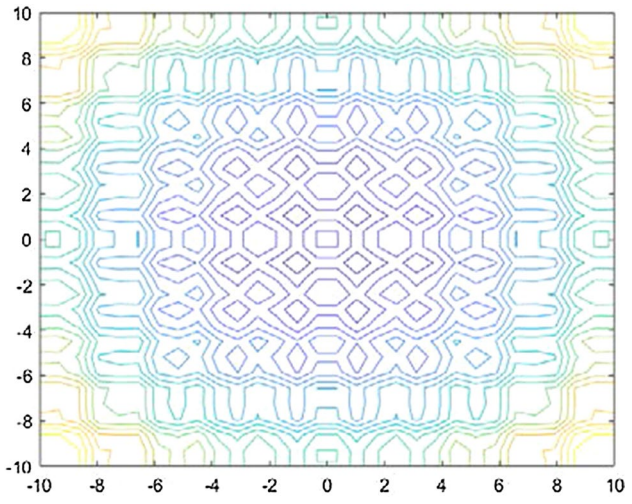


Fig. 4 Contour of Rastrigin function; gradient of yellow to purple denotes variation in depth wherein purple region is the deepest part and yellow region is the highest part (colour figure online)

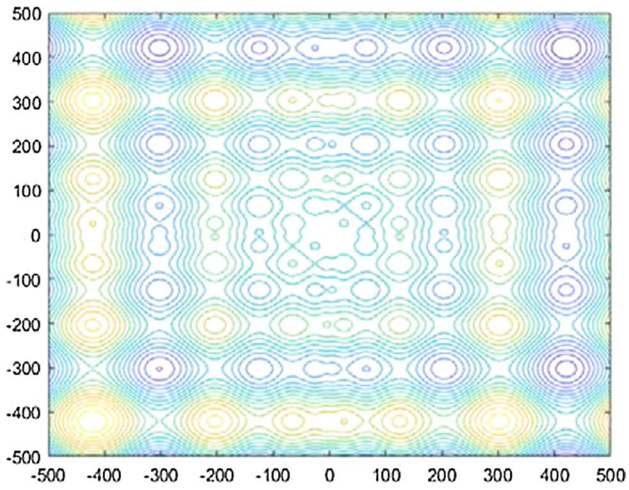


Fig. 5 Contour of Schwefel function; gradient of yellow to purple denotes variation in depth wherein purple region is the deepest part and yellow region is the highest part (colour figure online)

The constants are varied to get random constrained objective functions.

4.3 UAVs in the simulation

For the simulation, the following considerations were taken into for the surroundings and the UAV.

- If the final position of the UAV, calculated by the algorithm for the next iteration, is out of the bounded space, then the UAV's final position will be changed to the point

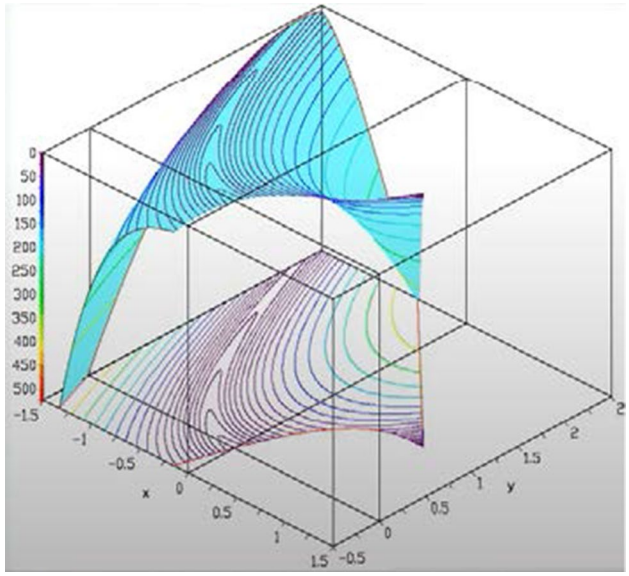


Fig. 6 Contour of randomized constrained optimization benchmarking function

in the domain, that is, on the line joining initial and the earlier final position (outside the bounded space) and the border of the space simultaneously.
- UAVs enter the search space from a launch site to the search space. UAVs then assume their positions, evenly distributed throughout the search space, before starting the actual search. This strategy helps in mitigating the problem of UAVs getting stuck in local optima.
- The "*a*" factor for the environment was taken as infinity to make the number of free planes equal to zero. This was done as there are no obstacles, other than the terrain, for the drones and they automatically shift the multi-plane system to an altitude that does not intersect with the terrain.

The UAVs use the combined PSO and collision avoidance algorithm in coherence as mentioned in the flowchart and algorithm steps of the methodology. The results obtained from the simulation and other inferences are presented in the later sections (Fig. 7).

4.4 Number of collisions

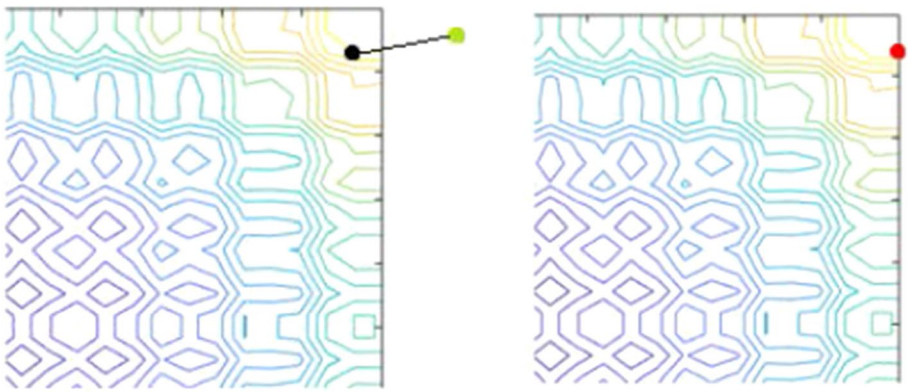
To calculate the number of collisions, cases were developed by varying swarm size, iterations, objective functions, height between the planes and the threshold value of Euclidean distance for the count of collisions. A total of 100 cases are made by randomly choosing the objective function among the ones listed in Sects. 4.1 and 4.2, randomly choosing swarm size between 10 and 1000, randomly choosing total iterations between 35 and 200, randomly varying height between the planes. Collisions were calculated had the UAVs been operating on original PSO and for the cases when UAVs were operating on the proposed algorithm. One sample case is shown in Fig. 8. The objective function for this case is the Rastrigin function, swarm size is 100, iterations are 100, height between planes is 0.1, and threshold Euclidean distance is 0.0001. The total original PSO collision is 17719, whereas total collisions with the proposed algorithm are 0. Collisions are calculated for each iteration using the threshold Euclidean distance; that is, if the Euclidean distance between a particle and an obstacle is less than the threshold value, then the particle is interpreted to have physically collided. The threshold distance can be varied according to the dimensions of. Similarly, 100 different cases with 200 data points are simulated, on the dataset of which the statistical z test is run to check the hypothesis that the proposed collision avoidance algorithm significantly reduced the number of collisions.

5 Results and discussions

In this section, the results of the simulated environment, described in the previous section, are presented. For the evaluation of the performance of the proposed modified UAV-PSO, several comparisons on different critical indicators, such as the probability of collisions and convergence plots, were made. All the comparison results clearly illustrate the effectiveness and advantages of the proposed improvements.

5.1 Probability of collisions

UAVs using PSO, without the proposed collision and obstacle avoidance system, are prone to a high probability



(a) (b)

Fig. 7 illustrates the case adjustment for the cases where the final position of a drone calculated by PSO is outside the boundary for the simulation. **a** represents one such case where black dot is the actual position of the UAV before the iteration and green dot is the calculated position after the iteration. Since the green dot is outside the boundary, the simulation changes the final position as shown in **b** where the red dot represents the intersection of the line joining initial (black dot) and the earlier final position (green dot) and the border (colour figure online)

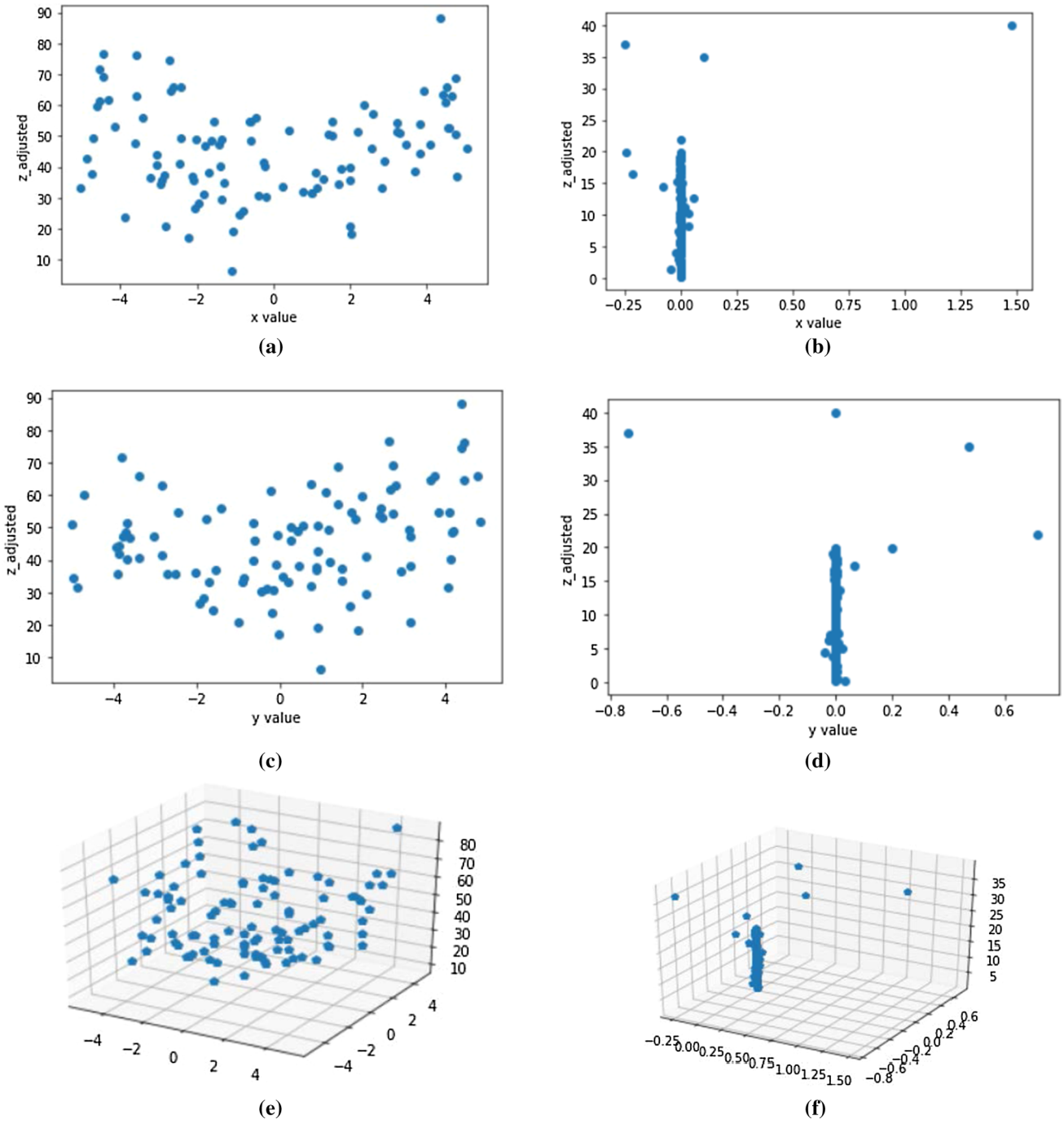


Fig. 8 **a**, **b**, **c** represents the initial state of the UAV particles before the proposed algorithm is run, whereas figure **d**, **e**, **f** represent the final state of UAV particle after all iterations. **a** and **b** represent the front view, **c** and **d** represent the side view, **d** and **f** represent 3d views

of collisions in physical search spaces which renders the use of unmanned aerial vehicles redundant. The proposed modifications reduce the probability of collisions between any two unmanned aerial vehicles to almost zero as the different UAVs operate in different non-intersecting planes shown in Fig. 8. However, challenging environments, such as spaces not having enough vertical room for the collision avoidance planes to exist and still the number of planes is not reduced, can increase the probability of collisions between the unmanned aerial vehicles as well. Moreover, the probability of the collision, while being low, also depends on how accurately the UAV's on-board obstacle detection system works.

The proposed obstacle avoidance system works for the case of static obstacles. For dynamic obstacles, the probability of avoiding collisions with obstacles decreases because the unmanned aerial vehicles are prone to collisions with obstacles in motion, after detection or while changing the planes. If a UAV detects dynamic obstacles, this algorithm can be used in conjunction with the dynamic path planning algorithms. The UAV can switch to a dynamic path planning algorithm for dodging a moving obstacle and use the obstacle avoidance, conceptualized in this work, for stationary obstacles. However, these dynamic path planning algorithms were not included in the current version.

5.2 Statistical test for collisions

To perform the statistical comparison, 100 samples were created by changing objective function, swarm size, iterations, the threshold of Euclidean distance between the obstacle and particle to count as collision, and the height between the planes.

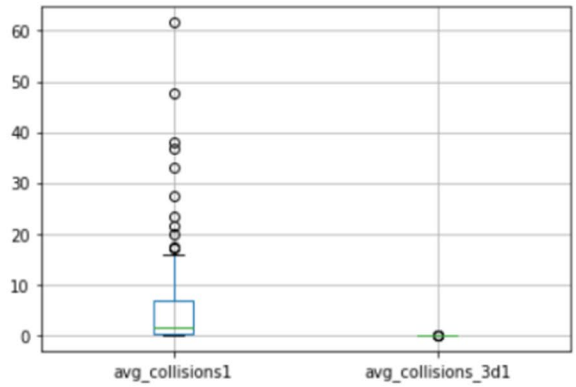
The statistical test used for comparison is the z-test. This Z-test is valid as the sample satisfies necessary conditions for the Z-test, i.e. data points are independent of each other, sample size greater than 30 and dataset is randomly selected from the population. Taking the null hypothesis that difference between the mean collisions per total iterations of PSO without collision avoidance algorithm and with collision avoidance algorithm is zero, i.e.

$$\text{Null hypothesis} \quad H_0 : \mu_1 - \mu_2 = 0$$
$$\text{Alternate hypothesis} \quad H_1 : \mu_1 - \mu_2 > 0$$

where μ_1 is mean total collisions per total iterations of samples with pure PSO, μ_2 is mean total collisions per total iterations of samples run on PSO with proposed collision avoidance.

In applying the statistic z-test, significance level = 0.05 is used. A modular collision avoidance algorithm is proposed to amend with PSO, the philosophy of which lies in the concept that each particle is given its own individual space while the PSO algorithm is run. This is achieved by a multi-plane approach wherein particles switch planes in difficult situations. The statistical Z-test was run on the mean of the total number of collisions in the original PSO and modified version of POS. p-value for the z test comes out to be $8.640335 \times 10^{-75.}$ With the results of the statistical z-test, it is concluded that there is significant evidence that the proposed approach significantly reduces the collisions of UAVs. The results of the statistical z test indicate giving each particle its own individual space to perform the search without altering the parent algorithm leads to reduced collisions by a large margin (Fig. 9).

```
In [6]: df.boxplot(column=['avg_collisions1','avg_collisions_3d1'])
Out[6]: <matplotlib.axes._subplots.AxesSubplot at 0x1db523b0e88>
```




```
In [10]: stats.ttest_rel(df.iloc[:,2],df.iloc[:,3])
Out[10]: Ttest_relResult(statistic=-53.40844335199274, pvalue=8.460335019249787e-75)
```

Fig. 9 Results of the statistical z test

5.3 Comparison of convergence

As shown in plot for average convergence in Fig. 10, the collision avoidance does not affect the convergence, there is no significant difference in the iterations to convergence in this conceptualization and modified PSO developed by Shi and Eberhart. The slight difference in the values calculated by the PSO with and without collision avoidance can be attributed to the random numbers, r_1^t and r_2^t, multiplied in Eq. (1) of Sect. 3. The value of the fitness value reaches within a tolerance of 0.001 to the global optima in around 40 and 220 iterations in approximately for both the objective functions. The exact value of the number of iterations to convergence changes for different runs of the algorithm.

Standardized residual plots give an idea of error in the global optima found by the algorithm more accurately as the values become asymptotic in the convergence plot as shown in Fig. 10.

The residuals show that there is no significant difference between the error in the values calculated by PSO and PSO with the anti-collision algorithm. The formula used for the calculation of standardized residuals, for each of the iterations, is mentioned below:

$$\mathrm{SR} = \frac{g_i - e}{\mathrm{S.D.E.}}$$

where SR = standardized residual, SDE = standardized error, g_i = Global minimum at the ith iteration and e = expected absolute minima.

5.4 Computational power and cost

For the case of zero obstacles in the search space, the computational power requirement for the proposed variant is almost equal to the computational power required of original particle swarm optimization. However, there is a requirement for additional memory to store the altitude of operation of each UAV. Practically, depending on the objective function, the UAVs might have to take images to process the fitness of a location, and cloud computing might be necessary for image processing. In such a case, the UAV at the higher altitude might need a camera with higher resolution to capture the location underneath accurately and better connectivity to send the data of larger size to the cloud server, increasing the overall cost.

For a case with obstacles, time for "crossing" the obstacle gets added to the total search time. This time depends on the density of static obstacles. As the number of obstacles increases, the time taken by the UAVs to complete the mission also increases.

5.5 Time complexity

Using the big-O notation, the time complexity of the original PSO is calculated as $O(m * N)$ where m is the population size and N is the number of dimensions. For the use case of search and rescue, N is generally 2 or 3, but at the cost of either additional time complexity of the native collision avoidance algorithm or at the cost of damage due to collisions. Original PSO cannot be implemented without a collision avoidance algorithm, which can significantly increase the time complexity and affect the searching algorithm path.

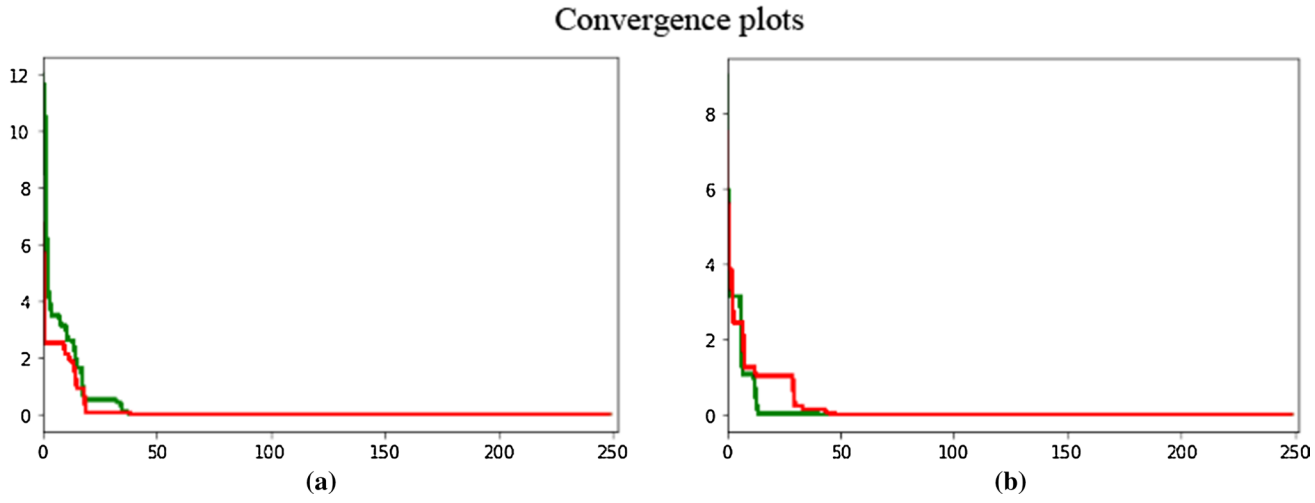


Fig. 10 **a** represents average fitness value plotted against number of iterations, **b** represents standardized residuals plotted against number of iterations, keeping swarm size and iterations constant, where green line depicts original PSO and red line depicts modified variant (colour figure online)

In the proposed algorithm, using the big O notation, time complexity comes out to be $O(m * P)$ where m is the population size and P is the number of planes according to the formula in Eq. (8), without compromising the parent PSO algorithm.

6 Conclusion

A novel collision avoidance algorithm is proposed for the application of a swarm of UAVs operating on particle swarm optimization for search and rescue operations. The collision avoidance algorithm is simulated for a total of 100 cases with 200 data points which were generated by varying swarm size, iterations, height between the planes, threshold Euclidean distance for collisions and objective functions. As the p-value of the statistical z test is 8.46×10^{-75}, results of the statistical z test run on the simulation dataset to favour the alternate hypothesis that the proposed algorithm significantly reduces the total number of collisions. The reason behind this result is that each UAV is allowed to move in its own independent space using a multi-plane approach without altering the parent algorithm. Moreover, the results of the simulations indicate that the new algorithm does not affect the efficacy of the original algorithm as the average of iterations to convergence and the trend of standardized residuals is the same for both original PSO and proposed algorithm.

6.1 Application areas

Potential applications of the proposed algorithm include various classes of search and rescue missions, wherein obstacles are mostly static. For example, maritime search and rescue cases are appropriate for the proposed algorithm as the environment for drones to operate in has minimum dynamic obstacles. Moreover, the use case of the proposed algorithm can be extrapolated to mining fields, exploring hazardous areas and other searching applications wherein a swarm of unmanned aerial vehicles can be used to search for the desired targets. The multi-plane approach of collision and obstacle avoidance can also be extrapolated to the newer metaheuristic algorithms featuring unique advantages.

6.2 Limitations and future works

The application of the proposed modified PSO model is limited to a simple search space with sufficient aerial space for the UAVs to switch planes. For a congested space, the algorithm is likely to fail or at least become highly inefficient. The collision avoidance which was integrated in the algorithm is a simple one and still requires provisions that would also prevent collisions with complex dynamic obstacles, such as birds. Till the time this new feature is incorporated, collision avoidance in this conceptualization can be used in combination with dynamic collision avoidance algorithm developed by other researchers. This way a bargain can be obtained between the advantages of both the algorithms. Also, near the final iterations of PSO, huge clusters form around the evident global minima. To avoid cluster formation, convergence algorithms can be proposed. The proposed algorithm in integration with the original PSO is prone to the original PSO's limitations of potentially falling in the local search trap; however, this can be solved in future works by integrating the proposed collision avoidance with hybrid versions of PSO and other metaheuristic algorithms as in the latest research works [29].

The model can also be extrapolated to search missions of various settings since the original searching algorithm remains theoretically unaffected by the inclusion of the collision avoidance algorithm. For example, in the field of mining, this implementation can be employed to search for the optimum mining zones in quarries. The use of UAVs would significantly reduce the search time. The proposed model can be extended to search on ocean beds as well where submarines can be treated as PSO particles working with the collision avoidance algorithm proposed in this paper.

In this work, the collision and obstacle avoidance algorithm is introduced for PSO. As a future line of work, the collision avoidance algorithm proposed in this document can be integrated with other algorithms. Moreover, the collision and avoidance algorithm will be extended to address the obstacles of high complexity. Further development to address dynamic obstacles can significantly widen the application domain of this algorithm but maintaining simplicity will still a challenge.

Author contributions All authors contributed to the study conception. Material preparation, coding and analysis were performed by AA, AD and AP. The first draft of the manuscript was written by GK, US and WKS. 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.

Code availability MATLAB and Jupyter Notebook software are used in this work.

Declarations

Conflict of interest The authors declare that there is no conflict of interest.

References

1. CRED, UNDRR, https://cred.be/sites/default/files/CredCrunch61-Humancost.pdf
2. Alotaibi E, Alqefari S, Koubaa A (2019) LSAR: multi-UAV collaboration for search and rescue missions. IEEE Access 7:55817–55832
3. Arnold R, Yamaguchi H, Tanaka T (2018) Search and rescue with autonomous flying robots through behavior-based cooperative intelligence. J Int Humanit Action 3:1–8
4. Lee K, Park J (2006) Application of particle swarm optimization to economic dispatch problem: advantages and disadvantages. In: 2006 IEEE PES power systems conference and exposition
5. Ghazali S, Anuar H, Zakaria S, Yusoff Z (2016) Determining position of target subjects in maritime search and rescue (MSAR) operations using rotary wing unmanned aerial vehicles (UAVs). In: 2016 international conference on information and communication technology (ICICTM)
6. Lee S, Morrison J (2015) Decision support scheduling for maritime search and rescue planning with a system of UAVs and fuel service stations. In: 2015 International conference on unmanned aircraft systems (ICUAS)
7. Rahmes M, Chester D, Hunt J, Chiasson B (2018) Optimizing cooperative cognitive search and rescue UAVs. In: Proceedings SPIE10643, autonomous systems: sensors, vehicles, security, and the internet of everything
8. Dinnbier N, Thueux Y, Savvaris A, Tsourdos A (2017) Target detection using Gaussian mixture models and fourier transforms for UAV maritime search and rescue. In: 2017 International conference on unmanned aircraft systems (ICUAS)
9. Kohlbrecher S, Meyer J, Graber T, Petersen K, Klingauf U, von Stryk O (2014) Hector open source modules for autonomous mapping and navigation with rescue robots. In: RoboCup 2013: robot world cup XVII. pp 624–631
10. Cui J, Phang S, Ang K, Wang F, Dong X, Ke Y, Lai S, Li K, Li X, Lin F, Lin J, Liu P, Pang T, Wang B, Wang K, Yang Z, Chen B (2015) Drones for cooperative search and rescue in post-disaster situation. In: 2015 IEEE 7th international conference on cybernetics and intelligent systems (CIS) and IEEE conference on robotics, automation and mechatronics (RAM).
11. Grissom C, Thomas F, James B (2006) Medical helicopters in wilderness search and rescue operations. Air Med J 25:18–25
12. Sun J, Li B, Jiang Y, Wen C (2016) A camera-based target detection and positioning UAV system for search and rescue (SAR) purposes. Sensors 16:1778
13. Al-Kaff A, Gómez-Silva M, Moreno F, de la Escalera A, Armingol J (2019) An appearance-based tracking algorithm for aerial search and rescue purposes. Sensors 19:652
14. XXXX (2021) Coverage algorithms for search and rescue with UAV drones. In: XIII AI*IA symposium on artificial intelligence, AIRO workshop
15. Wei C, Hindriks K, Jonker C (2016) Dynamic task allocation for multi-robot search and retrieval tasks. Appl Intell 45:383–401
16. Abdollahzadeh B, Gharehchopogh F, Mirjalili S (2021) African vultures optimization algorithm: a new nature-inspired metaheuristic algorithm for global optimization problems. Comput Ind Eng 158:107408. https://doi.org/10.1016/j.cie.2021.107408
17. Abdollahzadeh B, Soleimanian Gharehchopogh F, Mirjalili S (2021) Artificial gorilla troops optimizer: a new nature-inspired metaheuristic algorithm for global optimization problems. Int J Intell Syst 36(10):5887–5958. https://doi.org/10.1002/int.22535
18. Alhaqbani A, Kurdi H, Youcef-Toumi K (2020) Fish-inspired task allocation algorithm for multiple unmanned aerial vehicles in search and rescue missions. Remote Sens 13:27
19. Sánchez-García J, Reina D, Toral S (2019) A distributed PSO-based exploration algorithm for a UAV network assisting a disaster scenario. Future Gener Comput Syst 90:129–148
20. Duan H, Luo Q, Shi Y, Ma G (2013) Hybrid particle swarm optimization and genetic algorithm for multi-UAV formation reconfiguration. IEEE Comput Intell Mag 8:16–27
21. Ghamry K, Kamel M, Zhang Y (2017) Multiple UAVs in forest fire fighting mission using particle swarm optimization. In: 2017 international conference on unmanned aircraft systems (ICUAS)
22. Tsung-Ying S, Chih-Li H, Shang-Jeng T, Chan-Cheng L (2008) Optimal UAV flight path planning using skeletonization and particle swarm optimizer. In: 2008 IEEE congress on evolutionary computation (IEEE world congress on computational intelligence)
23. Shi W, Li J, Xu W, Zhou H, Zhang N, Zhang S, Shen X (2018) Multiple drone-cell deployment analyses and optimization in drone assisted radio access networks. IEEE Access 6:12518–12529
24. Hauert S, Winkler L, Zufferey J, Floreano D (2008) Ant-based swarming with positionless micro air vehicles for communication relay. Swarm Intell 2:167–188
25. Morgenthaler S, Braun T, Zhao Z, Staub T, Anwander M (2012) UAVNet: a mobile wireless mesh network using unmanned aerial vehicles. In: 2012 IEEE globecom workshops, 2012, pp 1603–1608
26. Carpin M, Rosati S, Khan M, Rimoldi B (2015) UAVs using bayesian optimization to locate WiFi devices. https://arxiv.org/abs/1510.03592
27. Hereford J, Siebold M, Nichols S (2007) Using the particle swarm optimization algorithm for robotic search applications. In: 2007 IEEE swarm intelligence symposium
28. Chou P, SauBor H (2011) Modified particle swarm optimization for sphere, rastrigin, schwefel and schafer fuctions. In: Proceedings of 2011 cross strait quad-regional radio science and wireless technology conference
29. Stacey A, Jancic M, Grundy I (2003) Particle swarm optimization with mutation. In: The 2003 congress on evolutionary computation. CEC '03
30. Peng Z, Li B, Chen X, Wu J (2012) Online route planning for UAV based on model predictive control and particle swarm optimization algorithm. In: Proceedings of the 10th world congress on intelligent control and automation
31. Geng Q, Zhao Z (2013) A kind of route planning method for UAV based on improved PSO algorithm. In: 2013 25th Chinese control and decision conference (CCDC)
32. Loscri V, Natalizio E, Mitton N (2014) Performance evaluation of novel distributed coverage techniques for swarms of flying robots. In: 2014 IEEE wireless communications and networking conference (WCNC)
33. Hayat S, Yanmaz E, Brown T, Bettstetter C (2017) Multi-objective UAV path planning for search and rescue. In: 2017 IEEE international conference on robotics and automation (ICRA)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

arXiv:2201.05811v1 [math.CV] 15 Jan 2022

ON A SUBFAMILY OF STARLIKE FUNCTIONS RELATED TO HYPERBOLIC COSINE FUNCTION

MRIDULA MUNDALIA AND S. SIVAPRASAD KUMAR

Abstract. A new subfamily of starlike functions $\mathcal{S}^*_\varrho$, is introduced. The class $\mathcal{S}^*_\varrho$ defined as

$$\mathcal{S}^*_\varrho := \left\{ f \in \mathcal{A} : \frac{zf'(z)}{f(z)} \prec \cosh\sqrt{z} =: \varrho(z), z \in \mathbb{D} \right\},$$

is associated with an analytic univalent function $\cosh\sqrt{z}$. We establish inclusion relations and sharp $\mathcal{S}^*_\varrho-$radii for some subclasses of analytic functions. Further we establish differential subordination implication results for the class $\mathcal{S}^*_\varrho$ and pointed out some special cases also.

1. Introduction

Let $\mathcal{H}$ be the class of all analytic functions defined on the unit disc $\mathbb{D} := \{z \in \mathbb{C} : |z| < 1\}$. Assume $\mathcal{A}_n \subset \mathcal{H}$ be a class of functions with Taylor series representation of the form $f(z) = z + a_{n+1}z^{n+1} + a_{n+2}z^{n+2} + \cdots$. For $n = 1$, let $\mathcal{A} := \mathcal{A}_1$. Let $\mathcal{S} \subset \mathcal{A}$ denote the class of univalent functions. If $f(z)$ and $g(z)$ are analytic functions in $\mathbb{D}$, then $f(z)$ is said to be subordinate to $g(z)$, written symbolically as $f \prec g$, if there exists a self-map $w(z)$ on $\mathbb{D}$ such that $w(0) = 0$ and $f(z) = g(w(z))$. For instance, if $g(z)$ is a univalent function in $\mathbb{D}$, then $f \prec g$ if and only if $f(0) = g(0)$ and $f(\mathbb{D}) \subset g(\mathbb{D})$. In 1992, Ma and Minda [16] investigated the following subclasses of $\mathcal{A}$ using the notion of subordination.

$$\mathcal{S}^*(\phi) = \left\{ f \in \mathcal{A} : \frac{zf'(z)}{f(z)} \prec \phi(z), z \in \mathbb{D} \right\}$$

and

$$\mathcal{C}(\phi) := \left\{ f \in \mathcal{A} : 1 + \frac{zf''(z)}{f'(z)} \prec \phi(z), z \in \mathbb{D} \right\}.$$

In the above defined classes, the expressions $zf'(z)/f(z)$ and $1 + z''(z)/f'(z)$ are subordinate to an analytic univalent function $\phi(z)$ such that $\phi'(0) > 0$ and $\operatorname{Re}\phi(z) > 0$ $(z \in \mathbb{D})$. Furthermore, $\phi(z)$ is symmetric with respect to the real axis and starlike with respect to $\phi(0) = 1$. Several authors have previously handled the Ma and Minda classes for various choices of $\phi(z)$, see [5, 6, 7, 8, 12, 21, 27, 28]. For instance, Janowski in [9] defined the class $\mathcal{S}^*[A, B] := \mathcal{S}^*((1 + Az)/(1 + Bz))$, a subclass of starlike functions, when $\phi(z) = (1 + Az)/(1 + Bz)$ $(-1 \leq B < A \leq 1)$. Especially for $A = 1 - 2\beta$ $(0 \leq \beta < 1)$ and $B = -1$, we get the class of starlike functions of order β, namely $\mathcal{S}^*(\beta)$, discovered by Robertson in [29]. The class $\mathcal{C}(\alpha)$ of convex functions of order α $(0 \leq \alpha < 1)$ is defined in a similar way. Infact for $\phi(z) = (1 + z)/(1 - z)$, we get the well-known class of starlike functions $\mathcal{S}^* := \mathcal{S}^*((1 + z)/(1 - z))$ and convex functions $\mathcal{C} := \mathcal{C}((1 + z)/(1 - z))$. It is worth mentioning that the classes $\mathcal{S}^*(\phi)$ and $\mathcal{C}(\phi)$, respectively, unify numerous subclasses of $\mathcal{S}^*$ and $\mathcal{C}$. The class $\mathcal{S}^*(q_\kappa) := \mathcal{S}^*(\sqrt{1 + \kappa z})$ where $0 < \kappa \leq 1$, was introduced by Sokół in [3]. It's important to note that $q_\kappa(\mathbb{D})$ lies in $\Omega_\kappa = \{u : |u^2 - 1| = \kappa\}$, which is the right loop of the Cassinian Ovals. For $\kappa = 1$, Sokół and Stankiewicz [31] introduced and examined the class $\mathcal{S}^*_L := \mathcal{S}^*(\sqrt{1 + z})$, related

2010 *Mathematics Subject Classification.* 30C45, 30C80.

Key words and phrases. Univalent functions, Starlike functions, Radius problems, Differential subordination.

The first author is supported by Delhi Technological University, New Delhi.

with the lemniscate of Bernoulli. In 2015, Raina and Sokół [22] introduced a new subclass of starlike functions, namely $S_q^* := \mathcal{S}^*(z+\sqrt{1+z^2})$. Interestingly $\mathcal{S}_q^*$ deals with a crescent like region $\Omega_q = \left\{u : |u^2-1|< 2|u|\right\}$, obtained as a consequence of $\left\{u : |u-1|< \sqrt{2}\right\} \cap \left\{u : |u+1|> \sqrt{2}\right\}$. Further, the class $\mathcal{S}_e^* := \mathcal{S}^*(e^z)$ was studied by Mendiratta et al. [18], which is related with the region $\Omega_e := \{u : |\log u|< 1\}$. Recently Kanas et al. [17] worked on the class $\mathcal{S}_s^*$, corresponding to a region bounded by concentric Limaçon's, defined as

$$\mathcal{S}_s^* := \left\{ f \in \mathcal{A} : \frac{zf'(z)}{f(z)} \prec (1+sz)^2 =: \mathcal{L}_s(z) \right\} \quad (-1 \leq s \leq 1).$$

The class $\mathcal{SS}^*(\beta)$ introduced by Stankiewicz [19], obtained when $\phi(z) = ((1+z)/(1-z))^\beta$, is the well known class of strongly starlike functions of order β $(0 \leq \beta < 1)$, where each $f \in \mathcal{A}$ is determined by the condition $|\arg(zf'(z)/f(z))|< \beta$. The classes $\mathcal{S}_e^*, \mathcal{S}_s^*$, $\mathcal{S}^*(q_\kappa)$, $\mathcal{S}^*[A,B]$ and $\mathcal{SS}^*(\beta)$ were widely studied in [3, 9, 17, 18, 25].

For the present investigation we have examined the function $\varrho_\sigma(z) := \cosh \sigma\sqrt{z}$ where $\sigma \in [-\pi/2, \pi/2] - \{0\}$. Note that $\varrho_\sigma(z)$ is an analytic univalent function with $\operatorname{Re} \varrho_\sigma(z) > 0$ and maps $\mathbb{D}$ onto a convex region. Further it is symmetric with respect to the real axis such that $\varrho_\sigma'(0) = \sigma^2/2 > 0$. Consequently, $\varrho_\sigma(z)$ is a Ma-Minda type function for each $\sigma \in [-\pi/2, \pi/2] - \{0\}$. In the recent years, cosine and cosine hyperbolic functions have been investigated, see [2, 4]. Note that $\varrho(z) := \varrho_1(z) = \cosh\sqrt{z}$, $\phi_1(z) = \cos z$ and $\phi_2(z) = \cosh z$ have identical images, however $\phi_1(z)$ $(\phi_1'(0) < 0)$ and $\phi_2(z)$ are non-univalent in $\mathbb{D}$, whereas $\varrho(z)$ $(\varrho'(0) = 1/2 > 0)$ is analytic univalent in $\mathbb{D}$. Thus the geometry of $\varrho_\sigma(z)$ piqued our interest in formulating the following definition, by means of subordination.

Definition 1.1. Let $\mathcal{S}_{\varrho_\sigma}^*$ be the class of normalized starlike functions, defined as follows:

$$\mathcal{S}_{\varrho_\sigma}^* := \left\{ f \in \mathcal{A} : \frac{zf'(z)}{f(z)} \prec \varrho_\sigma(z) := \cosh \sigma\sqrt{z}, z \in \mathbb{D} \right\} \quad (\sigma \in [-\pi/2, \pi/2] - \{0\}).$$

Note that $\overline{\varrho_\sigma(z)} = \varrho_\sigma(\overline{z})$ for each $z \in \mathbb{D}$, the domain $\varrho_\sigma(\mathbb{D})$ is symmetrical about real-axis. Under the conformal mapping $\varrho_\sigma : \mathbb{D} \to \mathbb{C}$, the unit disc $\mathbb{D}$ is mapped onto the region

$$\Omega_{\varrho_\sigma} := \{u \in \mathbb{C} : |\log(u + \sqrt{u^2-1})|^2 < \sigma^2\} \quad (\sigma \in [-\pi/2, \pi/2] - \{0\}),$$

defined on the principle branch of logarithm and square root functions. For each $\sigma \leq \hat{\sigma}$, observe that $\varrho_\sigma(\mathbb{D}) \subset \varrho_{\hat{\sigma}}(\mathbb{D})$. Note that the function $\varrho_\sigma(z)$ satisfies the following property for each circle $|z|= r$, $0 \leq r < 1$,

$$\begin{cases} \min\limits_{|z|=r} \operatorname{Re} \varrho_\sigma(z) = \min\limits_{|z|=r} |\varrho_\sigma(z)| = \varrho_\sigma(\sqrt{-r}) \\ \\ \max\limits_{|z|=r} \operatorname{Re} \varrho_\sigma(z) = \max\limits_{|z|=r} |\varrho_\sigma(z)| = \varrho_\sigma(\sqrt{r}). \end{cases} \tag{1.1}$$

Assume $\varrho(z) := \varrho_1(z)$. It is interesting to note that the series expansion of $\varrho(z)$ can be represented as a confluent hypergeometric limit function, namely ${}_0F_1\left(;1/2;z/4\right)$

$$\sum_{k=0}^{\infty} \frac{z^k}{(2k)!} = {}_0F_1\left(;1/2;z/4\right) = \varrho(z).$$

Infact $\varrho(z)$ is a linearly independent solution of the linear differential equation, $4zv''(z) + 2v'(z) - v(z) = 0$. In the present investigation we shall restrict our major workings to the following subclass of starlike functions,

$$\mathcal{S}_\varrho^* := \left\{ f \in \mathcal{A} : \frac{zf'(z)}{f(z)} \prec \varrho(z) := \cosh\sqrt{z}, z \in \mathbb{D} \right\}.$$

We determine radius constants and differential subordination implication results for the class $\mathcal{S}^*_\varrho$ along with some inclusion relations.

Example 1. *Let* $\psi_{\hat{\varrho}}(z) = 1 + z/3 + z^2/18$ *and* $\phi_{\hat{\varrho}}(z) = 1 + \sin(z/3)$. *Clearly* $\psi_{\hat{\varrho}}(z), \phi_{\hat{\varrho}}(z) \prec \varrho(z)$. *The corresponding functions*

$$f_1(z) = z\exp\left(\frac{z}{3} + \frac{z^2}{36}\right)$$

and

$$f_2(z) = ze^{Si(z)}, \text{ where } Si(z) = \int_0^z \frac{\sin t}{t}dt$$

belong to $\mathcal{S}^*_\varrho$.

In terms of integral representation, we have $f \in \mathcal{S}^*_\varrho$ if and only if

$$f(z) = z\exp\left(\int_0^z \frac{\hat{\varrho}(t) - 1}{t}dt\right) \tag{1.2}$$

where $\hat{\varrho}(z) \prec \varrho(z)$. Now using the representation in (1.2) we obtain different functions, those work as extremal functions for various results. If we consider $\varphi_{\varrho_n} \in \mathcal{A}$, $(n = 2, 3, 4...)$ satisfying

$$\frac{z\varphi'_{\varrho_n}(z)}{\varphi_{\varrho_n}(z)} = \varrho(z^{n-1}),$$

then $\varphi_{\varrho_n} \in \mathcal{S}^*_\varrho$ $(n = 2, 3, 4...)$. Note that for $n = 2$,

$$\varphi_{\varrho_2}(z) = z\exp\left(2(Chi(\sqrt{z}) - \log\sqrt{z} - \xi)\right),$$

where $Chi(\sqrt{z})$ is the cosine hyperbolic integral defined as

$$Chi(\sqrt{z}) = \xi + \log z + \int_0^z \frac{\cosh\sqrt{t} - 1}{t}dt$$

and $\xi \approx 0.5772$ is the Euler-Mascheroni constant. We denote $\varphi_\varrho := \varphi_{\varrho_2}$. For completeness of our class $\mathcal{S}^*_\varrho$, we give below a remark using the results of [16, 21].

Remark 1.2. For $f \in \mathcal{S}^*_\varrho$ and $|z| = r_0 < 1$, we have

(i) $-\varphi_\varrho(-r_0) \leq |f(z)| \leq \varphi_\varrho(r_0)$ (Growth Theorem).
(ii) $\varphi'_\varrho(-r_0) \leq |f'(z)| \leq \varphi'_\varrho(r_0)$ (Distortion Theorem).
(iii) $|\arg(f(z)/z)| \leq \max_{|z|=r_0} \arg(\varphi_\varrho(z)/z)$ (Rotation Theorem) .

Equality for (i)-(iii) holds for some $z_0 \neq 0$ if and only if $f(z)$ is a rotation of $\varphi_\varrho(z)$. If $f \in \mathcal{S}^*_\varrho$ then either f is a rotation of $\varphi_\varrho(z)$ or

$$f(\mathbb{D}) \supset \{v : |v| \leq -\varphi_\varrho(-1) = e^{-2(\xi - Ci(1))} \approx 0.619016\},$$

where $Ci(1)$ is the cosine integral defined as $Ci(1) = \xi + \int_0^1 (\cos\sqrt{t} - 1)/2t \; dt$ and ξ is the Euler-Mascheroni constant.

Further, from the results in [21] for each $f \in \mathcal{S}^*_\varrho$, (i) $|a_2| \leq 1/2$, (ii) $|a_3| \leq 1/4$, (iii) $|a_4| \leq 1/6$ and (iv) for any complex constant μ, the Fekete-Szegö functional is bounded by

$$|a_3 - \mu a_2^2| \leq \frac{1}{4}\max\left\{1, \left|\mu - \frac{7}{12}\right|\right\}.$$

These estimates are sharp. Equality in (i) holds for the function $\varphi_\varrho(z)$ and $\tilde{f}(z) = z + z^3/4$ is an extremal function for (ii) and (iv).

2. Examples and Properties of Hyperbolic Cosine function

We begin with a Lemma which demonstrates a maximal disc centered at a point $(c, 0)$ on the real line, that can be subscribed within $\varrho_\sigma(\mathbb{D})$.

Lemma 2.1. *Suppose $\sigma \neq 0$, then $\varrho_\sigma(z)$ satisfies the following inclusion*

$$\{u \in \mathbb{C} : |u - c| < r_{\sigma c}\} \subset \varrho_\sigma(\mathbb{D}) =: \Omega_{\varrho_\sigma} \quad (-\pi/2 \leq \sigma \leq \pi/2),$$

where

$$r_{\sigma c} = \begin{cases} c - \cos\sigma, & \cos\sigma < c \leq (\cosh\sigma + \cos\sigma)/2 \\ \cosh\sigma - c, & (\cosh\sigma + \cos\sigma)/2 \leq c < \cosh\sigma. \end{cases}$$

Proof. Let $\Gamma := \varrho_\sigma(e^{it})$ $(-\pi \leq t \leq \pi)$ be the boundary curve of the function $\varrho_\sigma(z)$. Clearly due to symmetricity of curve Γ with respect to the real-axis, it is enough to consider $0 \leq t \leq \pi$. Define $G_c(\tau)$ as follows:

$$G_c(\tau) := \left(c - \cosh\left(\sigma(\cos\tau)\right)\cos\left(\sigma(\sin\tau)\right)\right)^2 + \sinh^2\left(\sigma(\cos\tau)\right)\sin^2\left(\sigma(\sin\tau)\right),$$

where $\tau = t/2$. Observe that $G_c(\tau)$ (see Fig. 1 for different values of c) is the square of the distance from point $(c, 0)$ to Γ.

Case 1: For $\cos\sigma < c \leq 1$, $G_c(\tau)$ is monotonically decreasing on $[0, \pi/2]$, then

$$r_{\sigma c} = \min_{\tau \in [0,\pi/2]} \sqrt{G_c(\tau)} = \sqrt{G_c(\pi/2)} = c - \cos\sigma.$$

Case 2: When $1 \leq c \leq \sigma_0$, where $\sigma_0 < (\cosh\sigma + \cos\sigma)/2$ is a point at which $G_c(\tau)$ changes its character i.e $G_c(\tau)$ is monotonically decreasing for $1 \leq c \leq \sigma_0$ and has three critical points $\{0, \tau_{\tilde{c}}, \pi/2\}$ for $\sigma_0 < c \leq (\cosh\sigma + \cos\sigma)/2$, where $\tau_{\tilde{c}} \in (0, \pi/2)$ is the only root of the equation

$$\begin{aligned} &2c\tan\tau\cos(\sigma\sin\tau)\sinh(\sigma\cos\tau) + 2c\sin(\sigma\sin\tau)\cosh(\sigma\cos\tau) \\ &= \sin(2\sigma\sin\tau) + \tan\tau\sinh(2\sigma\cos\tau). \end{aligned} \tag{2.1}$$

Note that $\tau_c < \tau_{\tilde{c}}$ whenever $c < \tilde{c}$. Further

$$G_c(0) - G_c(\pi/2) = (\cosh\sigma - \cos\sigma)(\cos\sigma + \cosh\sigma - 2c) \geq 0.$$

Therefore this yields

$$r_{\sigma c} = \min_{\tau \in [0,\pi/2]} \left\{\sqrt{G_c(0)}, \sqrt{G_c(\tau_{\tilde{c}})}, \sqrt{G_c(\pi/2)}\right\} = \sqrt{G_c(\pi/2)} = c - \cos\sigma.$$

Case 3: For $(\cosh\sigma + \cos\sigma)/2 \leq c \leq \sigma_1$, where $\sigma_1 < \cosh\sigma$ is a point at which $G_c(\tau)$ changes its character i.e $G_c(\tau)$ has three critical points $\{0, \tau_{\hat{c}}, \pi/2\}$, where $\tau_{\hat{c}} \in (0, \pi/2)$ is the only root of equation (2.1) and $G_c(\tau)$ is an increasing function for $\sigma_1 < \sigma < \cosh\sigma$. Infact $G_c(0) \leq G_c(\pi/2)$. Therefore

$$r_{\sigma c} = \min_{\tau \in [0,\pi/2]} \left\{\sqrt{G_c(0)}, \sqrt{G_c(\tau_{\hat{c}})}, \sqrt{G_c(\pi/2)}\right\} = \sqrt{G_c(0)} = \cosh\sigma - c.$$

Hence the result follows. ∎

The inclusion results in Lemma 2.2, associated with Ω_{ϱ_σ}, can be easily deduced using equation (1.1) and Lemma 2.1.

Lemma 2.2. *For the region $\Omega_{\varrho_\sigma} := \varrho_\sigma(\mathbb{D})$, the following inclusion relations hold:*

(i) $\{u : |u - (\cosh\sigma + \cos\sigma)/2| < (\cosh\sigma - \cos\sigma)/2\} \subset \Omega_{\varrho_\sigma}$.

(ii) $\Omega_{\varrho_\sigma} \subset \{u : \cos\sigma < \operatorname{Re} u < \cosh\sigma\}$ *and* $\Omega_{\varrho_\sigma} \subset \{u : \cos\sigma < |u| < \cosh\sigma\}$.

(iii) $\Omega_{\varrho_\sigma} \subset \{u : |\operatorname{Im} u| < l\}$ *and* $\Omega_{\varrho_\sigma} \subset \{u : |u - (\cosh\sigma + \cos\sigma)/2| < l\}$, *where* $l = |\operatorname{Im}(\cosh(\sigma e^{it_0/2}))|$, *and t_0 is the root of the equation*

$$\cos\sigma + \cosh\sigma - 2\cos\left(\sigma\sin\left(t/2\right)\right)\cosh\left(\sigma\cos\left(t/2\right)\right) = 0.$$

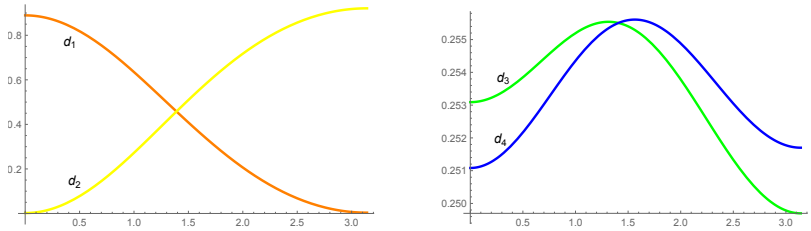


FIGURE 1. Graphs of $G_{0.6}(\tau) = d_1$, $G_{1.5}(\tau) = d_3$, $G_{1.04}(\tau) = d_3$, $G_{1.042}(\tau) = d_4$, (with $\sigma = 1$)

As a result of Lemma 2.1 (for $\sigma = 1$), we deduce the following theorem for the region $\Omega_{\varrho_1} =: \Omega_\varrho$.

Theorem 2.3. *The region* $\Omega_\varrho := \varrho(\mathbb{D}) \supset \{u \in \mathbb{C} : |u - c| < r_c\}$ *where*

$$r_c = \begin{cases} c - \cos 1, & \cos 1 < c \leq (\cosh 1 + \cos 1)/2 \\ \cosh 1 - c, & (\cosh 1 + \cos 1)/2 \leq c < \cosh 1. \end{cases} \tag{2.2}$$

Remark 2.4. Observe that Theorem 2.3 ensures that $D_c := |u - c| < r_c$, is the maximal disc subscribed in $\varrho(\mathbb{D})$, when $c = (\cosh 1 + \cos 1)/2$ and $r_c = (\cosh 1 - \cos 1)/2$. Thus $D_c \subset \varrho(\mathbb{D})$.

Note that for the sake of simplicity, throughout the paper we shall assume $c_0 := \cos 1$ and $c_1 := \cosh 1$. In the subsequent result we establish certain inclusion relations associated with the region Ω_ϱ.

Lemma 2.5. *For the region* $\Omega_\varrho := \varrho(\mathbb{D})$, *we have the following inclusion relations:*

(i) $\{u : |u - (c_0 + c_1)/2| < (c_1 - c_0)/2\} \subset \Omega_\varrho$.
(ii) $\Omega_\varrho \subset \{u : |\arg u| < m\}$, *where* $m \approx 0.506053 \approx (0.322163)\ \pi/2 \approx 28.9947°$.
(iii) $\Omega_\varrho \subset \{u : c_0 < \operatorname{Re} u < c_1\}$ *and* $\Omega_\varrho \subset \{u : c_0 < |u| < c_1\}$.
(iv) $\Omega_\varrho \subset \{u : |\operatorname{Im} u| < l\}$ *and* $\Omega_\varrho \subset \{u : |u - (c_0 + c_1)/2| < l\}$, *where* $l = |\operatorname{Im}(\cosh(e^{it_0/2}))|$ *and* t_0 *is the solution of the equation*

$$c_0 + c_1 - 2\cos(\sin(t/2))\cosh(\cos(t/2)) = 0.$$

Proof. We can obtain (i), $(iii) - (iv)$ from equations in (1.1), Remark 2.4 and Lemma 2.2 (for $\sigma = 1$). For part (ii) let $\Gamma := \partial(\varrho(z)) = \varrho(e^{it})$, $-\pi \leq t \leq \pi$, represents the boundary curve of $\varrho(z)$. Assume that

$$\operatorname{Re} \varrho(e^{it/2}) = \cos(\sin(t/2))\cosh(\cos(t/2)) =: X(t)$$

and

$$\operatorname{Im} \varrho(e^{it/2}) = \sin(\sin(t/2))\sinh(\cos(t/2)) =: Y(t).$$

Consider

$$\begin{aligned} |\arg \varrho(z)| &< \max_{|z|=1} |\arg \varrho(z)| = \max_{t\in[-\pi,\pi]} |\arg \varrho(e^{it/2})| = \max_{t\in[-\pi,\pi]} \tan^{-1}(Y(t)/X(t)) \\ &= \max_{t\in[-\pi,\pi]} \tan^{-1}(\tan(\sin(t/2))\tanh(\cos(t/2))) := m(t) \end{aligned}$$

Observe that $\tan^{-1} x$ is a monotonically increasing real valued function. Therefore it is enough to obtain the maximum of $m(t)$. The roots of

$$\begin{aligned} m'(t) = 0.5(&\cos(t/2)\tanh(\cos(t/2))\sec^2(\sin(t/2)) \\ &- \sin(t/2)\tan(\sin(t/2))\operatorname{sech}^2(\cos(t/2))) = 0 \end{aligned}$$

are $t_1 \approx -1.91672$ and $t_2 \approx 1.91672$. As $t_1 < t_2$, therefore maximum of $m(t)$ is attained at $t = t_2$. Hence the result for (ii) follows. ∎

In Theorem 2.6 and Corollary 2.7 we prove inclusion results pertaining to various classes along with the classes $\mathcal{ST}_p(\gamma)$, $\mathcal{S}^*_{hpl}(s)$, $k-\mathcal{ST}$ and $\mathcal{M}(\beta)$ [1, 10, 11, 32] defined below:

$$\mathcal{M}(\beta) := \left\{ f \in \mathcal{A} : \frac{zf'(z)}{f(z)} \prec \frac{1+(2\beta-1)z}{1+z}, \beta > 1 \right\},$$

$$\mathcal{ST}_p(\gamma) = \left\{ f \in \mathcal{A} : \operatorname{Re} \frac{zf'(z)}{f(z)} + \gamma > \left| \frac{zf'(z)}{f(z)} - \gamma \right|, \gamma > 0 \right\},$$

$$\mathcal{S}^*_{hpl}(s) := \left\{ f \in \mathcal{A} : \frac{zf'(z)}{f(z)} \prec (1-z)^{-s} = e^{-s\log(1-z)}, 0 < s \leq 1 \right\},$$

$$k-\mathcal{ST} := \left\{ f \in \mathcal{A} : \operatorname{Re} \frac{zf'(z)}{f(z)} > k \left| \frac{zf'(z)}{f(z)} - 1 \right|, k \geq 0 \right\},$$

Theorem 2.6. *Let $f \in \mathcal{S}^*_{\varrho_\sigma}$ then for each $\sigma \in [-\pi/2, \pi/2] - \{0\}$, following inclusions hold:*

(i) $\mathcal{S}^*_{\varrho_\sigma} \subset S^*(\zeta)$, *where* $\zeta = \cos\sigma$
(ii) $\mathcal{S}^*_{\varrho_\sigma} \subset \mathcal{M}(\beta)$, *where* $\beta = \cosh\sigma$.
(iii) $\mathcal{S}^*_{q_\kappa} \subset \mathcal{S}^*_{\varrho_\sigma}$, *whenever* $\kappa \leq 1 - \cos^2\sigma$.
(iv) $k-\mathcal{ST} \subset \mathcal{S}^*_{\varrho_\sigma}$, *whenever* $k \geq \cosh\sigma/(\cosh\sigma - 1)$.
(v) $\mathcal{S}^*_{\varrho_\sigma} \subset \mathcal{S}^*_{hpl}(s)$, *whenever* $\log(\sec\sigma)/\log 2 \leq s \leq 1$, $\sigma \in [-\pi/3, \pi/3] - \{0\}$.
(vi) $\mathcal{S}^*_{\varrho_\sigma} \subset \mathcal{S}^*_L(s)$, *whenever* $1 - \sqrt{\cos\sigma} \leq s \leq \frac{1}{\sqrt{2}}$.

Proof. Observe that, in equation (1.1), when r tends to 1^-, sharp bounds on real part and modulus of $\varrho_\sigma(z)$ are obtained. As a result of this observation and Lemma 2.2, we conclude that the inclusion results in $(i)-(ii)$ hold true for the class $\mathcal{S}^*_{\varrho_\sigma}$. We know that $q_\kappa(z) = \sqrt{1+\kappa z}$ where $0 < \kappa \leq 1$, is associated with the region $|u^2 - 1| < \kappa$. Therefore part (iii) can be easily established as $q_\kappa(\mathbb{D})$ lies in Ω_{ϱ_σ}, if and only if, $\sqrt{1-\kappa} \geq \cos\sigma$, which implies $\kappa \leq 1 - \cos^2\sigma$. For part (iv), consider the curve $\Gamma_k = \{u \in \mathbb{C} : \operatorname{Re} u > k|u-1|\}$, where $k \geq 0$. When $k > 1$, the set Γ_k represents the interior of an ellipse,

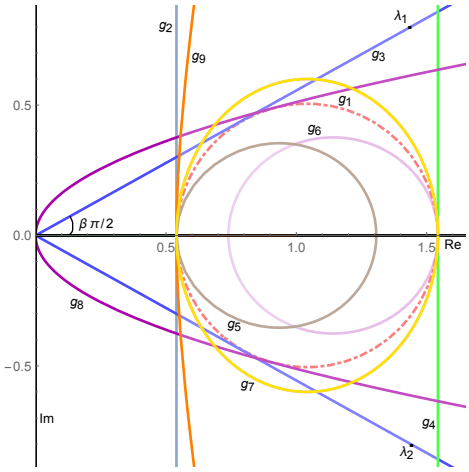
$$\gamma_k := \left\{ (x, y) : \frac{(x-x_1)^2}{a_1^2} + \frac{y^2}{b_1^2} = 1 \right\},$$

where $x_1 = k^2/(k^2-1)$, $a_1 = k/(k^2-1)$ and $b_1 = 1/\sqrt{k^2-1}$. For γ_k to lie in Ω_{ϱ_σ} we must have $x_1 + a_1 \leq \cosh\sigma$, which gives a sufficient condition for γ_k to lie in Ω_{ϱ_σ}, this leads us to the required condition. From [10] we know that $\operatorname{Re}(1-z)^{-s} > 2^{-s}$. Therefore for (v) to hold true $2^{-s} \leq \cos\sigma$, which gives $\log(\sec\sigma)/\log 2 \leq s \leq 1$, provided $-\pi/3 \leq \sigma \leq \pi/3$. Furthermore, it was demonstrated in [17], that $L_S(\mathbb{D}) \supset \{u : |u-1| < 1 - (1-s)^2\}$, where $0 < s \leq 1/\sqrt{2}$. Thus for (vi) to hold true we must have $1 - (1-s)^2 \geq 1 - \cos\sigma$. We conclude $\mathcal{S}^*_{\varrho_\sigma} \subset \mathcal{S}^*_L(s)$ for each $s \geq 1 - \sqrt{\cos\sigma}$. ∎

In the following Corollary we prove inclusion results for the class $\mathcal{S}^*_\varrho$.

Corollary 2.7. *For each function $f \in \mathcal{S}^*_\varrho$ the following inclusions hold:*

(i) $\mathcal{S}^*_\varrho \subset \mathcal{S}^*(c_0)$.
(ii) $\mathcal{S}^*_\varrho \subset \mathcal{M}(c_1)$.
(iii) $\mathcal{S}^*_\varrho \subset \mathcal{SS}^*(\beta)$, *where* $\beta \approx 0.3222163$.
(iv) $\mathcal{S}^*_{q_\kappa} \subset \mathcal{S}^*_\varrho$, *whenever* $\kappa \leq 1 - c_0^2$.
(v) $k-\mathcal{ST} \subset \mathcal{S}^*_\varrho$, *whenever* $k \geq c_1/(c_1-1)$.
(vi) $\mathcal{S}^*_\varrho \subset \mathcal{S}^*_{hpl}(s)$, *whenever* $-\log c_0/\log 2 \leq s \leq 1$.
(vii) $\mathcal{S}^*_\varrho \subset \mathcal{S}^*_L(s)$, *whenever* $1 - \sqrt{c_0} \leq s \leq \frac{1}{\sqrt{2}}$
(viii) $\mathcal{S}^*_\varrho \subset \mathcal{ST}_p(\gamma)$, *whenever* $\gamma \geq \gamma_0 \approx 0.0654238$.



Legend -

$g_1 : \varrho(z) = \cosh\sqrt{z}$

$g_2 : \operatorname{Re} u = c_0$

$g_3 : |\arg u| = \beta\pi/2$

$\arg\lambda_1 = -\arg\lambda_2 = \beta\pi/2$

$\beta \approx 0.322163$

$g_4 : \operatorname{Re} u = c_1$

$g_5 : \sqrt{1+(1-c_0^2)z}$

$g_6 : \operatorname{Re} u = \frac{c_1}{c_1-1}|u-1|$

$g_7 : \frac{(\operatorname{Re} u - \frac{c_0+c_1}{2})^2}{(\frac{c_1-c_0}{2})^2} + \frac{(\operatorname{Im} u)^2}{(c_2)^2}$,

$c_2 = 0.65$

$g_8 : \operatorname{Re} u + \gamma = |u-\gamma|$,

$\gamma \approx 0.0654238$

$g_9 : \frac{1}{(1-z)^{s_0}}$, $s_0 = \frac{\log c_0^{-1}}{\log 2}$

FIGURE 2. Inclusion graphs in context of Corollary 2.7 associated with $\varrho(z)$.

Proof. Clearly parts $(i)-(ii)$ and $(iv)-(vii)$ can be obtained as a result of Theorem 2.6 for $\sigma = 1$. Part (iii) holds true due to Lemma 2.5, for the class $\mathcal{S}_\varrho^*$ (see Fig. 2). For $(viii)$ in order to show $\mathcal{S}_\varrho^* \subset \mathcal{ST}_p(\gamma)$, we must have $|u-\gamma| - \operatorname{Re} u < \gamma$, where $u(z) = \cosh\sqrt{z}$. For $z = e^{it}$ we have

$$H(\tau) := \frac{\sin^2(\sin\tau)\sinh^2(\cos\tau)}{4\cos(\sin\tau)\cosh(\cos\tau)} < \gamma,$$

where $\tau = t/2$. Clearly $H'(\tau)$ vanishes on $\{0, \tilde{\tau}, \pi/2\}$, with $\tau = \tilde{\tau} \approx 0.832934$ as the only root of the equation

$$\begin{aligned}\tan(\sin\tau)\tanh(\cos\tau)((\cos\tau(\cos(2\sin\tau)+3)\sinh(\cos\tau)\sec(\sin\tau))\\ -\sin\tau\sin(\sin\tau)(\cosh(2\cos\tau)+3)\operatorname{sech}(\cos\tau)) = 0\end{aligned}$$

in $(0, \pi/2)$. Therefore $\max_{\tau\in[0,\pi/2]} H(\tau) = H(\tilde{\tau}) \approx 0.0654238$. Observe that $\mathcal{ST}_p(\gamma_1) \subset \mathcal{ST}_p(\gamma_2)$ whenever $\gamma_1 < \gamma_2$. Thus this leads us to the required inclusion relation. ∎

Remark 2.8. Fig. 2 displays various inclusion relations related to the region $\Omega_\varrho := \Omega_{\varrho_1}$. Note that a vertical ellipse enclosing the region Ω_ϱ is $(x-x_2)^2/a_2^2 + y^2/b_2^2 = 1$, where $x_2 = c_0/2$, $a_2 = c_1/2$ and $c_2 \geq \max \operatorname{Im}\varrho(z)$. For visual purposes we have illustrated this ellipse (g_7), for $c_2 = 0.65$. Note that Fig. 2 depicts the sharpness of inclusion results in Corollary 2.7.

Let $\mathcal{P}_n(\alpha)$ denote the class of functions $p(z)$ of the type $p(z) = 1 + p_n z^n + p_{n+1}z^{n+1} + \ldots$ such that $\operatorname{Re} p(z) > \alpha$ $(0 \leq \alpha < 1)$. Clearly the class $\mathcal{P}_n(\alpha) \subset \mathcal{P}_n$. For $\alpha = 0$, $\mathcal{P}_n := \mathcal{P}_n(0)$. If a function $p(z)$ of the form $p(z) = 1 + p_n z^n + p_{n+1}z^{n+1} + \cdots$ satisfies $p(z) \prec (1+Az)/(1+Bz)$, for $A \neq B$ and $|B| \leq 1$, then $p \in \mathcal{P}_n[A, B]$. We state a few lemmas in connection with these classes.

Lemma 2.9. [23] *If $p \in \mathcal{P}_n[A, B]$, then for $|z| = r$*

$$\left| p(z) - \frac{1-ABr^{2n}}{1-B^2r^{2n}} \right| \leq \frac{|A-B|r^n}{1-B^2r^{2n}}.$$

Particularly, if $p \in \mathcal{P}_n(\alpha)$, *then*

$$\left| p(z) - \frac{1+(1-2\alpha)r^{2n}}{1-r^{2n}} \right| \leq \frac{2(1-\alpha)r^n}{1-r^{2n}}.$$

Lemma 2.10. [26] *If* $p \in \mathcal{P}_n(\alpha)$, *then for* $|z|= r$

$$\left| \frac{zp'(z)}{p(z)} \right| \leq \frac{2(1-\alpha)nr^n}{(1-r^n)(1+(1-2\alpha)r^n)}$$

Theorem 2.11. *Let* $p(z) = (1+Az)/(1+Bz)$, *where* $-1 < B < A \leq 1$, *then* $p(z) \prec \cosh\sqrt{z}$, *if and only if*

$$A \leq \begin{cases} 1-(1-B)c_0 & \text{if } 2(1-AB) \leq (c_0+c_1)(1-B^2) \\ (1+B)c_1 - 1 & \text{if } 2(1-AB) \geq (c_0+c_1)(1-B^2). \end{cases} \tag{2.3}$$

Proof. We know from Lemma 2.9 that the transformation $p(z) = (1+Az)/(1+Bz)$, where $-1 < B < A \leq 1$, maps $\mathbb{D}$ onto the disc

$$\left| p(z) - \frac{1-AB}{1-B^2} \right| \leq \frac{A-B}{1-B^2}.$$

By Theorem 2.3, $p(z) \prec \cosh\sqrt{z}$ if and only if the above disc lies within Ω_ϱ. The conditions in (2.3) gives $(1+A) \leq (1+B)c_1$, whenever $2(1-AB) \geq (c_0+c_1)(1-B^2)$ holds. The factor $(1-B)$ when multiplied to $(1+A) \leq (1+B)c_1$ gives $(A-B)/(1-B^2) \leq c_1 - (1-AB)/(1-B^2)$, which leads to $(A-B)/(1-B^2) \leq c_1 - c$ whenever $2c \geq c_0 + c_1$ where $c = (1-AB)/(1-B^2)$. Also from (2.3), $(1-A) \geq (1-B)c_0$ whenever $2(1-AB) \leq (c_0+c_1)(1-B^2)$. Equivalently $(A-B)/(1-B^2) \leq c - c_0$ whenever $2c \leq c_0 + c_1$. This establishes the fact that $p(z)$ lies in the disc $|u-c| < r_c$, where r_c is as given in (2.2). Thus the proof follows. ∎

Corollary 2.12. *If conditions* (2.3) *given in Theorem 2.11 hold for* $-1 < B < A \leq 1$, *then* $\mathcal{S}^*[A,B] \subset \mathcal{S}^*_\varrho$.

3. Radius Problems

Let X and Y be two subclasses of $\mathcal{A}$. The radius problems corresponding to these classes deals with determining R, the least upper bound of all numbers r such that for each $f \in Y$, $r^{-1}f(rz) \in X$ for each $0 < r \leq R$. The number $R =: \mathcal{R}_X(Y)$ is called the X-radius for the class Y. Similarly Y-radius for the class X can be defined. If there exists a function $F \in Y$ such that $r^{-1}F(rz) \notin X$ for $r > R$, then F is called the extremal function, thus R is called the sharp X-radius for the class Y. In the field of geometric function theory, radius problems have been an active area of research. Some of the pioneering work in this direction have been discussed by several authors, see [3, 18, 23, 25, 31]. Motivated by the aforestated work, we derive radius results involving some subclasses of $\mathcal{A}$ and $\mathcal{A}_n$. Further we define the classes $\mathcal{S}^*_n(\varrho)$, $\mathcal{S}^*_n[A,B]$ $(-1 \leq B < A \leq 1)$ and $\mathcal{M}_n(\beta)$ as follows:

$$\mathcal{S}^*_n(\varrho) = \left\{ f \in \mathcal{A}_n : \frac{zf'(z)}{f(z)} \prec \cosh\sqrt{z} =: \varrho(z) \right\},$$

$$\mathcal{S}^*_n[A,B] = \left\{ f \in \mathcal{A}_n : \frac{zf'(z)}{f(z)} \prec \frac{1+Az}{1+Bz} \right\}$$

and

$$\mathcal{M}_n(\beta) = \left\{ f \in \mathcal{A}_n : \frac{zf'(z)}{f(z)} \prec \frac{1+(1-2\beta)z}{1-z}, \beta > 1 \right\}.$$

In the sequel we apply lemmas stated in Section 2, so as to obtain sharp $\mathcal{S}^*_n(\varrho)-$radius, $\mathcal{S}^*_n[A,B]-$radius and $\mathcal{M}_n(\beta)-$radius for the class $\mathcal{S}^*_\varrho$. The following theorem is obtained as a consequence of Lemma 2.5 and equations in (1.1).

Theorem 3.1. *The class $\mathcal{S}^*_\varrho \subset \mathcal{M}(\beta)$ for $|z|< r_\beta$, where*

$$r_\beta = \begin{cases} r(\beta), & 1 < \beta < c_1 \\ 1, & \beta \geq c_1. \end{cases}$$

and $r(\beta) \in (0,1)$ is the smallest root of the equation $\cosh\sqrt{r} = \beta$. Equality holds when $f(z) = \varphi_\varrho(z)$.

Theorem 3.2. *Suppose $f \in \mathcal{S}^*_\varrho$, then $f \in \mathcal{S}^*(\zeta)$, where $\zeta \in [0,1)$ and $0 \leq r \leq r_\zeta < 1$ is the least positive root of the equation $\cos\sqrt{r} = \zeta$. This radius result is sharp.*

Proof. As $f \in \mathcal{S}^*_\varrho$, then we have $zf'(z) = f(z)\cosh\sqrt{w(z)}$, where $w(z)$ is a Schwarz function with $w(0) = 0$ such that for $-\pi \leq t \leq \pi$, $w(z) = Re^{it}$. Due to the symmetricity of $\varrho(z)$ with respect to the real axis, we restrict the range of t to $0 \leq t \leq \pi$. Clearly for each $R = |w(z)| \leq |z| = r < 1$, we get $\cos\sqrt{R} \geq \cos\sqrt{r}$, and as a result of equations in (1.1)

$$\operatorname{Re}\frac{zf'(z)}{f(z)} \geq \min_{|z|=r} \operatorname{Re}\varrho(w(z)) = \cos\sqrt{r} \geq \zeta.$$

If $s(r,\zeta) := \cos\sqrt{r} - \zeta$, then there exist $r_{\zeta_0} < r_{\zeta_1}$ such that $s(r_{\zeta_0},\zeta) > 0$ and $s(r_{\zeta_1},\zeta) < 0$, holds. Thus a least positive root r_ζ for the equation $s(r,\zeta) = 0$, will serve the purpose. In particular, at $z_0 = -r$, we have $\operatorname{Re}(z_0\tilde{f}'(z_0)/\tilde{f}(z_0)) = \cos\sqrt{r} = \zeta$, then function $\tilde{f}(z) = \varphi_\varrho(z)$ is the extremal function. ∎

Theorem 3.3. *Let $f \in \mathcal{S}^*_\varrho$, then $f \in \mathcal{C}(\alpha)$, where $\alpha \in [0,1)$, provided $|z| \leq r_0$, where $r_0 \in [0,1)$ is the least positive root of the following equation,*

$$2(1-r^2)\cos\sqrt{r} - \sqrt{r}\tan\sqrt{r} = \alpha.$$

Proof. As $f \in \mathcal{S}^*_\varrho$, then by definiton of $\mathcal{S}^*_\varrho$, we have

$$\frac{zf'(z)}{f(z)} = \cosh\sqrt{w(z)} \tag{3.1}$$

where $w(z)$ is a Schwarz function such that $w(0) = 0$. On logarithmically differentiating (3.1) and applying triangular inequality we deduce that

$$\begin{aligned} \operatorname{Re}\left(1 + \frac{zf''(z)}{f'(z)}\right) &= \operatorname{Re}\frac{zf'(z)}{f(z)} + \operatorname{Re}\left(\frac{zw'(z)\tanh\sqrt{w(z)}}{2\sqrt{w(z)}}\right) \\ &\geq \cos\sqrt{r} - |z||w'(z)|\left|\frac{\tanh\sqrt{w(z)}}{2\sqrt{w(z)}}\right| \quad (|z| = r < 1). \end{aligned} \tag{3.2}$$

By Schwarz Pick Lemma, we have

$$-|z||w'(z)|\left|\frac{\tanh\sqrt{w(z)}}{\sqrt{w(z)}}\right| \geq -|z|\frac{1-|w(z)|^2}{1-|z|^2}\left|\frac{\tanh\sqrt{w(z)}}{\sqrt{w(z)}}\right|. \tag{3.3}$$

Assume $w(z) = Re^{it}$, $t \in [-\pi,\pi]$ where $R \leq r$, then inequality (3.3) yields

$$\operatorname{Re}\left(\frac{zw'(z)\tanh\sqrt{w(z)}}{2\sqrt{w(z)}}\right) \leq \frac{\sqrt{r}\tan\sqrt{r}}{2(1-r^2)}. \tag{3.4}$$

Thus from inequalities (3.2) and (3.4) we conclude that

$$\operatorname{Re}\left(1 + \frac{zf''(z)}{f(z)}\right) \geq \cos\sqrt{r} - \frac{\sqrt{r}\tan\sqrt{r}}{2(1-r^2)}.$$

Hence the least positive root of the equation of $2(1-r^2)\cos\sqrt{r}-\sqrt{r}\tan\sqrt{r}=\alpha$ will serve the purpose. ∎

Theorem 3.4. *The sharp $\mathcal{S}_n^*(\varrho)-$radius for the class $\mathcal{S}_n^*[A,B]$ is*

(i) $\mathcal{R}_{\mathcal{S}_n^*(\varrho)}(\mathcal{S}_n^*[A,B])=\min\{1;((1-c_0)/(A-Bc_0)^{1/n}\}=:\mathcal{R}_0$, *where* $0\le B<A\le 1$.

(ii) $\mathcal{R}_{\mathcal{S}_n^*(\varrho)}(\mathcal{S}_n^*[A,B])=\begin{cases}\mathcal{R}_0, & \mathcal{R}_0\le\mathcal{R}_1\\ \mathcal{R}_2, & \mathcal{R}_0>\mathcal{R}_1,\end{cases}$ *where* $-1\le B<0<A\le 1$.

where

$$\mathcal{R}_1=\left(\frac{c_0-2}{B(c_0B-2A)}\right)^{1/2n},\quad \mathcal{R}_2=\min\left\{1;\left(\frac{c_1-1}{A-Bc_1}\right)^{1/n}\right\}.$$

Proof. Let $p(z)=zf'(z)/f(z)$ lies in $|p(z)-c|<R$, where

$$c=\frac{1-ABr^{2n}}{1-B^2r^{2n}}\quad\text{and}\quad R=\frac{(A-B)r^n}{1-B^2r^{2n}}.$$

For the case when $B\ge 0$, it follows that $c\le 1$. As $f\in\mathcal{S}_n^*(\varrho)$, then by Theorem 2.3 and Lemma 2.9 we have

$$\frac{(A-B)r^n}{1-B^2r^{2n}}\le\frac{1-ABr^{2n}}{1-B^2r^{2n}}-c_0.$$

On solving the above inequality we get $r\le R_0$. Equality here holds for $f(z)=\tilde{f}(z)$ is of the following form

$$\tilde{f}(z)=\begin{cases}z(1+Bz^n)^{(A-B)/nB}, & B\ne 0\\ z\exp\left(Az^n/n\right), & B=0.\end{cases}\tag{3.5}$$

This proves the first part. Now assume $-1\le B<0<A\le 1$ and $\mathcal{R}_0\le\mathcal{R}_1$, then $c\le(c_0+c_1)/2$ if and only if $r\le\mathcal{R}_1$. Therefore, for $0\le r\le\mathcal{R}_0$, we deduce that $c\le(c_0+c_1)/2$. Due to Theorem 2.3 for each $f\in\mathcal{S}_n^*(\varrho)$, we have $(A-B)r^n/(1-B^2r^{2n})\le c-c_0$ which is equivalent to $r\le\mathcal{R}_0$. Furthermore assume that $\mathcal{R}_0>\mathcal{R}_1$. Note that $c\ge(c_0+c_1)/2$ if and only if $r\ge\mathcal{R}_1$. In particular for $r\ge\mathcal{R}_0$, we have $c\ge(c_0+c_1)/2$. Thus by Theorem 2.3 for each $f\in\mathcal{S}_n^*(\varrho)$, the inequality, $(A-B)r^n/(1-B^2r^{2n})\ge c_1-c$ is equivalent to $r\le\mathcal{R}_2$. The function $\tilde{f}(z)$ given in (3.5) works as the extremal function for this case as well. ∎

Theorem 3.5. *Let $\beta>1$, then the sharp $\mathcal{S}_n^*(\varrho)-$radius for the class $\mathcal{M}_n(\beta)$, is given by*

$$\mathcal{R}_{\mathcal{S}_n^*(\varrho)}(\mathcal{M}_n(\beta))=\left(\frac{1-c_0}{2\beta-(1+c_0)}\right)^{1/n}.$$

Proof. Since $f\in\mathcal{M}_n(\beta)$, then by definition of $\mathcal{M}_n(\beta)$, we have $zf'(z)/f(z)\prec(1+(1-2\beta)z)/(1-z)$. Now $(1+(1-2\beta)r^{2n})/(1-r^{2n})\le 1$ for each $\beta>1$. Further by Lemma 2.9, we get

$$\left|\frac{zf'(z)}{f(z)}-\frac{1+(1-2\beta)r^{2n}}{1-r^{2n}}\right|\le\frac{2(\beta-1)r^n}{1-r^{2n}}.$$

On applying Theorem 2.3, we have

$$\frac{2(\beta-1)r^n}{1-r^{2n}}\le\frac{1+(1-2\beta)r^{2n}}{1-r^{2n}}-c_0$$

or equivalently $r^{2n}((1-2\beta)+c_0)-2(\beta-1)r^n+1-c_0\ge 0$, which gives $r\le\mathcal{R}_{\mathcal{S}_n^*(\varrho)}(\mathcal{M}_n(\beta))$. The required extremal function is

$$\tilde{f}(z)=\frac{z}{(1-z^n)^{2(1-\beta)/n}},$$

as $z_0\tilde{f}'(z_0)/\tilde{f}(z_0)=c_0$ at the point $z_0=\mathcal{R}_{\mathcal{S}_n^*(\varrho)}(\mathcal{M}_n(\beta))$. ∎

Recently, Lecko et al. [14] investigated the classes $\mathrm{Re}(1-z^2)f(z)/z>0$ and $\mathrm{Re}(1-z)^2f(z)/z>0$, involving the starlike functions $z/(1-z^2)$ and $z/(1-z)^2$. In 2019, Cho et al.[5] estimated radii constants for classes characterised by the ratio of two analytic functions $f(z)$ and $g(z)$ along with some conditions associated with $g(z)$, namely $\mathrm{Re}\, g(z)/z>\alpha$ where $\alpha=0$ or $1/2$, such that $\mathrm{Re}\, f(z)/g(z)>0$. Further for $\mathrm{Re}\, g(z)/z>0$, the inequality $|f(z)/g(z)-1|<1$ was also discussed by Cho et al.[5]. Likewise Khatter et al. [15], Kumar et al. [12] and Kumar [13] have also discussed similar expressions. Motivated by these classes, here below we define some subclasses of $\mathcal{A}_n$, where $g\in\mathcal{A}_n$,

$$\mathfrak{F}_1(\beta):=\left\{f\in\mathcal{A}_n:\left|\frac{f(z)}{g(z)}-1\right|<1 \text{ and } \mathrm{Re}\,\frac{g(z)}{z}>\beta\right\}\quad(\beta\in\{0,1/2\})$$

and

$$\mathfrak{F}_2:=\left\{f\in\mathcal{A}_n:\left|\frac{f(z)}{g(z)}-1\right|<1 \text{ and } g\in\mathcal{C}\right\}.$$

Definition 3.6. Let $-1\le A\le 1$ and $g\in\mathcal{A}_n$, then for each $n=1,2,\ldots,$ $\mathfrak{F}_3\subset\mathcal{A}_n$, be defined as:

$$\mathfrak{F}_3:=\left\{f\in\mathcal{A}_n:\mathrm{Re}\,\frac{f(z)}{g(z)}>0 \text{ and } \mathrm{Re}\,\frac{(1-z^n)^{(1+A)/n}g(z)}{z}>0\right\}.$$

Remark 3.7. The functions $\tilde{f}(z)=z(1+(1-2\beta)z^n)$ and $\tilde{g}(z)=z(1+(1-2\beta)z^n)/(1-z^n)$ defined on $\mathbb{D}$ satisfies $|(\tilde{f}(z)/\tilde{g}(z))-1|=|z|^n<1$ and $\mathrm{Re}\,\tilde{g}(z)/z=\mathrm{Re}(1+(1-2\beta)z^n)/(1-z^n)>\beta$. Therefore $\tilde{f}\in\mathfrak{F}_1(\beta)$, where $\beta\in\{0,1/2\}$. If $\tilde{f}(z)=z(1+z^n)/(1-z^n)^{1/n}$ and $\tilde{g}(z)=z/(1-z^n)^{1/n}$, then $f\in\mathfrak{F}_2$. Similarly when $\tilde{f}(z)=z(1+z^n)^2/(1-z^n)^{2+(1+A)/n}$ and $\tilde{g}(z)=z(1+z^n)/(1-z^n)^{1+(1+A)/n}$, then $\tilde{f}\in\mathfrak{F}_3$. Therefore class $\mathfrak{F}_3$ is non-empty.

Theorem 3.8. *The sharp $\mathcal{S}_n^*(\varrho)-$ radii for the classes $\mathfrak{F}_1(0),\mathfrak{F}_1(1/2)$ and $\mathfrak{F}_2$, are respectively given by*

(i) $$\mathcal{R}_{\mathcal{S}_n^*(\varrho)}(\mathfrak{F}_1(0))=\left(\frac{\sqrt{9n^2-4(c_0-1)(1+n-c_0)}-3n}{2(1+n-c_0)}\right)^{1/n}.$$

(ii) $$\mathcal{R}_{\mathcal{S}_n^*(\varrho)}(\mathfrak{F}_1(1/2))=\left(\frac{1-c_0}{2n-(c_0-1)}\right)^{1/n}.$$

(iii) $$\mathcal{R}_{\mathcal{S}_n^*(\varrho)}(\mathfrak{F}_2)=\left(\frac{\sqrt{1+n(n+6)+4c_0(c_0-(1+n))}-(1+n)}{2(n-c_0)}\right)^{1/n}.$$

Proof. Assume $f(z)/g(z)=p_1(z)$ and $g(z)/z=p_2(z)$, where $f(z)$ and $g(z)$ are analytic functions in $\mathbb{D}$.

(i) As $f\in\mathfrak{F}_1(0)$, then $p_2\in\mathcal{P}_n(0)$. We know that $|p_1(z)-1|<1$ holds if $\mathrm{Re}(1/p_1(z))>1/2$ and vice-versa. Assume $f(z)=zp_1(z)p_2(z)$. Now using the expressions of $p_1(z)$, $p_2(z)$ and by applying Theorem 2.3 and Lemma 2.10 we have

$$\left|\frac{zf'(z)}{f(z)}-1\right|=\left|\frac{zp_2'(z)}{p_2(z)}-\frac{zp_1'(z)}{p_1(z)}\right|\le\frac{(3+r^n)nr^n}{1-r^{2n}}\le 1-c_0.$$

The above inequality leads to $r^{2n}(n+1-c_0)+3nr^n-1+c_0\le 0$, provided $r\le\mathcal{R}_{\mathcal{S}_n^*(\varrho)}(\mathfrak{F}_1(0))$. The functions $\tilde{f}(z)=z(1+z^n)/(1-z^n)^2$ and $\tilde{g}(z)=z(1+z^n)/(1-z^n)$ at $z_0=\mathcal{R}_{\mathcal{S}_n^*(\varrho}(\mathfrak{F}_1(0))e^{i\pi/n}$ gives

$$\frac{z_0\tilde{f}'(z_0)}{\tilde{f}(z_0)}-1=\frac{(3+z_0^n)nz_0^n}{1-z_0^{2n}}=1-c_0.$$

Thus $\tilde{f}$ is the extremal function.

(ii) As $f \in \mathcal{R}_{\mathcal{S}_n^*(\varrho)}(\mathfrak{F}_1(1/2))$, then $1/p_1, p_2 \in \mathcal{P}_n(1/2)$. Proceeding as in (i), on applying Theorem 2.3 and Lemma 2.10 we get

$$\left|\frac{zf'(z)}{f(z)} - 1\right| \leq \frac{2nr^n}{1-r^n} \leq 1 - c_0.$$

This holds true whenever $r \leq \mathcal{R}_{\mathcal{S}_n^*(\varrho)}(\mathfrak{F}_1(1/2))$. For sharpness, consider $\tilde{f}(z) = z$ and $\tilde{g}(z) = z/(1-z^n)$, then at $z_0 = \mathcal{R}_{\mathcal{S}_n^*(\varrho)}(\mathfrak{F}_1(1/2))e^{i\pi/n}$, we get

$$\frac{z_0\tilde{f}'(z_0)}{\tilde{f}(z_0)} - 1 = \frac{2nz_0^n}{1-z_0^n} = 1 - c_0.$$

(iii) Let $f(z)/g(z) = p(z)$ be a function defined in $\mathbb{D}$. As $f \in \mathfrak{F}_2$, then $|1/p(z) - 1| < 1$ if and only if $\operatorname{Re} p(z) > 1/2$. Further as $g \in \mathcal{C}$, then due to Marx-Strohhäcker theorem, $g \in \mathcal{S}^*(1/2)$. Therefore due to Lemma 2.9,

$$\left|\frac{zg'(z)}{g(z)} - \frac{1}{1-r^{2n}}\right| \leq \frac{r^n}{1-r^{2n}}.$$

On logarithmically differentiating $f(z)$ and applying Theorem 2.3, we get

$$\begin{aligned}\left|\frac{zf'(z)}{f(z)} - \frac{1}{1-r^{2n}}\right| &= \left|\frac{zg'(z)}{g(z)} - \frac{zp'(z)}{p(z)} - \frac{1}{1-r^{2n}}\right| \\ &\leq \frac{nr^{2n} + (1+n)r^n}{1-r^{2n}} \leq \frac{1}{1-r^{2n}} - c_0.\end{aligned}$$

This leads us to the inequality $r^{2n}(n-c_0) + r^n(1+n) - 1 + c_0 \leq 0$, provided $r \leq \mathcal{R}_{\mathcal{S}_n^*(\varrho)}(\mathfrak{F}_2)$. The functions $\tilde{f}(z) = z(1+z^n)/(1-z^n)^{1/n}$ and $\tilde{g}(z) = z/(1-z^n)^{1/n}$ at $z_0 = \mathcal{R}_{\mathcal{S}_n^*(\varrho)}(\mathfrak{F}_2)e^{i\pi/n}$ gives $|z_0\tilde{f}'(z_0)/\tilde{f}(z_0)| = c_0$. Hence the result is sharp.

∎

Theorem 3.9. *Let $-1 \leq A \leq 1$ and $0 \leq r < 1$, then the sharp $\mathcal{S}_n^*(\varrho)-$radius for the class $\mathfrak{F}_3$ is given by*

$$\mathcal{R}_{\mathcal{S}_n^*(\varrho)}(\mathfrak{F}_3) = \begin{cases} \mathcal{R}_0, & r \leq \mathcal{R}_0, \\ \mathcal{R}_1, & r \geq \mathcal{R}_0, \end{cases}$$

where

$$\mathcal{R}_0 = \begin{cases} \left(\dfrac{1+A+4n+\sqrt{(1+A+4n)^2 - 4(1-c_0)(A+c_0)}}{2(A+c_0)}\right)^{1/n} & \text{if } -1 \leq A < -c_0, \\ \left(\dfrac{1-c_0}{1+4n-c_0}\right)^{1/n} & \text{if } A = -c_0, \\ \left(\dfrac{1+A+4n-\sqrt{(1+A+4n)^2 - 4(1-c_0)(A+c_0)}}{2(A+c_0)}\right)^{1/n} & \text{if } -c_0 < A \leq 1, \end{cases}$$

and

$$\mathcal{R}_1 = \left(\frac{\sqrt{(1+A+4n)^2 + 4(A+c_1)(c_1-1)} - (1+A+4n)}{2(A+c_1)}\right)^{1/n}.$$

Proof. Let $f \in \mathfrak{F}_3$, then $\operatorname{Re} f(z)/g(z) > 0$ and $\operatorname{Re}((1-z^n)^{(1+A)/n} g(z)/z) > 0$, where $g \in \mathcal{A}_n$. Define $g(z)/f(z) = p_1(z)$ and $(1-z^n)^{(1+A)/n} g(z)/z = p_2(z)$, where $p_1(z)$ and $p_2(z)$ are analytic

in $\mathbb{D}$. Since $A<1$, then for $|z|=r<1$, the inequality $(1+Ar^{2n})\geq 1-r^{2n}$, holds true. Further on logarithmically differentiating $zp_1(z)p_2(z)(1-z^n)^{-(1+A)/n}=f(z)$, we get

$$\frac{zf'(z)}{f(z)}=\frac{1+Az^n}{1-z^n}+\frac{zp_1'(z)}{p_1(z)}+\frac{zp_2'(z)}{p_2(z)}.$$

Due to Lemmas 2.9 - 2.10, for $|z|=r$, we infer

$$\left|\frac{zf'(z)}{f(z)}-\frac{1+Ar^{2n}}{1-r^{2n}}\right|\leq\frac{4nr^n}{1-r^{2n}}+\frac{(1+A)r^n}{1-r^{2n}}. \tag{3.6}$$

Assume $c=(1+Ar^{2n})/(1-r^{2n})$. Then $c\leq(c_0+c_1)/2$ leads to $r\leq\mathcal{R}$ and vice-versa, where $\mathcal{R}=((c_0-2)/(2A+c_0))^{1/2n}$. Algebraically, for each $n=1,2,3,\ldots$, it can be observed that, for the given range of A, we have $\mathcal{R}_0<\mathcal{R}_1<\mathcal{R}$. In particular, if $r\leq\mathcal{R}_0$, then we have $c\leq(c_0+c_1)/2$. Furthermore due to Theorem 2.3 the inequality (3.6) gives

$$\frac{4nr^n}{1-r^{2n}}+\frac{(1+A)r^n}{1-r^{2n}}\leq\frac{1+Ar^{2n}}{1-r^{2n}}-c_0,$$

whenever $r\leq\mathcal{R}_0$. Moreover if $c\geq(c_0+c_1)/2$, then $r\geq\mathcal{R}_0$. Infact, when $r\geq\mathcal{R}_0$, then we have $c\geq(c_0+c_1)/2$. Now (3.6) together with Theorem 2.3 yields

$$\frac{4nr^n}{1-r^{2n}}+\frac{(1+A)r^n}{1-r^{2n}}\leq c_1-\frac{1+Ar^{2n}}{1-r^{2n}},$$

whenever $r\leq\mathcal{R}_1$. Hence the following functions, mentioned in Remark 3.7

$$\tilde{f}(z)=\frac{z(1+z^n)^2}{(1-z^n)^{2+(1+A)/n}}\quad\text{and}\quad\tilde{g}(z)=\frac{z(1+z^n)}{(1-z^n)^{1+(1+A)/n}},$$

serve as the extremal function for both the cases. ∎

4. Certain estimates for hyperbolic cosine function

In this section certain sufficient conditions for the class $\mathcal{S}_\varrho^*$ are established. We begin with the following criterion.

Theorem 4.1. *Let $f\in\mathcal{A}$, then $f\in\mathcal{S}_\varrho^*$ if and only if*

$$\frac{1}{z}\left(f(z)*\frac{z-kz^2}{(1-z)^2}\right)\neq 0 \tag{4.1}$$

where $k=\cosh e^{it/2}/(\cosh e^{it/2}-1)$ for $t\in[-\pi,\pi]$. Moreover, $f\in\mathcal{S}_\varrho^*$ if and only if

$$1-\sum_{n=2}^{\infty}\frac{(n-\cosh e^{it/2})a_n}{\cosh e^{it/2}-1}z^{n-1}\neq 0. \tag{4.2}$$

Proof. Since $f\in\mathcal{S}_\varrho^*$, then $zf'(z)/f(z)=\cosh\sqrt{w(z)}$, where $w(z)$ is a Schwarz function with $w(0)=0$. Equivalently for $w(z)=e^{it}$, $-\pi\leq t\leq\pi$, we have

$$\frac{zf'(z)}{f(z)}\neq\cosh e^{it/2}\Leftrightarrow zf'(z)-(\cosh e^{it/2})f(z)\neq 0\quad\text{for}\quad t\in[-\pi,\pi],$$

Eventually it leads to $zf'(z)-k(zf'(z)-f(z))\neq 0$. Thus through simple computations (4.1) can be established. The condition in (4.2) can be deduced using (4.1). ∎

Corollary 4.2. *Let $f \in \mathcal{A}$ satisfy the following*

$$\sum_{n=2}^{\infty} \left| \frac{n - \cosh e^{it/2}}{\cosh e^{it/2} - 1} \right| |a_n| < 1, \tag{4.3}$$

then $f \in \mathcal{S}^*_\varrho$.

Proof. Consider the following inequality with $k = \cosh e^{it/2}/(\cosh e^{it/2} - 1)$,

$$\left| 1 - \sum_{n=2}^{\infty} (n(k-1) - k) a_n z^{n-1} \right| \geq 1 - \sum_{n=2}^{\infty} |n(k-1) - k||a_n|.$$

Thus from (4.3) we establish

$$\left| 1 - \sum_{n=2}^{\infty} (n(k-1) - k) a_n z^{n-1} \right| > 0,$$

Hence due to Theorem 4.1 we conclude that $f \in \mathcal{S}^*_\varrho$. ∎

Theorem 4.3. *Let $f \in \mathcal{S}^*_\varrho$ then the following inequality holds*

$$c_1^2 - 1 \geq \sum_{k=2}^{\infty} (k^2 - {c_1}^2)|a_k|^2.$$

Proof. Since $f \in \mathcal{S}^*_\varrho$, then $zf'(z) = \cosh(\sqrt{w(z)})f(z)$, for a Schwarz function $w(z)$ with $w(0) = 0$. For $0 \leq |z| = r < 1$, we get the following

$$\begin{aligned}
2\pi \sum_{k=1}^{\infty} k^2 |a_k|^2 r^{2k} &= \int_0^{2\pi} \left| re^{i\theta} f'(re^{i\theta}) \right|^2 d\theta \\
&= \int_0^{2\pi} \left| \cosh\left(\sqrt{w(re^{i\theta})} \right) f(re^{i\theta}) \right|^2 d\theta \\
&\leq \int_0^{2\pi} \cosh^2\left(\sqrt{|w(re^{i\theta})|} \right) |f(re^{i\theta})|^2 d\theta \\
&\leq \int_0^{2\pi} (\cosh^2 r) |f(re^{i\theta})|^2 d\theta \\
&= 2\pi (\cosh^2 r) \sum_{k=1}^{\infty} |a_k|^2 r^{2k}.
\end{aligned} \tag{4.4}$$

Thus when r tends to 1^-, we at once obtain the required inequality. ∎

Example 2. *Let $f \in \mathcal{A}$, then following functions are members of $\mathcal{S}^*_\varrho$.*

(i) $f(z) = z + a_n z^n \in \mathcal{S}^*_\varrho$, *provided* $|a_n| \leq (1 - c_0)/(n - c_0)$, $n \in \mathbb{N} - \{1\}$.
(ii) $f(z) = z/(1 - Az)^2 \in \mathcal{S}^*_\varrho$, *provided* $|A| \leq (c_1 - 1)/(c_1 + 1)$.
(iii) $f(z) = z/(1 - Az) \in \mathcal{S}^*_\varrho$, *provided* $|A| \leq (c_1 - 1)/c_1$.
(iv) $f(z) = ze^{Az} \in \mathcal{S}^*_\varrho$, *provided* $|A| \leq 1 - c_0$.

Proof. (i) To establish this result, we use Theorem 2.3. We require that $zf'(z)/f(z) = (1 + na_n z^{n-1})/(1 + a_n z^{n-1})$ must lie in the disc $\{u : |u - c| < r_c\} \subset \varrho(\mathbb{D})$, centered at c, where r_c is defined in (2.2). We know that the function $f(z) = z + a_n z^n$ is univalent if and only if

$|a_n| \leq 1/n$. Therefore $c = (1-n|a_n|^2)/(1-|a_n|^2) \leq 1$. Now if $u = (1+na_nz^{n-1})/(1+a_nz^{n-1})$ and $r_c = (1-n|a_n|^2)/(1-|a_n|^2) - c_0$, then due to Theorem 2.3,

$$\frac{(n-1)|a_n|}{1-|a_n|^2} \leq \frac{1-n|a_n|^2}{1-|a_n|^2} - c_0.$$

Equivalently we have $|a_n| \leq (1-c_0)/(n-c_0)$.

(ii) Let $zf'(z)/f(z) = (1+Az)/(1-Az)$. We know that the extremities of diameter of the disc

$$\left|\frac{1+Az}{1-Az} - \frac{1+|A|^2}{1-|A|^2}\right| \leq \frac{2|A|}{1-|A|^2}$$

are $(1-|A|)/(1+|A|)$ and $(1+|A|)/(1-|A|)$. Therefore due to Theorem 2.3, $u = (1+Az)/(1-Az)$ lies in the region $\{u : |u-c| < r_c\} \subset \varrho(\mathbb{D})$, provided $(1+|A|)/(1-|A|) \leq c_1$. Equivalently $|A| \leq (c_1-1)/(c_1+1)$.

(iii) As $zf'(z)/f(z) = 1/(1-Az)$ maps $\mathbb{D}$ onto the disc

$$\left|\frac{1}{1-Az} - \frac{1}{1-|A|^2}\right| \leq \frac{|A|}{1-|A|^2}.$$

Now proceeding as in part (ii), we get the desired bound on $|A|$.

(iv) We know that $zf'(z)/f(z) = 1+Az$ maps $\mathbb{D}$ onto the disc $|u-1| < |A|$. Then due to Theorem 2.3, $u = 1+Az$ lies in $\{u : |u-c| < r_c\} \subset \varrho(\mathbb{D})$, provided $|A| \leq 1-c_0$.

∎

5. Subordination Results

Let $h(z)$ be a univalent function in $\mathbb{D}$ and $\psi : \mathbb{C}^3 \times \mathbb{D} \to \mathbb{C}$ be an analytic function. Suppose $p(z)$ is analytic in $\mathbb{D}$ and satisfies the differential subordination

$$\psi(p(z), zp'(z), z^2p''(z); z) \prec h(z), \tag{5.1}$$

then $p(z)$ is called a solution of the differential subordination (5.1). The univalent function $q(z)$ is called a dominant of the solutions of (5.1). Amongst all the dominants $\tilde{q}(z)$ which satisfy $\tilde{q} \prec q$ for all dominants $q(z)$ of (5.1), is called the best dominant of (5.1). Observe that the best dominant $\tilde{q}(z)$ is unique upto rotations. Miller and Mocanu in [20] have extensively discussed the theory of differential subordination. In the recent past several authors have established conditions on η, so that $1+\beta zp'(z)/p^n(z) \prec \phi(z)$ $(n = 0, 1, 2)$ implies $p(z) \prec \psi(z)$, for different choices of $\phi(z)$ and $\psi(z)$, namely e^z, $\sqrt{1+z}$ and $(1+Az)/(1+Bz)$ $(-1 \leq B < A \leq 1)$, see [2, 8, 18, 27, 28, 30]. In this section we have established certain differential subordination results related to functions $e^z, (1+Az)/(1+Bz), z+\sqrt{1+z}$ and $(1+sz)^2$, where $0 < s \leq 1/\sqrt{2}$. Furthermore results related to Briot-Bouquet type differential subordination have also been discussed. In sequel we require following lemmas and definition.

Definition 5.1. Let $\mathcal{Q}$ be the set of functions q that are analytic and injective on $\overline{\mathbb{D}} \backslash E(q)$, where

$$E(q) = \left\{\zeta \in \partial\mathbb{D} : \lim_{z\to\zeta} g(z) = \infty\right\},$$

such that $q'(\zeta) \neq 0$ for $\zeta \in \partial\mathbb{D} \backslash E(q)$.

Lemma 5.2. [20] *Let q be analytic in $\mathbb{D}$ and let φ be analytic in domain $\mathcal{D}$ containing $q(\mathbb{D})$ with $\varphi(w) \neq 0$ when $w \in q(\mathbb{D})$. Set $\mathcal{Q}(z) := zq'(z)\varphi(q(z))$ and $h(z) := \nu(q(z)) + \mathcal{Q}(z)$. Suppose (i.) either h is convex, or $\mathcal{Q}$ is starlike univalent in $\mathbb{D}$ and (ii.)* $\mathrm{Re}(zh'(z)/\mathcal{Q}(z)) > 0$ *for $z \in \mathbb{D}$. If p is analytic in $\mathbb{D}$, with $p(0) = q(0)$, $p(\mathbb{D}) \subset \mathcal{D}$ and*

$$\nu(p(z)) + zp'(z)\varphi(p(z)) \prec \nu(q(z)) + zq'(z)\varphi(q(z)),$$

then $p(z) \prec q(z)$, and q is the best dominant.

Lemma 5.3. [24, Lemma 1.3, p.28] *Let w be a meromorphic function in $\mathbb{D}$, $w(0) = 0$. If for some $z_0 \in \mathbb{D}$, $\max_{|z|\leq|z_0|} |w(z)| = |w(z_0)|$, then it follows that $z_0 w'(z_0)/w(z_0) \geq 1$.*

Theorem 5.4. *Suppose η satisfies $|\eta| \geq 2|A-B|/(1-|B|)\tanh 1$, where $(A, B) \in \mathbb{C} \times \mathbb{C}$, with $A \neq B$ and $|B| < 1$. Let $p(z)$ be an analytic function in $\mathbb{D}$, with $p(0) = 1$, satisfying the subordination $1 + \eta(zp'(z))/p(z) \prec (1+Az)/(1+Bz)$, then $p(z) \prec \cosh\sqrt{z} =: \varrho(z)$.*

Proof. Let $T(z) = z\varrho'(z)/\varrho(z) = (\sqrt{z}\tanh\sqrt{z})/2$, then clearly it is a starlike function in $\mathbb{D}$ and $\mathrm{Re}(\sqrt{z}\,\mathrm{csch}\,2\sqrt{z}) > 0$. Let $v(u) = 1$ and $\psi(u) = \eta/u$. Note that $\mathrm{Re}\, z\mathfrak{h}'(z)/T(z) - 1 = \mathrm{Re}\,\sqrt{z}\,\mathrm{csch}\,2\sqrt{z} > 0$, where $\mathfrak{h}(z) = v(\varrho(z)) + T(z)$. Assume $\Psi(z) = (1+Az)/(1+Bz)$. Then $\Psi^{-1}(u) = (u-1)(A-Bu)$. Let

$$M(z) = 1 + \frac{\eta}{2}\sqrt{z}\tanh(\sqrt{z}).$$

We show that $\Psi(z) \prec M(z)$. For this purpose we shall prove that $|\Psi^{-1}(M(e^{it}))| \geq 1$ for each $-\pi \leq t \leq \pi$. On the boundary of $\mathbb{D}$, we have

$$|\Psi^{-1}(M(e^{it}))| \geq \frac{|\eta||\tanh(e^{it/2})|}{2|A-B| + |\eta B\tanh(e^{it/2})|}.$$

As $|\tanh(e^{it/2})|$ attains its minimum at $t = 0$, we get the required lower bound on $|\eta|$. Thus by applying Lemma 5.2 the result follows. ∎

Upon substituting $p(z) = zf'(z)/f(z)$, we deduce the following corollary.

Corollary 5.5. *Assume A and B be as defined in Theorem 5.4 and let $f \in \mathcal{A}$ such that*

$$1 + \eta\left(1 + \frac{zf''(z)}{f'(z)} - \frac{zf'(z)}{f(z)}\right) \prec \frac{1+Az}{1+Bz}$$

then $f \in \mathcal{S}^*_\varrho$.

Theorem 5.6. *Let $p(z)$ be an analytic function in $\mathbb{D}$ such that $1 + \eta zp'(z) \prec \cosh\sqrt{z}$ with*

$$\kappa = \frac{Chi(1) + Ci(1) - 2\xi}{Chi(1) - Ci(1)},$$

where

$$Chi(1) = \xi + \int_0^1 (\cosh\sqrt{t} - 1)/t \ dt, \quad Ci(1) = \xi + \int_0^1 (\cos\sqrt{t} - 1)/t \ dt.$$

then

(i) $p(z) \prec e^z$, whenever $(e-1)\eta \geq 2e(\xi - Ci(1))$.
(ii) $p(z) \prec z + \sqrt{1+z^2}$, whenever $(\sqrt{2}-1)\eta \geq \sqrt{2}(\xi - Ci(1))$.
(iii) $p(z) \prec (1+Az)/(1+Bz)$, where $-1 < B < A \leq 1$, whenever $\eta \geq \eta_0$ where

$$\eta_0 = \begin{cases} \dfrac{2}{A-B}(1-B)(\xi - Ci(1)) & \text{if } -1 < B \leq -\kappa, \\ \dfrac{-2}{A-B}(1+B)(\xi - Chi(1)) & \text{if } -\kappa < B < 1. \end{cases} \tag{5.2}$$

(iv) $p(z) \prec (1+sz)^2$ *whenever*

$$\eta_0 = \begin{cases} \dfrac{2}{s(s+2)}(Chi(1)-\xi) & \text{if } 0 < s \leq 2\kappa, \\ \dfrac{2}{s(s-2)}(Ci(1)-\xi) & \text{if } 2\kappa < s \leq 1/\sqrt{2}. \end{cases} \tag{5.3}$$

Proof. Let q_η be an analytic function, $q_\eta : \overline{\mathbb{D}} \to \mathbb{C}$ defined by

$$q_\eta(z) = 1 + \frac{1}{\eta}\left(2Chi\sqrt{z} - \log z - 2\xi\right).$$

Then $q_\eta(z)$ is a solution of the differential equation $1 + \eta z p'(z) = \cosh\sqrt{z}$. Let $v(u) = 1$ and $\psi(u) = \eta$. Define an analytic function $T : \overline{\mathbb{D}} \to \mathbb{C}$ by

$$T(z) = z q_\eta'(z)\psi(q_\eta(z)) = \cosh\sqrt{z} - 1.$$

Note that $T(z)$ is a starlike function in $\mathbb{D}$. Observe that $\mathfrak{h}(z) = v(q_\eta(z)) + T(z)$ satisfies

$$\operatorname{Re}\left(\frac{z\mathfrak{h}'(z)}{T(z)}\right) = \operatorname{Re}\left(\frac{1}{2}\sqrt{z}\coth\left(\frac{\sqrt{z}}{2}\right)\right) > 0.$$

Thus by Lemma 5.2 we deduce that $p(z) \prec q_\eta(z)$. In each of the above parts we shall show that $p(z) \prec \mathcal{G}(z)$, for various choices of $\mathcal{G}(z) : e^z, z + \sqrt{1+z^2}, (1+Az)/(1+Bz)$ or $(1+sz)^2$. Further if $q_\eta(z) \prec \mathcal{G}(z)$, then

$$\mathcal{G}(-1) \leq q_\eta(-1) < q_\eta(1) \leq \mathcal{G}(1) \tag{5.4}$$

holds true. In each of the following parts we shall obtain a condition on η for which (5.4) holds.

(i) Let $\mathcal{G}(z) = e^z$, then the inequalities (5.4), yields the following

$$e^{-1} \leq 1 + \frac{2}{\eta}(Ci(1)-\xi) < 1 + \frac{2}{\eta}(Chi(1)-\xi) \leq e, \tag{5.5}$$

Inequality (5.5) holds true, provided $\eta \geq \eta_0$, where $\eta_0 = 2e(\xi - Ci(1))/(e-1)$. Thus, for each $\eta \geq \eta_0$ the result follows.

(ii) For $\mathcal{G}(z) = z + \sqrt{1+z^2}$, the expression in (5.4), gives

$$\sqrt{2} - 1 \leq 1 + \frac{2}{\eta}(Ci(1)-\xi) < 1 + \frac{2}{\eta}(Chi(1)-\xi) \leq \sqrt{2} + 1, \tag{5.6}$$

which further gives that (5.6) holds for each $\eta \geq \eta_0$, where $\eta_0 = \sqrt{2}(\xi - Ci(1))/\left(\sqrt{2}-1\right)$. Therefore for each $\eta \geq \eta_0$, $q_\eta(z) \prec \mathcal{G}(z)$. This leads us to the conclusion that $p(z) \prec \mathcal{G}(z)$.

(iii) Let $\mathcal{G}(z) = (1+Az)/(1+Bz)$. Then for (5.4) to hold true, proceeding as in the previous cases, we must have $\eta \geq \eta_0$ where η_0 is as given in (5.2).

(iv) For $\mathcal{G}(z) = (1+sz)^2$, if η in (5.4) satisfies the conditions given in (5.3), then it yields the required subordination implication.

∎

Theorem 5.7. *Let* $-1 < B < A \leq 1$ *and* $\mu = (c_1 - 1)/(1 - c_0)$. *Suppose* $p(z)$ *is an analytic function in* $\mathbb{D}$ *such that* $1 + \eta z p'(z) \prec (1+Az)/(1+Bz)$, *where* η *satisfies the following*

$$\begin{cases} \eta B(c_1 - 1) \leq (A-B)\log(1+B) \text{ if } -1 < B \leq B_0, \\ \eta B(c_0 - 1) \geq (A-B)\log(1-B) \text{ if } B_0 < B < 0, \\ \eta B(c_0 - 1) \leq (A-B)\log(1-B) \text{ if } 0 < B < 1, \end{cases} \tag{5.7}$$

where B_0 is the root of the equation $(1+B)(1-B)^{\mu}=1$. Moreover, for $B=0$, if η satisfies $2\eta \geq A\csc^2(1/2)$, then $p(z) \prec \cosh\sqrt{z}$.

Proof. Let $q_\eta(z)$ be an analytic function, $q_\eta : \overline{\mathbb{D}} \to \mathbb{C}$ defined as follows:

$$q_\eta(z) = 1 + \frac{A-B}{\eta B}\log(1+Bz),$$

be a solution of the differential equation $1+\eta z p'(z) = (1+Az)/(1+Bz)$. Let $v(u)=1$ and $\psi(u)=\eta$. Define an analytic function $T : \overline{\mathbb{D}} \to \mathbb{C}$ by $T(z) = zq'_\eta(z)\psi(q_\eta(z))$. Clearly for the given choice of A and B, $T(z) = (A-B)z/(1+Bz)$ is a starlike function in $\mathbb{D}$. Note that the function $\mathfrak{h}(z) = v(q_\eta(z)) + T(z)$ gives $z\mathfrak{h}'(z)/T(z) = 1/(1+Bz)$. Thus $\operatorname{Re} z\mathfrak{h}'(z)/T(z) > 0$. Further by Lemma 5.2 we deduce that $p(z) \prec q_\eta(z)$. We need to prove that $p(z) \prec \cosh\sqrt{z} =: \varrho(z)$. Graphically observe that due to the images of $q_\eta(z)$ and $\varrho(z)$, the implication $q_\eta(z) \prec \varrho(z) \Rightarrow \varrho(-1) \leq q_\eta(-1) < q_\eta(1) \leq \varrho(1)$. On taking suitable values of η, this inequality becomes a necessary as well as sufficient for $q_\eta(z) \prec \varrho(z)$. For $B \neq 0$, we have

$$c_0 \leq 1 + \frac{A-B}{\eta B}\log(1-B) < 1 + \frac{A-B}{\eta B}\log(1+B) \leq c_1.$$

On solving the above inequalities the conditions on η given in (5.7) can be established. Further for $B=0$, we get $m_\eta(z) = 1+Az/\eta$ such that inequality $c_0 \leq m_\eta(-1) < m_\eta(1) \leq c_1$ holds. This inequality holds if and only if $A \leq 2\eta\sin^2(1/2)$. Hence $p(z) \prec \cosh\sqrt{z}$. ∎

Theorem 5.8. *Suppose $(\eta,\gamma) \in \mathbb{R}\times\mathbb{R} - \{(\eta,-\eta)\}$ satisfy the following inequalities:*

$$0 \leq \gamma + \eta c_1 \leq \frac{\sinh 1}{2\left(1+\sqrt{2}-c_1\right)} \text{ and } 0 \geq \gamma + \eta c_0 \geq -\frac{\sin 1}{2\left(1+\sqrt{2}-c_0\right)}. \tag{5.8}$$

Let $p(z)$ be an analytic function, such that $p(0)=1$ and satisfies

$$p(z) + \frac{zp'(z)}{\eta p(z)+\gamma} \prec z + \sqrt{1+z^2} \tag{5.9}$$

then

$$p(z) \prec \cosh\sqrt{z}.$$

Proof. We define analytic functions $M(z)$ and $w(z)$, as $M : \mathbb{D} \to \mathbb{C}$ and $w : \mathbb{D} \to \mathbb{C}$ defined by

$$M(z) = p(z) + \frac{zp'(z)}{\eta p(z)+\gamma} \quad \text{and} \quad p(z) = \cosh\sqrt{w(z)} \tag{5.10}$$

where $w(z)$ is defined on the principle branch of logarithmic and square root function. Clearly $w(z)$ is analytic with $w(0)=0$. We need to prove that $|w(z)|<1$. Let $z_0 \in \mathbb{D}$ such that $\max_{|z|\leq|z_0|}|w(z)| = |w(z_0)| = 1$. Further using Lemma 5.3, $z_0w'(z_0) = kw(z_0)$ where $k \geq 1$. Assume $w(z) = e^{it}$, $-\pi \leq t \leq \pi$. Consider the expression

$$\begin{aligned}
|M(z_0)-1|^2 &= \left|\cosh\sqrt{w(z_0)} - 1 + \frac{z_0w'(z_0)\sinh\sqrt{w(z_0)}}{2\sqrt{w(z_0)}(\eta\cosh\sqrt{w(z_0)}+\gamma)}\right|^2 \\
&= \left|\cosh e^{it/2} - 1 + \frac{ke^{it/2}\sinh e^{it/2}}{2(\eta\cosh e^{it/2}+\gamma)}\right|^2 \\
&= \frac{A_N(\tau)}{A_D(\tau)} \quad (\tau = t/2, -\pi \leq \tau \leq \pi),
\end{aligned}$$

where

$$\begin{aligned}A_N(\tau) = {} & (4\eta^2\cos^3(\sin\tau)\cosh^3(\cos\tau) - \eta\cosh^2(\cos\tau)(k\sin\tau\sin(2\sin\tau)\\ & + 4(\eta-2\gamma)\cos^2(\sin\tau)) + \eta\sinh^2(\cos\tau)(k\sin\tau\sin(2\sin\tau) - 4\eta\sin^2(\sin\tau))\\ & + 2\gamma k\cos\tau\cos(\sin\tau)\sinh(\cos\tau) + \eta k\cos\tau\sinh(2\cos\tau) - 4\gamma^2\\ & + \cosh(\cos\tau)(4\eta^2\sin^2(\sin\tau)\cos(\sin\tau)\sinh^2(\cos\tau) - 2\gamma k\sin\tau\sin(\sin\tau)\\ & + 4\gamma(\gamma-2\eta)\cos(\sin\tau)))^2 + (3\eta^2\sin(\sin\tau)\sinh(\cos\tau)\cosh^2(\cos\tau)\\ & + 2\gamma k\sin(\sin\tau)\cos\tau\cosh(\cos\tau) + \eta\sinh(2\cos\tau)(k\sin\tau + 2\gamma\sin(2\sin\tau))\\ & + \sinh(\cos\tau)(2\gamma k\sin\tau\cos(\sin\tau) - 2(\eta^2-2\gamma^2)\sin(\sin\tau) + \eta^2\sin(3\sin\tau))\\ & + \eta k\sin(2\sin\tau)\cos\tau + \eta^2\sin(\sin\tau)\sinh^3(\cos\tau))^2\end{aligned}$$

and

$$A_D(\tau) = 16((\gamma + \eta\cos(\sin\tau)\cosh(\cos\tau))^2 + \eta^2\sin^2(\sin\tau)\sinh^2(\cos\tau))^2$$

Define $F(\tau) = A_N(\tau) - 2A_D(\tau)$ on $[-\pi/2, \pi/2]$, then clearly $F(\tau)$ is an even function. Therefore we verify that $F(\tau)$ is non-negative for each $\tau \in [0, \pi/2]$ and minimum of $F(\tau)$ is attained at either $\tau = 0$ or $\tau = \pi/2$. Observe that

$$F(0) = (k(\eta\sinh 2 + 2\gamma\sinh 1) - 4(1-c_1)(\gamma+\eta c_1)^2)^2 - 32(\gamma+\eta c_1)^4$$

and

$$F\left(\frac{\pi}{2}\right) = (k(\eta\sin 2 + 2\gamma\sin 1) - 4(c_0-1)(\gamma+\eta c_0)^2)^2 - 32(\gamma+\eta c_0)^4.$$

For the sake of simplicity, we assume $\eta_0 = \sinh 1/(2c_1\left(1+\sqrt{2}-c_1\right)) - \gamma/c_1$, $\eta_1 = -\gamma/c_1$, $\eta_2 = -\sin 1/(2c_0\left(1+\sqrt{2}-c_0\right)) - \gamma/c_0$ and $\eta_3 = -\gamma/c_0$. As per the conditions in (5.8) it can be deduced that $F(0) \geq (4(c_1-1)(\gamma+\eta c_1)^2 + (\eta\sinh 2 + 2\gamma\sinh 1)^2)^2 - 32(\gamma+\eta c_1)^4 =: X(\eta)$. From inequalities in (5.8) we deduce that $X(\eta) \geq 0$, provided $\eta_0 \leq \eta \leq \eta_1$. Thus $F(0)$ is a non-negative quantity. Further as $\eta \leq \eta_3$, so we have $F(\pi/2) \geq ((\eta\sin 2 + 2\gamma\sin 1) - 4(c_0-1)(\gamma+\eta c_0^2)^2 - 32(\gamma+\eta c_0)^4 =: Y(\eta)$. Note that $\eta_2 \leq \eta \leq \eta_3$, leads to $(\eta\sin 2 + 2\gamma\sin 1) - 4(c_0-1)(\gamma+\eta c_0)^2 \geq 4\sqrt{2}\left(\gamma+\eta c_0\right)^2$. Thus $Y(\eta) \geq 0$ and eventually $F(\pi/2)$ is a non-negative quantity. Hence we conclude that $|M(z_0)-1|^2 \geq 2$. This contradicts the assumption in (5.9). Thus $p(z) \prec \cosh\sqrt{z}$. ∎

In Corollary 5.9, we give an illustration for Theorem 5.8. It can be verified that for $\vartheta \in \mathcal{A}$ the Briot-Bouquet differential equation [20] is related to an integral operator $\mathcal{L}$ defined as follows:

$$\mathcal{L}(z) := \frac{\eta+\gamma}{z^\gamma}\int_0^z \vartheta^\eta(t)t^{\gamma-1}dt.$$

Corollary 5.9. *Let $(\eta,\gamma) \in \mathbb{R}\times\mathbb{R} - \{(\eta,-\eta)\}$ satisfy the conditions given in (5.8). If $\vartheta \in \mathcal{S}_q^*$, then $\mathcal{L} \in \mathcal{S}_\varrho^*$.*

Proof. Suppose $p(z) = z\mathcal{L}'(z)/\mathcal{L}(z)$ be analytic in $\mathbb{D}$, then through this substitution we obtain the Briot-Bouquet type differential equation of the form

$$p(z) + \frac{zp'(z)}{\eta p(z)+\gamma} = \frac{z\vartheta'(z)}{\vartheta(z)}.$$

Thus by applying Theorem 5.8 the result holds true. ∎

Theorem 5.10. *Suppose $(\eta,\gamma) \in \mathbb{R}\times\mathbb{R} - \{(\eta,-\eta)\}$ satisfy the following conditions:*

$$0 < \gamma + \eta c_1 \leq \frac{\sinh 1}{2(e-c_1)} \text{ and } -\frac{\sin 1}{2(e-c_0)} \leq \gamma + \eta c_0 < 0, \tag{5.11}$$

Let $p(z)$ be an analytic function, such that $p(0) = 1$ and satisfies

$$p(z) + \frac{zp'(z)}{\eta p(z) + \gamma} \prec e^z \tag{5.12}$$

then

$$p(z) \prec \cosh \sqrt{z}.$$

Proof. Suppose $w(z)$ and $M(z)$ be analytic functions defined in (5.10). Then proceeding as in Theorem 5.8 we have $z_0 w'(z_0) = kw(z_0)$, where $k \geq 1$ and let $w(z) = e^{it}$, $-\pi \leq t \leq \pi$. For the purpose of proving this result we need to show that

$$|\log P(z_0)|^2 \geq 1. \tag{5.13}$$

Consider

$$\begin{aligned} A_N(\tau) = \kappa_\tau^{-1} &(\cosh(\cos\tau)(4\eta^2 \sin^2(\sin\tau)\cos(\sin\tau)\sinh^2(\cos\tau) + 4\gamma^2\cos(\sin\tau)) \\ &- 2\gamma k \sin(\sin\tau)\sin\tau + \eta\cosh^2(\cos\tau)(8\gamma\cos^2(\sin\tau) - k\sin(2\sin\tau)\sin\tau) \\ &+ k(\eta\cos\tau\sinh(2\cos\tau) + \eta\sin(2\sin\tau)\sin\tau\sinh^2(\cos\tau) \\ &+ 2\gamma\cos\tau\cos(\sin\tau)\sinh(\cos\tau)) + 4\eta^2\cos^3(\sin\tau)\cosh^3(\cos\tau)) \end{aligned}$$

and

$$\begin{aligned} A_D(\tau) = \kappa_\tau^{-1} &(\eta\sinh(2\cos\tau)(k\sin\tau + 2\gamma\sin(2\sin\tau)) + \eta k\sin(2\sin\tau)\cos\tau \\ &+ 2\gamma k\sin(\sin\tau)\cos\tau\cosh(\cos\tau) + 3\eta^2\sin(\sin\tau)\sinh(\cos\tau)\cosh^2(\cos\tau) \\ &+ \eta^2\sin(\sin\tau)\sinh^3(\cos\tau) + \sinh(\cos\tau)(2\gamma k\sin\tau\cos(\sin\tau) \\ &- 2(\eta^2 - 2\gamma^2)\sin(\sin\tau) + \eta^2\sin(3\sin\tau))), \end{aligned}$$

where $\kappa_\tau := 4((\gamma + \eta\cos(\sin\tau)\cosh(\cos\tau))^2 + \eta^2\sin^2(\sin\tau)\sinh^2(\cos\tau))$ and $\tau = t/2$. In order to show (5.13) for $-\pi/2 \leq \tau \leq \pi/2$, we equivalently prove that

$$F(\tau) := 4|\log p(z_0)|^2 - 4 = \log^2(A_N^2(\tau) + A_D^2(\tau)) + 4\left(\tan^{-1}\frac{A_N(\tau)}{A_D(\tau)}\right)^2 - 4 \geq 0.$$

Clearly $F(\tau)$ is an even function, so we shall restrict our findings to $[0, \pi/2]$. We show that $F(\tau) \geq 0$ on $[0, \pi/2]$. Observe that

$$F(0) = \left(\log\left(c_1 + \frac{k(\eta\sinh 2 + 2\gamma\sinh 1)}{4(\gamma + \eta c_1)^2}\right)^2\right)^2 - 4$$

and

$$F\left(\frac{\pi}{2}\right) = \left(\log\left(c_0 - \frac{k(\eta\sin 2 + 2\gamma\sin 1)}{4(\gamma + \eta c_0)^2}\right)^2\right)^2 - 4$$

Since $\log x$ is a monotonically increasing function, so it is enough to obtain the minimum of $X(k) := c_1 + k(\eta\sinh 2 + 2\gamma\sinh 1)/4(\gamma + \eta c_1)^2$. Observe that $X'(k) = (2\gamma\sinh 1 + \eta\sinh 2)/4(\gamma + \eta c_1)^2 > 0$, provided $\eta > -\gamma/c_1$. Due to the conditions in (5.11), $X(k) \geq X(1) = c_1 + (2\gamma\sinh 1 + \eta\sinh 2)/4(\gamma + \eta c_1)^2 \geq e$. Therefore we deduce that $(\log(X^2(k)))^2 \geq (\log e^2)^2 = 4$. Thus we have $F(0) \geq 0$. Similarly as $-\sin 1/(2c_0(e - c_0)) \leq \eta + \gamma/c_0 < 0$, we deduce that $Y(k) \geq Y(1) = c_0 - (\eta\sin 2 + 2\gamma\sin 1)/(4(\gamma + \eta c_0)^2) \geq e$. Thus we obtain $(\log(Y^2(k)))^2 \geq (\log e^2)^2 = 4$, that leads to $F(\pi/2) \geq 0$. Hence we get a contradiction to (5.12). Hence $p(z) \prec \cosh\sqrt{z}$. ∎

In Example 3, we illustrate a particular case of the above theorem. Best dominant for the differential subordination implication is obtained, that supports Theorem 5.10, for suitable values of η and γ.

Example 3. *Substitute $\eta = 1/2$, in Theorem 5.10 we get that $-c_0(1+\sin 1/c_0(e-c_0)) \leq 2\gamma \leq -c_1(1-\sinh 1/c_1(e-c_1))$. Upon taking $n=a=1, \eta=1/2$ and $h(z)=e^z$ in* [20, Theorem 3.2d, p.86] *we obtain $\mathrm{Re}(\eta a+\gamma) = \mathrm{Re}(\gamma+1/2) > 0$. Further suppose $\gamma + e^z/2 = \eta h(z) + \gamma \prec R_{\eta a+\gamma,n}(z) = R_{\gamma+1/2,1}(z)$, where $R_{\gamma+1/2,1}(z) = (\gamma+1/2)(1+z)/(1-z) + 2z/(1-z^2)$ is an open door function, that is univalent in $\mathbb{D}$. Therefore by an application of* [20, Theorem 3.2d, p.86], *we deduce that*

$$q_\gamma(z) = \left(\frac{1}{2}\int_0^1 t^{\gamma-1}\left(e^{Chi(tz)-Chi(z)+Shi(tz)-Shi(z)}\right)^{1/2} dt\right)^{-1} - 2\gamma,$$

where $Chi(z) := \xi + \log z + \displaystyle\int_0^z (\cosh t - 1)/t \; dt$ and $Shi(z) = \displaystyle\int_0^z \sinh t/t \; dt$, is a solution of the differential equation $q(z) + (zq'(z))/(\eta q(z)+\gamma) = e^z$, then $q_\gamma(z) \prec \cosh\sqrt{z}$. Moreover, as e^z is a convex function, then due to [20, Theorem 3.2j, p.97], *$q_\gamma(z)$ is univalent. For $n=1$, if $p \in \mathcal{H}(1,1)$ ($\mathcal{H}(a,n) := \{p \in \mathcal{H} : p(z) = a + \sum_{k=n}^{\infty} a_k z^k, a \in \mathbb{C}\}$) satisfies* (5.12), *then $p \prec q_\gamma$ and $q_\gamma(z)$ is the best (1,1)-dominant.*

Corollary 5.11. *If $\vartheta \in \mathcal{S}_e^*$ and $(\eta,\gamma) \in \mathbb{R}\times\mathbb{R} - \{(\eta,-\eta)\}$ satisfy the conditions given in* (5.11), *then $\mathcal{L} \in \mathcal{S}_\varrho^*$.*

Theorem 5.12. *Suppose $0 < s \leq 1/\sqrt{2}$ and $(\eta,\gamma) \in \mathbb{R}\times\mathbb{R} - \{(\eta,-\eta)\}$ satisfy the following inequalities:*

$$\eta_3 \leq \eta - \eta_2 \leq 0 \text{ and } \begin{cases} \eta \geq \eta_0 & \text{if } 0 < s \leq -1+\sqrt{c_1}, \\ 0 \leq \eta - \eta_0 \leq \eta_1 & \text{if } -1+\sqrt{c_1} < s \leq 1/\sqrt{2}, \end{cases} \tag{5.14}$$

where

$$\eta_0 = -\frac{\gamma}{c_1}, \quad \eta_1 = \frac{\sinh 1}{2c_1((1+s)^2 - c_1)},$$

$$\eta_2 = -\frac{\gamma}{c_0} \quad \text{and} \quad \eta_3 = -\frac{\sin 1}{2c_0((1+s)^2 - c_0)}.$$

If $p(z)$ is an analytic function such that $p(0)=1$ and satisfies

$$p(z) + \frac{zp'(z)}{\eta p(z) + \gamma} \prec (1+sz)^2 \tag{5.15}$$

then

$$p(z) \prec \cosh\sqrt{z}.$$

Proof. Define two analytic functions $M(z)$ and $w(z)$ in $\mathbb{D}$, as in (5.10). Consider the expression

$$\begin{aligned} |M(z_0) - 1|^2 &= \left|\cosh e^{it/2} - 1 + \frac{ke^{it/2}\sinh e^{it/2}}{2(\eta\cosh e^{it/2} + \gamma)}\right|^2 \\ &= \frac{A_N(\tau)}{A_D(\tau)} \quad (-\pi/2 \leq \tau = t/2 \leq \pi/2) \end{aligned}$$

where

$$\begin{aligned}A_N(\tau) = {} & (\cosh(\cos\tau)(4\eta^2\sin^2(\sin\tau)\cos(\sin\tau)\sinh^2(\cos\tau) + 4\gamma(\gamma-2\eta)\cos(\sin\tau)\\ & - 2\gamma k\sin(\sin\tau)\sin\tau - \eta k\sin(2\sin\tau)\sin\tau + \eta k\cos\tau\sinh(2\cos\tau)\\ & + 2\gamma k\cos\tau\cos(\sin\tau)\sinh(\cos\tau) + 4\eta^2\cos^3(\sin\tau)\cosh^3(\cos\tau) - 4\gamma^2\\ & - 4\eta(\eta\sin^2(\sin\tau)\sinh^2(\cos\tau) + (\eta-2\gamma)\cos^2(\sin\tau)\cosh^2(\cos\tau)))^2\\ & + (\sinh(\cos\tau)(2\gamma k\sin\tau\cos(\sin\tau) - 2(\eta^2-2\gamma^2)\sin(\sin\tau) + \eta^2\sin(3\sin\tau))\\ & + \eta\sinh(2\cos\tau)(k\sin\tau + 2\gamma\sin(2\sin\tau)) + \eta k\sin(2\sin\tau)\cos\tau\\ & + 2\gamma k\sin(\sin\tau)\cos\tau\cosh(\cos\tau) + 3\eta^2\sin(\sin\tau)\sinh(\cos\tau)\cosh^2(\cos\tau)\\ & + \eta^2\sin(\sin\tau)\sinh^3(\cos\tau))^2\end{aligned}$$

and

$$A_D(\tau) = 16((\gamma + \eta\cos(\sin\tau)\cosh(\cos\tau))^2 + \eta^2\sin^2(\sin\tau)\sinh^2(\cos\tau))^2.$$

Assume $F(\tau) = A_N(\tau) - ((s+1)^2-1)^2 A_D(\tau)$, where $0 \le \tau \le \pi/2$. Now

$$\begin{aligned}F(0) = {} & (4(c_1-1)(\eta c_1+\gamma)^2 + k(\eta\sinh 2 + 2\gamma\sinh 1))^2\\ & - 16((s+1)^2-1)^2(\eta c_1+\gamma)^4 \end{aligned} \tag{5.16}$$

and

$$\begin{aligned}F\left(\frac{\pi}{2}\right) = {} & (4(c_0-1)(\eta c_0+\gamma)^2 - k(\eta\sin 2 + 2\gamma\sin 1))^2\\ & - 16((s+1)^2-1)^2(\eta c_0+\gamma)^4.\end{aligned}$$

Let $X(k,s) := F(0)$ and $Y(k,s) := F(\pi/2)$. For each $k \ge 1$, it can be easily verified that $X(k,s)$ is a monotonically increasing function of k, provided $\eta \ge \eta_0$, therefore $X(k,s) \ge X(1,s)$. Moreover due to the conditions given in (5.14), $\eta\sinh 2 + 2\gamma\sinh 1 + 4(c_1-(s+1)^2)(\eta c_1+\gamma)^2 \ge 0$, holds for each $0 < s \le \sqrt{c_1}-1$, provided $\eta \ge \eta_0$. Thus as a result of the equation in (5.16), $X(1,s) \ge 0$, for each $0 < s \le \sqrt{c_1}-1$. Similarly for each $\sqrt{c_1}-1 < s \le 1/\sqrt{2}$, the inequality $4((s+1)^2-c_1)(\eta c_1+\gamma)^2 \le \eta\sinh 2 + 2\gamma\sinh 1$ is valid if $\eta_0 \le \eta \le \eta_0+\eta_1$. Eventually, (5.16) gives $F(0) \ge 0$. Further $Y(k,s) \ge Y(1,s)$ provided $\eta \le \eta_2$. The condition $\eta \ge \eta_2 + \eta_3$, leads to $Y(1,s) \ge 0$, for each $0 < s \le 1/\sqrt{2}$. Therefore $F(\pi/2) \ge 0$. Hence $F(\tau) \ge 0$ for each $\tau \in [-\pi/2, \pi/2]$. This gives a contradiction to (5.15). Hence $p(z) \prec \cosh\sqrt{z}$. ∎

Corollary 5.13. *Let $0 < s \le 1/\sqrt{2}$ and $(\eta,\gamma) \in \mathbb{R}\times\mathbb{R} - \{(\eta,-\eta)\}$ satisfy the conditions given in (5.14), if $\vartheta \in \mathcal{S}^*_L(s)$, then $\mathcal{L} \in \mathcal{S}^*_\varrho$.*

Example 4. *Let $\gamma = -3/5$ and $s = 1/2$, then Theorem 5.12 gives*

$$\frac{3}{5c_0} - \frac{2\sin 1}{c_0(9-4c_0)} \le \eta \le \frac{3}{5c_1} + \frac{2\sinh 1}{c_1(9-4c_1)}.$$

Choose $a = 1 = n$ and $h(z) = (1+(1/2)z)^2$ in [20, Theorem 3.2d, p.86]*, then $\eta h(z) + \gamma = \eta(1+(1/2)z)^2 - 3/5$ and $\mathrm{Re}(\eta a + \gamma) = \mathrm{Re}(\eta - 3/5) > 0$. Also $\eta(1+(1/2)z)^2 - 3/5 \prec R_{\eta-3/5,1}(z)$ where $R_{\eta-3/5,1}(z) = (\eta-3/5)(1+z)/(1-z) + 2z/(1-z^2)$ is an open door function, that is univalent in $\mathbb{D}$. Then*

$$q_\eta(z) = \left(\eta\int_0^1 t^{-8/5}e^{(t-1)(z\eta(z(1+t)+8))/8}dt\right)^{-1} + \frac{3}{5\eta}$$

is a solution of the differential equation $q(z) + zq'(z)/(\eta q(z)+\gamma) = (1+(1/2)z)^2$, then $q_\eta(z) \prec \cosh\sqrt{z}$. Furthermore as $h(z)$ is a convex function, so by [20, Theorem 3.2j, p.97] *$q_\eta(z)$ is univalent. For $n = 1$, if $p \in \mathcal{H}(1,1)$ satisfies (5.15), then $p \prec q_\eta$ and $q_\eta(z)$ is the best $(1,1)-$dominant.*

The proof of Theorem 5.14 is much akin to the previous theorems, so we omit the proof.

Theorem 5.14. *Let $-1 \leq B < A \leq 1$ and $(\eta, \gamma) \in \mathbb{R} \times \mathbb{R} - \{(\eta, -\eta)\}$ satisfy the following conditions:*

(i) $(1 - B^2)\sinh 1 + 2(\gamma + \eta c_1)(c_1 - 1 + B(A - Bc_1)) \geq 0$
(ii) $(\sinh 1 + 2(c_1 - 1)(\gamma + \eta c_1))^2 \geq (B \sinh 1 - 2(A - Bc_1)(\gamma + \eta c_1))^2$
(iii) $(1 - B^2)\sin 1 + 2(\gamma + \eta c_0)(1 - c_0 - B(A - Bc_0)) \geq 0$
(iv) $(\sin 1 + 2(c_0 - 1)(\gamma + \eta c_0))^2 \geq (B \sin 1 + 2(A - Bc_0)(\gamma + \eta c_0))^2$

Let $p(z)$ be an analytic function, such that $p(0) = 1$ and satisfies

$$p(z) + \frac{zp'(z)}{\eta p(z) + \gamma} \prec \frac{1 + Az}{1 + Bz}$$

then

$$p(z) \prec \cosh \sqrt{z}.$$

Corollary 5.15. *Let $\vartheta \in \mathcal{S}^*(A, B)$ and parameters A, B and $(\eta, \gamma) \in \mathbb{R} \times \mathbb{R} - \{(\eta, -\eta)\}$ satisfy the conditions given in Theorem 5.14 then $\mathcal{L} \in \mathcal{S}^*_\varrho$.*

Corollary 5.16. *If an analytic function $p(z)$ with $p(0) = 1$, satisfies:*

$$\left| p(z) + \frac{zp'(z)}{\eta p(z)} - 1 \right| < 1,$$

where $-\sinh 1 \leq 2\eta c_1^2 \leq c_1 \sinh 1/(2 - c_1)$ or

$$\left| p(z) + \frac{zp'(z)}{p(z) + \gamma} - \frac{4}{3} \right| < \frac{2}{3},$$

where $-c_0 + (\sin 1)/2(c_0 - 2) \leq \gamma \leq -c_1 - (\sinh 1)/2(c_1 - 2)$. Then

$$p(z) \prec \cosh \sqrt{z}.$$

Proof. By replacing $A = 1$ and $B = 0$ with $\gamma = 0$ or $A = 0$ and $B = -1/2$ with $\eta = 1$ in Theorem 5.14 we establish the required result. ∎

References

[1] Ali, R.M., Subramanian, K.G., Ravichandran V., Ahuja, O.P.: Neighborhoods of starlike and convex functions associated with parabola. J. Inequal. Appl. **346279**, 9 (2008)
[2] Alotaibi A., Arif M., Alghamdi, M.A., Hussain, S.: Starlikness associated with cosine hyperbolic function. Mathematics **8**(7), 1118(2020). https://doi.org/10.3390/math8071118
[3] Aouf, M. K., Dziok J., Sokół, J.: On a subclass of strongly starlike functions. Appl. Math. Lett. **24**(1), 27–32(2011)
[4] Bano, K., Raza, M.: Starlike functions associated with cosine functions. Bull. Iranian Math. Soc. **47**(5), 1513–1532(2021)
[5] 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)
[6] Ebadian, A., Bulboacă, T., Cho, N.E., Adegani, E. A.: Coefficient bounds and differential subordinations for analytic functions associated with starlike functions. Rev. R. Acad. Cienc. Exactas Fís. Nat. Ser. A Mat. RACSAM **114**(3), 1–19(2020)
[7] 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
[8] 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)
[9] Janowski,W.: Some extremal problems for certain families of analytic functions. I. Ann. Polon. Math. **28**(3), 297–326(1973)
[10] Kanas, S., Masih, V. S., Ebadian, A.: Relations of a planar domains bounded by hyperbolas with families of holomorphic functions. J. Inequal. Appl, **2019**(1), 1–14(2019)

[11] Kanas, S., Wiśniowska, A.: Conic domains and starlike functions. Rev. Roumaine Math. Pures Appl. **45**(4), 647–658(2000)
[12] Kumar, S. S., Kamaljeet, G.: A cardioid domain and starlike functions. Anal. Math. Phys. **11**(2), 1-34(2021)
[13] Kumar, V.: Moduli difference of successive inverse coefficients for certain classes of close-to-convex functions. Ricerche di Matematica, (2021). https://doi.org/10.1007/s11587-021-00682-1
[14] Lecko, A., Sim, Y. J.: Coefficient problems in the subclasses of close-to-star functions. Results Math. **74**(3), 104(2019)
[15] Lee, S. K., Khatter, K., Ravichandran, V.: Radius of starlikeness for classes of analytic functions. Bull. Malays. Math. Sci. Soc. **43**(6), 4469–4493(2020)
[16] 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
[17] Masih, V. S., Kanas, S.: Subclasses of starlike and convex functions associated with the Limaçon Domain. Symmetry **12**, 942(2020)
[18] Mendiratta, R., Nagpal, S., Ravichandran, V.: On a subclass of strongly starlike functions associated with exponential function. Bull. Malays. Math. Sci. Soc. **38**(1), 365–386(2015)
[19] Stankiewicz, J.: Quelques problèmes extrémaux dans les classes des fonctions α-angulairement étoilées, Ann. Univ. Mariae Curie-Skłodowska Sect. A **20**(1966), 59–75(1971).
[20] Miller, S. S., Mocanu, P. T.: Differential subordinations, Theory and Applications, vol.225. Marcel Dekker, New York (2000)
[21] Mundalia, M., Shanmugam, S. K.: Coefficient bounds for a unified class of holomorphic functions. In Mathematical analysis. I. Approximation theory, pp. 197–210, Springer Proc. Math. Stat., 306, Springer, Singapore.
[22] 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)
[23] 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)
[24] Ruscheweyh, S.: Convolutions in Geometric Function Theory. Les Presses De L'Universite De Montreal, Montreal (1982)
[25] Saliu, A., Noor, K. I., Hussain, S., Darus, M.: Some results for the family of univalent functions related with limaçon domain. AIMS Math. **6**(4), 3410–3431(2021)
[26] Shah, G. M.: On the univalence of some analytic functions. Pac. J. Math. **43**, 239–250(1972)
[27] Ravichandran, V., Kumar, S. S.: On sufficient conditions for starlikeness. Southeast Asian Bull. Math. **29**(4), 773–783(2005)
[28] Kumar, S., Kumar, V., Ravichandran, V., Cho, N. E.: Sufficient conditions for starlike functions associated with the lemniscate of Bernoulli. J. Inequal. Appl. **2013**, 13(2013).
[29] Robertson, M.I.S.: On the theory of univalent functions. Ann. Math. **37**(2), 374–408(1936)
[30] Sharma, K., Aghalary, R., Ravichandran, V.: Briot-Bouquet differential subordination and Bernardi's integral operator. Preprint at https://arxiv.org/pdf/2101.06041v1.pdf (2021)
[31] Sokół, J., Stankiewicz, J.: Radius of convexity of some subclasses of strongly starlike functions. Zeszyty Nauk. Politech. Rzeszowskiej Mat. **19**, 101–105(1996).
[32] Uralegaddi, B. A., Ganigi, M. D., Sarangi, S. M.: Univalent functions with positive coefficients. Tamkang J. Math. **25**(3), 225–230(1994)

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

arXiv:2201.06939v1 [math.NT] 3 Jan 2022

ON POWER VALUES OF SUM OF DIVISORS FUNCTION IN ARITHMETIC PROGRESSIONS

SAI TEJA SOMU AND VIDYANSHU MISHRA

Abstract. Let $a \geq 1, b \geq 0$ and $k \geq 2$ be any given integers. It has been proven that there exist infinitely many natural numbers m such that sum of divisors of m is a perfect kth power. We try to generalize this result when the values of m belong to any given infinite arithmetic progression $an + b$. We prove that there exist infinitely many natural numbers n such that sum of divisors of $an + 1$ is a perfect kth power. We also prove that, either sum of divisors of $an + b$ is not a perfect kth power for any natural number n or sum of divisors of $an + b$ is a perfect kth power for infinitely many natural numbers n.

1. Introduction

Sum of divisors function (denoted by σ) is an important function in number theory. Perfect numbers, natural numbers n satisfying $\sigma(n) = 2n$, have been studied as early as 300 BC. Euclid showed in [E] that $\frac{q(q+1)}{2}$ is an even perfect number whenever q is a prime number of the form $2^p - 1$, for natural number p. Euler (in [LE]) proved that all even perfect numbers are of this form. It is still not known whether there are infinitely many perfect numbers and whether there exists an odd perfect number.

The problem of finding natural numbers n such that $\sigma(n)$ is a perfect kth power has been considered by 17th century mathematicians like Fermat, Wallis, and Frenicle. Fermat in 1657, asked two problems: (i) Find a cube $m = n^3$ for which $\sigma(m)$ is a perfect square (ii) Find a square $m = n^2$ for which $\sigma(m)$ is a perfect cube (see Chapter 2, p.54 of [D66] for detailed history).

In [B12] (see also [F12], [Z18]) it has been proven that for any natural number $k \geq 2$, $\sigma(m)$ is a perfect kth power for infinitely many natural numbers m. There are several sequences in Neil Sloane's online encyclopedia of integer sequences [S] dedicated to this problem, namely A006532 ($\sigma(m)$ is perfect square), A020477 ($\sigma(m)$ is perfect cube), A019422 ($\sigma(m)$ is perfect fourth power), A019423 ($\sigma(m)$ is perfect fifth power), A019424 ($\sigma(m)$ is

Date: January 19, 2022.
2020 *Mathematics Subject Classification.* Primary: 11A25, Secondary: 11B25.
Key words and phrases. Sum of divisors, Power Values, Arithmetic Progressions.

perfect sixth power), A048257 ($\sigma(m)$ is perfect seventh power) and A048258 ($\sigma(m)$ is perfect eighth power).

In this paper, we investigate this problem when the values of m belong to any given infinite arithmetic progression $\{an + b : n \in \mathbb{N}\}$ (here $\mathbb{N} = \{1, 2, 3, \ldots\}$). We study the following problem.

Q. *For any given integers $a \geq 1, b \geq 0, k \geq 2$ do there exist infinitely many natural numbers n such that $\sigma(an + b)$ is a perfect kth power?*

In the next section, we give an answer to the above problem when $b = 1$. We prove the following theorem.

Theorem 1. *For any natural number a and $k \geq 2$ there exist infinitely many natural numbers n such that $\sigma(an + 1)$ is a perfect kth power.*

For general b, we prove the following theorem.

Theorem 2. *For any given integers $a \geq 1, b \geq 0$ and $k \geq 2$, either $\sigma(an+b)$ is not a perfect kth power for any natural number n or $\sigma(an+b)$ is a perfect kth power for infinitely many natural numbers n.*

In the final section, we give few open questions for further research.

2. Proofs of Theorem 1 and Theorem 2

We require some lemmas in order to prove Theorem 1 and Theorem 2. We will use a group theoretic result mentioned in [A94] (Theorem 2 of [A94]).

Theorem(Alford et al.) 3. *If G is a finite abelian group in which the maximal order of an element is m, then in any sequence of at least $m(1 + \log(|G|/m))$ (not necessarily distinct) elements of G, there is a nonempty subsequence whose product is the identity.*

Using the above theorem we can prove this lemma.

Lemma 1. *Let $k \geq 2$ be any natural number and P be any nonempty finite set of primes. Let A be any finite set of natural numbers such that all the prime factors of all elements of A belong to P. If $|A| \geq k|P| \log k + 1$ then there exists a nonempty subset $A' \subset A$ such that $\prod_{n \in A'} n$ is a perfect kth power.*

Proof. Let $n_1, \ldots, n_{|A|}$ be distinct elements of A, and let $p_1, \ldots, p_{|P|}$ be distinct elements of P. Let $G = (\mathbb{Z}/k\mathbb{Z})^{|P|}$. As all the prime factors of

n_i are in P there exist nonnegative integers $a_{i,1}, \ldots, a_{i,|P|}$ such that $n_i = p_1^{a_{i,1}} \cdots p_{|P|}^{a_{i,|P|}}$. Now define $e_i \in G$ in the following way:

$$e_i = (a_{i,1} \bmod k, \ldots, a_{i,|P|} \bmod k),$$

for $1 \leq i \leq |A|$. Now for group G, we have that the maximal order(m) of group equal to k, and number of elements in the group is equal to $|G| = k^{|P|}$. As

$$\begin{aligned} m(1 + \log(|G|/m)) &= k(1 + \log(k^{|P|-1})) \\ &= k|P| \log k + k - k \log k \\ &\leq k|P| \log k + 1 \\ &\leq |A|, \end{aligned}$$

we have, number of elements in the sequence e_i, $|A| \geq m(1 + \log(|G|/m))$, therefore we can apply above theorem. From Theorem 3 (Theorem 2 of [A94]), there exists a nonempty subsequence $e_{m_1}, \ldots, e_{m_s}$ such that

$$\sum_{i=1}^{s} e_{m_i} = (0, \ldots, 0).$$

Let $A' = \{n_{m_1}, \ldots, n_{m_s}\} \subset A$ then it follows from definition of e_i and $\sum_{i=1}^{s} e_{m_i} = (0, \ldots, 0)$ that all the exponents in the prime factorization $\prod_{n \in A'} n$ are multiples of k, which implies that $\prod_{n \in A'} n$ is a perfect kth power. □

Let ϕ denote Euler's totient function. Using the fact that the infinite product over primes $\prod_p \left(1 - \frac{1}{p}\right)^{-1}$ diverges we prove the following lemma.

Lemma 2. *For any given natural numbers $a, k \geq 2$ there exists a natural number b which is relatively prime to a, and satisfying the inequality*

$$\frac{b}{\phi(b)} > (k \log k + 1)\phi(a).$$

Proof. As the product over primes $\prod_{p \nmid a} \left(1 - \frac{1}{p}\right)^{-1}$ diverges, there exist primes $p_1, \ldots, p_t$ with $p_i \nmid a$ for $1 \leq i \leq t$ and

$$\prod_{i=1}^{t} \left(1 - \frac{1}{p_i}\right)^{-1} > (k \log k + 1)\phi(a).$$

Let $b = \prod_{i=1}^{t} p_i$, as $p_i \nmid a$ for $1 \leq i \leq t$ we can conclude that b is relatively prime to a. Now,

$$\frac{b}{\phi(b)} = \prod_{i=1}^{t} \left(\frac{p_i}{p_i - 1}\right) = \prod_{i=1}^{t} \left(1 - \frac{1}{p_i}\right)^{-1} > (k \log k + 1)\phi(a),$$

which proves the lemma. □

Using Lemma 1 and Lemma 2, we can prove the following lemma.

Lemma 3. *For any natural numbers a and $k \geq 2$ there exists a positive real number $x_0(a)$ such that for all $x \geq x_0(a)$ there exists a natural number n such that all the prime factors of $an+1$ are in the interval $(\sqrt{x}, x]$ and $\sigma(an+1)$ is a perfect kth power.*

Proof. Given natural numbers a and $k \geq 2$, from Lemma 2, there exists a natural number b relatively prime to a such that

$$\frac{b}{\phi(b)} > (k \log k + 1)\phi(a).$$

For any $x \geq b^2$, define

$$A = \{p+1 : p \text{ prime}, \sqrt{x} < p \leq x, p \equiv 1 \bmod a, p \equiv -1 \bmod b\},$$

and

$$P = \left\{p : p \text{ prime}, p \leq \frac{x+1}{b}\right\}.$$

If $p+1 \in A$ then as $p \equiv -1 \bmod b$ we have $p+1 = b(\frac{p+1}{b})$, as $b \leq \frac{x+1}{b}$ and $\frac{p+1}{b} \leq \frac{x+1}{b}$, all prime factors of $p+1$ are less than or equal to $\frac{x+1}{b}$. Therefore, all prime factors of $p+1$ lie in the set P. From Prime Number Theorem and Prime Number Theorem for Arithmetic progressions we have $|A| \sim \frac{x}{\phi(a)\phi(b)\log x}$ and $|P| \sim \frac{x}{b \log x}$, which implies $k|P| \log k + 1 \sim \frac{kx \log k}{b \log x}$.

As $\frac{b}{\phi(b)} > (k \log k+1)\phi(a)$ we have $\frac{1}{\phi(a)\phi(b)} > \frac{k \log k}{b}$. Since, $|A| \sim \frac{x}{\phi(a)\phi(b)\log x}$ and $k|P| \log k+1 \sim \frac{kx \log k}{b \log x}$, we must have $|A| \geq k|P| \log k+1$ for sufficiently large $x \geq x_0(a)$, hence we can apply Lemma 1 for $x \geq x_0(a)$. Therefore, for $x \geq x_0(a)$, there exists a nonempty subset $A' \subset A$ such that

$$\prod_{p+1 \in A'} (p+1) = \sigma\left(\prod_{p+1 \in A'} p\right)$$

is a perfect kth power.

Since, $p+1 \in A'$ implies $p \equiv 1 \bmod a$ we have $\prod_{p+1 \in A'} p \equiv 1 \bmod a$. Hence, $\prod_{p+1 \in A'} p$ is of the form $an+1$ whose sum of divisors is a perfect kth power and all prime divisors of $\left(\prod_{p+1 \in A'} p\right)$ lie in the interval $(\sqrt{x}, x]$. □

We can now prove Theorem 1 and Theorem 2 using Lemma 3.

2.1. **Proof of Theorem 1.**

Proof. Assume for a contradiction that there are only finitely many n for which $\sigma(an+1)$ is a perfect kth power, then there exists an $x_1 > 0$ such that $\sigma(an+1)$ is not a perfect kth power for any $n \geq x_1$. This implies that there cannot be any $an+1$ all of whose all prime factors lie in the range $(\sqrt{x}, x]$

for $x = \max\{x_0(a), (ax_1+1)^2\}$ and whose sum of divisors is a perfect kth power, contradicting Lemma 3. □

2.2. **Proof of Theorem 2.**

Proof. If $\sigma(an+b)$ is not a perfect kth power for all natural numbers n, there is nothing to prove. If there exists a natural number n' such that $\sigma(an'+b)$ is perfect kth power then define a sequence x_n with $x_1 = \max\{(an'+b)^2, x_0(a)\}$, $x_{n+1} = x_n^2$ for $n \geq 1$.

From Lemma 3, for all natural numbers m there exists n_m such that $\sigma(an_m+1)$ is a perfect kth power and all prime factors of an_m+1 lie in the interval $(\sqrt{x_m}, x_m]$. Since, $(\sqrt{x_m}, x_m]$ are disjoint intervals for $m \in \mathbb{N}$, natural numbers an_m+1 along with $an'+b$ are pairwise relatively prime and hence distinct.

Define $s_m = (an'+b)(an_m+1)$ for all natural numbers m. We can observe that s_m is of the form $an+b$ and

$$\sigma(s_m) = \sigma(an'+b)\sigma(an_m+1)$$

is a perfect kth power as product of perfect kth powers is also a perfect kth power. Therefore, if there exists a natural number n' such that sum of divisors of $an'+b$ is a perfect kth power then there exist infinitely many natural numbers n such that sum of divisors of $an+b$ is a perfect kth power. □

3. Future prospects

We ask five questions Q1 to Q5 in this section.

Given a, b and $k \geq 2$, Theorem 2 leaves us with two possibilities, either $\sigma(an+b)$ is perfect kth power for infinitely many n or $\sigma(an+b)$ is never a perfect kth power. This implies that if there exists a natural number n such that $\sigma(an+b)$ is perfect kth power then there are infinitely many n such that $\sigma(an+b)$ is a perfect kth power. So we ask the following question.

Q1. *Do there exist integers $a \geq 1, b \geq 0, k \geq 2$ such that $\sigma(an+b)$ is not a perfect kth number for any natural number n?*

If the answer to Q1 is no, then from Theorem 2, it follows that for all integers $a \geq 1, b \geq 0, k \geq 2$ there exist infinitely many n such that $\sigma(an+b)$ is a perfect kth power.

For natural numbers a, $k \geq 2$ and integer b such that $0 \leq b \leq a-1$, let $N(x; a, b, k)$ denote number of natural numbers $n \leq x$ such that $n \equiv a \mod b$ and $\sigma(n)$ is perfect kth power.

Q2. *Given natural numbers* $a, 0 \leq b_1, b_2 \leq a-1, k \geq 2$ *such that* $N(x; a, b_1, k) > 0$, $N(x; a, b_2, k) > 0$ *for some real number* x. *Is it true that*

$$\liminf_{x\to\infty} \frac{N(x; a, b_1, k)}{N(x; a, b_2, k)} > 0?$$

We will now ask a stronger version of the above question. Are perfect kth power values of $\sigma(n)$ equally distributed among residue classes modulo a?

Q3. *Given natural numbers* $a, 0 \leq b_1, b_2 \leq a-1, k \geq 2$, *is it true that there exists a real number* x *such that* $N(x; a, b_1, k) > 0$, $N(x; a, b_2, k) > 0$ *and*

$$\lim_{x\to\infty} \frac{N(x; a, b_1, k)}{N(x; a, b_2, k)} = 1?$$

Numerical evidence suggests that Q3 may not always be true, for example when we consider perfect square values of $\sigma(n)$ for $n \leq x$ and residue classes $n \equiv 0 \bmod 7$ and $n \equiv 1 \bmod 7$, we have the following data.

x	$N(x; 7, 0, 2)$	$N(x; 7, 1, 2)$	$\frac{N(x;7,0,2)}{N(x;7,1,2)}$
100	2	2	1.0000...
1000	16	5	3.2000...
10000	68	28	2.4285...
100000	307	147	2.0884...
1000000	1508	748	2.1299...
10000000	7562	3811	1.9842...
100000000	37652	18882	1.9940...

TABLE 1. Comparison of $N(x; 7, 0, 2)$ and $N(x; 7, 1, 2)$

Table 1 suggests that, there might be larger number of natural numbers n of the form $7k$ compared to $7k+1$ for which $\sigma(n)$ is a perfect square. We ask the following question.

Q4. *Is it true that*

$$\limsup_{n\to\infty} \frac{N(x; 7, 0, 2)}{N(x; 7, 1, 2)} > 1?$$

If the answer for question Q4 is yes, then the answer for Q3 is no.

Till now, we have been dealing with linear polynomials $an+b$. A natural generalization would be the following question.

Q5. *Let* $p(n)$ *be any given polynomial with nonnegative integer coefficients and* $k \geq 2$ *be any natural number. Is* $\sigma(p(n))$ *perfect* k*th power for infinitely many* n*?*

Question 2.5 (when $p(n) = n^l$), Conjecture 2.6 (when $p(n) = n^2$), Conjecture 2.7 (when $p(n) = n^3$) of [B12] are special cases of the above question. In Neil Sloane's online encyclopedia of integer sequences [S] we find few sequences related to this problem, namely, A008847 ($\sigma(n^2)$ is a perfect square, $p(n) = n^2$, $k = 2$), A008850 ($\sigma(n^2)$ is a perfect cube, $p(n) = n^2$, $k = 3$) and A008849 ($\sigma(n^3)$ is a perfect square, $p(n) = n^3$, $k = 2$).

References

[A94] W. R. Alford, A. Granville, C. Pomerance, *There are infinitely many Carmichael numbers*, Annals of Mathematics, **139** (1994) 703–722.

[B12] F. Beukers , F. Luca , F. Oort, *Power values of divisor sums*, The American Mathematical Monthly 119, **5** (2012), 373–380.

[D66] L. E. Dickson, *History of the Theory of Numbers*, Vol. I, Divisibility and Primality, Chelsea Publishing, New York, 1966.

[E] Euclid, *Elements*, Book IX, Proposition 36.

[LE] L. Euler, *De numeris amicibilibus*, Commentationes arithmeticae (in Latin, published posthumously), **2** (1849), 627–636.

[F12] T. Freiberg, *Products of shifted primes simultaneously taking perfect power values*, J. Aust. Math. Soc. (special issue dedicated to Alf van der Poorten) **92** (2012), 145–154.

[S] N. J. Sloane, *The Online Encyclopedia of integer sequences*, available at http://oeis.org.

[Z18] B. Żmija, *Power values of arithmetic functions*, Colloquium Mathematicum **151** (2018), 19–25.

Aspireal Technologies Pvt. Limited, Hyderabad, India
Email address: somuteja@gmail.com

Department of Applied Mathematics, Delhi Technological University, Delhi, India
Email address: vidyanshurpvvyv@gmail.com

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAI.2022.3144651, IEEE Transactions on Artificial Intelligence

Optimal Feature Selection for Imbalanced Text Classification

Anshu Khurana and Om Prakash Verma

***Abstract*—Textual data suffers from two main problems; large number of features and class imbalance. Many conventional approaches and their variants exist in literature to solve both these problems. The classical Synthetic Minority Over-sampling Technique (SMOTE) is the most explored technique for balancing the dataset. We introduced a new algorithm to balance the dataset, named Distributed SMOTE (D_SMOTE), which overcomes the problem of lack of density and reducing the formation of small disjuncts. Further, another problem handled is the large number of features or high-dimensionality. To solve high-dimensionality, a novel feature selection technique is introduced known as modified Biogeography-Based Optimization (M_BBO). The proposed model, M_BBO, performs modification in ranking of variables using feature weighting algorithm rather than randomly ranking. We have proposed two new expressions in D_SMOTE and one new expression in M_BBO. The extensive experimental results are computed out on four text classification datasets with four machine learning classifiers. The results are concluded using three performance measures: 1) Area Under Curve, 2) G-mean and 3) F1-score. Our empirical and statistical observation for four class-imbalanced datasets shows that the proposed D_SMOTE outperforms the other similar oversampling technique. We have also compared our proposed algorithm, M_BBO+D_SMOTE, with other models on seventeen imbalanced text classification datasets. Our model outperformed the other models in fourteen datasets. We have also compared our model with Bidirectional Encoder Representations from Transformers (BERT). To validate the experimental analysis, statistical Friedman test is employed.**

***Impact Statement*—Text classification has become a major avenue in generating valuable insights. It is being vastly used to solve real world problems by performing sentimental analysis, detecting frauds and patterns in various sectors like healthcare, e-commerce, sports etc. In Big Data, the performance of text classification can be improved by selecting relevant features and handling of imbalance problems between the distribution of classes in the dataset. In the past, the research work has mostly been done on optimizing the classifier and tuning the parameters and has deviated from the natural distribution of the data itself. There has now been a radical shift in this approach with the emergence of data science, where the focus is now on understanding the data and feature selection. The proposed algorithm overcomes the limitation of random ranking of SIV variables used in Biogeography-based optimization algorithm. It also handles the problem of lack of density in SMOTE to solve imbalance problem. The technique has been tested on seventeen datasets for AUC score and performed best for fourteen datasets, when compared to other models.**

***Index Terms*—Text Classification, Distributed SMOTE, Modified Biogeography-Based Optimization, Class imbalance and Feature Selection.**

Department of Computer Science and Engineering
Delhi Technological University

Department Of Electronics and Communication
Delhi Technological UniversityE-mail: see opverma.dce@gmail.com

I. Introduction

As there is a rise in research areas, advancement in Information Communication and Technology (ICT) gives opportunities to all users to access the information quickly and at a faster rate. Due to an increase in demand for the information, there is a simultaneous increase in the number of text documents as most of the information is stored in text form [43]. Text classification is widely used for data analysis and prediction. Many machine learning algorithms are applied in different application areas of text classification such as Naïve Baye's (NB) [19], [15], K- Nearest Neighbor (KNN) [15], Support Vector Machine (SVM) [8], [17], [15], Random Forest (RF) [26], Decision Tree (DT) [26] and Ensemble classifiers [57]. Some real-world applications in text classification are disease prediction [49], Text categorization [16] and fraud detection [61].

During text classification process, there are many related issues that arise specifically are class imbalance and high dimensional dataset [68]. Class imbalance [40] has become a major issue in determining the classification performance, hence many researchers are working on this problem as imbalance causes adverse effect on the correct predicted results [3]. Many techniques have been developed for balancing the binary imbalanced dataset [49]. But however, the main challenge which is faced by researchers is imbalanced data classification for multi-class dataset [5]. The existing solutions to multi-class imbalanced data is generally the extension of algorithm of binary class imbalance classifier. The construction of Decision Trees [41] and Ensemble based classifiers also proved to give better results in multi-class [35]. The common oversampling technique used is Synthetic Minority Oversampling Technique (SMOTE) and its different variations [24]. The classical SMOTE and its variant suffer with a problem of small disjuncts which lead to less training data, hence lack of information. In our research, we overcome the problem of lack of information caused by lack of density or class disjuncts.

The other problem that leads to the classification performance is high-dimensionality. The low weight features or noise, that is features which are not essential with respect to data, decreases the efficiency of the algorithm [26], hence degrading the performance of text classification. Feature selection is a process of selecting only desired features related to the data which helps in increasing the efficiency of the model [22]. The use of metaheuristic algorithm in the literature, as feature selection technique is proved as beneficial to improve classification Accuracy [78], [34], [68]. Various optimization algorithms, such as Genetic algorithm (GA) [70], Ant colony optimization (ACO) [55], Firefly algorithm [72],

2691-4581 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 31,2022 at 08:55:19 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAI.2022.3144651, IEEE Transactions on Artificial Intelligence

Particle Swarm optimization (PSO) [75], Biogeography-based optimization(BBO) [65] and many more are used as feature selection techniques. The BBO algorithm for feature selection was proposed in [65], where the proposed technique solved the problem of best selection of sensor for aircraft engine and BBO outperformed other nature optimization techniques such as GA, PSO and Differential Evolution. In [66], results again proved that BBO outperformed GA. In many classification problems, BBO is used for optimization and has proved best among others [59], [58], [4]. The new model for language representation was proposed by [18] known as BERT (Bidirectional Encoder Representations from Transformers). It is widely used for many artificial intelligence real world applications [32]. But, main limitation of using BERT arises when dealing with low dimensional dataset.

The proposed technique for balancing the dataset by overcoming the problem of class disjuncts or lack of density is referred as Distributed SMOTE (D_SMOTE). For reducing the problem of high-dimensionality, a novel modified Biogeography based optimization approach (M_BBO) is proposed in this paper. The proposed modification is based on changing the selection procedure of suitability index variable (SIV) in the mutation process. The training is performed using four machine learning algorithms namely: Naive Baye's (NB), Random Forest (RF), Logistic regression (LR) and Support Vector Machine (SVM). The performance measures used to validate the model are: Area Under Curve (AUC), G-mean and F1-score. We compared our model with other techniques, and found that our model for optimal text classification is the most suitable among all. The proposed model is also compared with the other state-of-the-art models available in the literature and the new technique BERT.

A. Contribution of the paper

1) The main motivation arises from the widespread problem of multi-class imbalance problem, and solutions to solve the problem. Existing solution to the problem is the model-oriented approach or data-oriented approach. Our main focus lies towards the data-oriented oversampling technique. The SMOTE, which is extensively used for over-sampling the minority instances, is one of the emerging and efficient technique [9]. The proposed approach is a modification in the existing SMOTE algorithm and named as Distributed SMOTE (D_SMOTE) which overcomes the problem of lack of density that leads to small disjuncts. Hence, two new expressions are introduced in the proposed algorithm. The proposed technique is compared with six oversampling techniques and our technique outperformed the other techniques in three datasets.

2) High number of features is also the main problem of dataset which cause noise in text and affects the classification performance of the model. The relevant features generate more appropriate results, and hence improving the classification performance. For feature selection, we have proposed a modified BBO algorithm (M_BBO). BBO is a nature-based optimization [66] algorithm. In Classical BBO mutation, the selection of a suitable index variable is a random process. Randomly choosing the variable can neglect the relevant variables and hence affects the performance of the classification. So, we ranked these variables using Recursive Feature Elimination (RFE) method, and the lowest rank SIV is mutated with higher rank SIV. Hence, relevant features are considered which helps in improving the classification efficiency. The BBO has high convergence and good exploration properties. Therefore, choosing BBO was a wise decision to improve the classification performance.

3) The proposed technique of Distributed SMOTE is compared with other variants of oversampling techniques. Our technique outperformed in three datasets. The proposed model M_BBO + D_SMOTE is compared with other combinations of models for the validation of the model. Our model is compared with five state-of-the-art techniques on seventeen datasets. The model is also compared with BERT on four datasets. The datasets considered for experimenting are low-dimensional as well as high-dimensional to validate the generalizability.

4) The extensive experiments are performed with 50 runs on four high-dimensional datasets. The four classifiers used for classification namely; Naive Baye's (NB), Random Forest (RF), Logistic Regression (LR) and Support Vector Machine (SVM). All results are validated using 10-fold-cross validation. To prove statistical significance Friedman rank test is employed.

The remaining part of this paper is organized in the following order. State-of-the-art-literature is presented in Section II. The theoretical concept of the proposed methodology is elaborated in Section III. Performance measure, used to evaluate the model and datasets description, is explained in Section IV. Next Section V throws light on the complete analysis of results with statistical comparison. Finally, the proposed technique with future work is concluded in section VI.

II. State-of-the-Art Literature

The focus of this section is on the related research carried on different techniques of class imbalance problem [40] and feature selection [1]. In the first subsection, all work related to class imbalance is elaborated. In the further subsection, techniques on high-dimensionality used in the literature are discussed.

A. Class imbalance techniques

Class imbalance degrades the performance of classification due to more instances of one type of class [62]. To solve the imbalance problem there exist different approaches which can be categorized mainly into :-1) Data level approaches [51], 2) Algorithm level approaches [50], [14], 3) Ensemble level approaches [8], [57] and 4) Cost sensitive learning [28], [50]. The data level approaches are based on all the pre-processing techniques required for balancing the class distribution [30]. The sampling techniques mainly consist of two types namely: undersampling and oversampling [74]. In under-sampling strategy, the desired number of majority classes are removed randomly which causes the loss of useful information [29]. Some of the most under-sampling techniques used in the literature for balancing are: Wilson's Edited Nearest Neighbor rule [ENN] [23], One Sided Selection [48], Condensed Nearest

2691-4581 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 31,2022 at 08:55:19 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAI.2022.3144651, IEEE Transactions on Artificial Intelligence

Neighbor Rule [56] and Tomek Links [33], [69].
Oversampling is performed by generating the synthetic or duplicate the minority instances. The main risk in this technique lies in the fact that it is vulnerable to over-fitting [10]. To circumvent the probability of over-fitting new technique was introduced for balancing the dataset, SMOTE [9]. SMOTE is the most common technique used in the literature as it can combined with any classifier [10]. The main drawback of SMOTE was over-generalization [40], [52] as the generation of artificial new instances does not consider the nearest neighbor of the majority class, hence, result in class disjuncts and class overlapping [52], [71], [11]. The different variants of SMOTE are described in Table I. Apart from the discussed SMOTE variants, the technique is still filled with abundant research. The work on borderline instances was proposed by Chen et al. [13], where they performed collaboration of clustering and data before pre-processing techniques to determine the samples which always lies in the same cluster (class). Such type of samples is known as center samples. They classify those samples which change class cluster as border samples. So, authors apply SMOTE for minority classes in the border samples. For multi-class imbalanced problem Fernandez et al. [25] proposed to apply decomposition algorithms, one versus all (OAA) [63] and one versus one (OAO) [38]. The study indicated the improved performance of classification. The study [1] focused on data-oriented technique and proposed novel approach based on Mahahlanobis Distance for Oversampling (MDO). This technique uses class mean and covariance, and generates samples for only those minority classes which are at the same distance from the considered class. They also proved that by using this technique risk of overlapping of classes decreases. This technique outperformed when compared with other over-sampling techniques (SMOTE, random oversampling, ADASYN and Borderline SMOTE) using assessment metrics M-measure, F1-score, Precision, Recall and G-measure. But, their proposed model works with low dimensional numeric datasets. The multi-class imbalanced problem was discussed in [77] where the researchers have proposed modified K-nearest neighbor SMOTE algorithm (SMOM). Modification in the technique was to assign weights to all the nearest neighboring minority instances, and instead of randomly selecting the minority instance, instances must be selected according to weight. Thus, the problem of over-generalization was removed by selecting instances according to the weights. Experiment was carried out on 27 real-world datasets with performance measure as MG, MAUC, Precision, Recall and F1-score. Their model works with continuous dataset and not tested on nominal and ordinal attributes.

B. High Dimensionality and Class-imbalance

The problem of high-dimensional features generally leads to decrease the classification performance [47]. The high-dimension features with class-imbalance can cause problems to gain accurate results. Hence, the identification of relevant features can reduce class-overlapping [10]. Many researchers have worked to solve both problems to improve the classification results. In [42], Jiang and other researchers proposed a novel technique by introducing the use of GA with SMOTE technique known as GASMOTE. It overcomes the disadvantage of re-sampling of minority class by reconsidering different rates for sampling of different instances taken, and found the optimal values of sampling rate to combine the results. They compared the results in terms of F-measure and G-mean of ten datasets with classical SMOTE and Borderline SMOTE. GASMOTE proved to produce the best results among the other algorithms. Their proposed model calculated the results on low-dimensional datasets. Other method to solve imbalance problem is the use of Ensemble techniques namely KNORA-Union (KNU) [64]. The study combines the Ensemble technique(dynamic selection of classifier) with preprocessing techniques used are SMOTE, Random Balance [20] and RAMO [12].
The preprocessing techniques used by researchers are oversampling technique. They performed experiment on 26 multi-class datasets and compare their results with static Ensembles. The conclusion of their study was that RAMO is the best preprocessing technique when used with dynamic selection. Another novel approach for multi-class imbalance problem was solved by proposing new algorithm in [5] known as Diversified Error correcting output codes (DECOC). The main technique used in this algorithm was combining Error correcting output codes (ECOC) with Ensemble technique. The ECOC is used to balance the imbalance data. They have used 19 datasets and calculated the results with performance measures: G-mean, Area Under Curve, F-measure and Accuracy. They have compared results with 17 state-of-the-art algorithms and proved that DECOC gave best performance in all measures. Their model computation takes lot of time to get the results. A novel approach using ensemble technique with the use of differential evolution for optimizing weights was proposed in [76] to solve the imbalanced problem. When the results of 12 datasets were compared on the basis of G-mean with vote-based Ensemble technique and non-Ensemble technique, the proposed model proved to be the best.
The study in [54] proposed a model which uses different distance measures with SMOTE algorithm for high dimensional datasets known as Synthetic minority oversampling technique with Minkowski distance (SMOTE-SF). They proposed a new distance metric, Minkowski distance [44] to determine the neighbor of minority class. The various distance measure used are: Chebyshev, Euclidean and Manhattan distances. These distances are ranked with measures namely: Mutual Information, Eigenvector, Fisher Score, Correlation Score and Centrality. The main problem model faced was of over generalization, as researchers ignored the majority class while constructing the synthetic samples. The studies in literature does not solve the problem of lack of density caused due to SMOTE, hence proposed technique is the main endowment to further studies.

III. Theoretical Foundations

A. Classical Biogeography Based Optimization

Biogeography based optimization (BBO) [65] is a population-based nature inspired algorithm which focused on solution of a problem of global optimization. The working

2691-4581 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 31,2022 at 08:55:19 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAI.2022.3144651, IEEE Transactions on Artificial Intelligence

TABLE I: SMOTE variants

Variant	Characteristics	Pros	Cons
Borderline-SMOTE (SMOTE-B)[36]	It creates the synthetic samples from the minority class that are closed to the borderline between two classes	Does not create sample for noise instances	Evade some important instances in minority class
Safe-Level SMOTE [6]	It defines a 'safe level' for each minority sample and create new instances closer to this safe level by providing weights to the new samples	It assures the synthetic instances are not located close to the positive instance which has a lower safelevel value but close to the one with a higher safelevel value	Prone to overfitting
Adaptive Synthetic Sampling Approach (ADASYN) [39]	It increases the weight of newly added samples	The selection process is based on the minority class distribution	It does not identify noisy instances

of algorithm depends on the distribution of species. Each "Habitat" (candidate solution) act as an individual species and every species have its Habitat Suitability Index (HSI), which indicates the degree of goodness. High HSI indicates good solutions for species and low HSI indicates poor solutions. Information sharing among species is possible with migration of species from one island to the other island. Migration of species and HSI values depends on some desirable factors like temperature, rainfall etc. known as Suitability Index Variables (SIV). Immigration of species and emigration of species which resides on different islands are important factors for migration, evolving two operators, immigration operator and emigration operator. Emigration increases when suitability on island increases and hence raising HSI. High HSI indicates good solutions for species and low HSI indicates poor solutions. In BBO, mutation and migration [65], [31] of species is calculated using SIV vector. By considering the values of SIV, optimal solution is obtained. The immigration rate λ, which signifies the living of species on other island and emigration rate μ signifies the leaving of species from native island are the objective functions.

B. Modified Biogegraphy based Optimization (M_BBO)

Biogeography based optimization (BBO) [65] is a population-based nature inspired algorithm which focused on solution of a problem of global optimization. In the literature, BBO is used as feature selection technique in [47], [46], [45], and proved as best feature selection technique. Hence, in our proposed model, we chose to improve mutation operator in BBO and named technique as modified BBO (M_BBO). M_BBO is used for feature selection to reduce the high dimensionality and hence, improved classification performance. For K-th candidate solution x_k, λ_k is the immigration rate. After performing random selection with roulette wheel and independent variable candidate (which is chosen) has to be replaced, then the probability of emigrating candidate solution (selected) is proportional to the emigration rate, μ_k, which is performed using roulette wheel selection, and expressed in Equation 1.

$$Prob(x_i)\textit{selected for emigration} = \mu_i / \sum_{k=1}^{A} \mu_k \qquad (1)$$

where, i=1,2...A, A is presented as the number of candidate solutions in the population.

We have proposed modification in mutation operator of BBO for the improvement in selection of candidate's solution, and migration operator works similar as in classical BBO. Migration uses probability among species. Due to some low suitability factors on islands like lack of resources, species tend to migrate to habitat with low HSI. The migration process modifies the habitat of the island. The immigration probability is denoted by H_t which is directly proportional to λ, the immigration rate. The probability of emigration habitat is H_k, where $H_k \propto \mu$, emigration rate. Pseudocode for migration is formalized in Fig 1

To determine the Habitat (H_t) with the probability $\propto \lambda$
if H_t *is chosen* **then**
  if $rand.random(0.0,1.0) < \lambda$ **then**
    for $k \leftarrow 1$ **to** N **do**
      Determine the Habitat (H_k) with the probability $\propto \mu$)
      if *(H_k) is chosen* **then**
        if $rand.random(0.0, 1.0) < \mu$ **then**

$$H_k(SIV) \leftarrow H_t(SIV) \qquad (2)$$

        end
      end
    end
  end
end

Fig. 1: BBO Migration

2691-4581 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 31,2022 at 08:55:19 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAI.2022.3144651, IEEE Transactions on Artificial Intelligence

In classical mutation process, BBO calculates the probability of individual to create variations in population. This operator is less effective in terms of difference between actual habitat and resultant habitat. This operator is not frequent and intensifies the solution of search space. In this operator, one habitat is chosen based on mutation probability. Afterward, one index is chosen randomly of this habitat, and put one new SIV/element between [0, 2^n] (where n is the total number of sequences) in place of the previous element. The SIVs of the ith habitat H_i can be randomly modified by the mutation operator according to the habitat's priori probability P_i. The mutation probability m_i of the ith habitat H_i is expressed in Equation 3.

$$m_i = m_{max} \times (1 - P_i/P_{max}) \tag{3}$$

where m_{max} is a user-defined parameter and P_{max} = max(P_i) and i = 1, 2... n. In modified BBO mutation operator, in place of random generation, the SIV feature set is ranked using recursive feature elimination (RFE) method and the lowest rank SIV is mutated with higher rank SIV. RFE [73] is the effective method for recognising the higher rank and lower rank. To revamp the search space and handle the outliers, so that exploration can be controlled, the deciding variable adapts the boundary rank value. It can be lower rank value or higher rank value. This modified mutation process helps in achieving higher fitness for habitat overall. The proposed Mean Ranking method is expressed in Equation 4:

$$SIV_l'(j) = (SIV_l(j) + SIV_H(j))/2 \tag{4}$$

where, SIV_l = updated SIV with the lowest rank for jth individual solution, SIV_H = feature with highest weightage, SIV_l' = new feature formed with weighted mean of SIV_l and SIV_H. Mutation follows the pseudocode described in Fig 2:- In this paper, we have proposed and implemented M_BBO as

for $y = 1$ *to* m **do**
  The parameters λ and μ are used to calculate probability P_i
  Select Habitat H_t(y)(SIV) with the probability $\propto P_i$
  if $H_t(y) == selected$ **then**
    Apply Recursive feature elimination to rank features of SIV

$$H_t(y)(SIV_l') = \frac{H_t(y)(SIV_l) + H_t(y)(SIV_h)}{2} \tag{5}$$

  end
end

Fig. 2: Modified BBO mutation operator

feature selection algorithm. We have considered parameters of an algorithm mentioned in Table II :-

TABLE II: Parameters for BBO feature selection

Parameters	Symbols
Number of habitats, habitat[]	T
The feature set	SIV[]
Maximum number of iterations	K
Immigration	λ[]
Emigration	μ[]
Maximum immigration	λ = 500
Maximum emigration	μ=600
Array to store crucial/redundant SIV factor	gb[]

The working of modified BBO (M_BBO) algorithm for feature selection is explained in Fig. 3:-

C. Proposed Distributed SMOTE-D_SMOTE

The original SMOTE technique works by selecting examples that are close in the feature space, drawing a line between the examples in the feature space and drawing a new sample at a point along that line. The need to remove the problem of lack of data by oversampling the minority class instances by SMOTE method leads to the problem of small disjuncts and badly distributed data (with lack of density already a problem). To overcome this problem, we propose a modified distributed SMOTE (D_SMOTE) technique, wherein the synthetic data is generated outside the line joining two present minority class instances. This helps in better distribution of minority class data point and also overcomes the problem of small disjuncts, which in turn, improves the performance of the classifier. The classical SMOTE plots the synthetic sample using K-NN algorithm in between the two minority class instances which causes less instance spatial exploration, hence causing lack of density problem. The other similar technique of oversampling, known as Density-based Synthetic Minority Over-sampling Technique (DBSMOTE) [7] works on creating the artificial sample near the cluster define by DBSCAN. Their synthetic samples are densely populated near centroid and sparse at far locations. Hence, their technique suffered from small disjunct problem, and our proposed technique overcame this limitation. So, in D_SMOTE synthetic point is generated outwards from the seed samples using Equation 6 and Equation 7:

$$X_{new} = X_{lower} - (X_{higher} - X_{lower}) \times \delta \tag{6}$$

$$X_{new}{}' = X_{higher} + (X_{higher} - X_{lower}) \times \delta \tag{7}$$

where, X_{higher} is upper coordinates, X_{lower} is lower coordinates of chosen minority samples, δ is a random number between [0,1]. X_new and $X_{new}{}'$ are two new coordinates respectively on both the sides of chosen minority samples. For the better distribution of synthetic points we plot the points at varied angle. Consider, X_new = $(x_1, x_2, x_3....., x_r, ..x_n)$ where X_new is feature vector in n-dimensions, n corresponds to number of independent features/dimensions. To plot the new synthetic data point of minority class in a different direction to its parent data points, we randomly select a direction for the feature say x_r = X_new in direction r (r$\leqslant$n).
Hence, new synthetic features are:-

$$X = x_1, x_2, x_3....., x'_r, ..x_n \tag{8}$$

2691-4581 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 31,2022 at 08:55:19 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAI.2022.3144651, IEEE Transactions on Artificial Intelligence

Input: habitat[], λ, μ
Output: Optimized feature Set
1. Initialize parameters
2. Generate initial population
2.1 Randomize each habitat population (habitat [ ]): H_1, H_2,...,H_T
2.2 Link each habitat index to SIV values(feature)
2.3 λ [ ],initial immigration rate of each habitat
2.4μ[ ],initial emigration rate of each habitat
2.5 Assign index to habitat associated to a feasible solution
3. Calculate fitness of each habitat (solution)
3.1 Calculate using classification rate with new subset of features.
3.2 Compute each habitat HSI value: hab_index [ ]
3.3 Rank habitats using calculated HSI value.
4. Set iteration variable x: depicts the maximum number of iterations and is gradually increased.
5. **while** $x < T$ **do**
Calculate emigration rate (μ_j) and immigration rate (λ_i) for each habitat. Here, j = 1,2,...T and i = 1,2,...T
H_j is selected where selection criterion based on emigration rate (μ_j).
H_i is selected where selection criterion based on immigration rate (λ_i)
Perform migration operation
Perform mutation operation on H_i
Produce new population by replacing previous (old) H_i with new H_i
Re-calculate habitat values, compute corresponding HSI values
Increment x
end
6. Evolve fittest habitats based on threshold value.
7. Perform ranking second time with respect to population
// SIV value that causes the change in hab_index [], is a crucial feature (removed redundant features),
for $x < N$ **do**
if *SIV[x]==crucial value* **then**
gb[x]=1;
else gb[x]=0 ;
end
end
8. Based on gb[]=1, keep crucial feature and remove redundant features
9. Dataset with reduced features is stored in SIV[]

Fig. 3: M_BBO Feature Selection

where, x'_r = $x_r.\cos\theta$, and $\cos\theta \neq 1$, where, θ is random angle between [0°,360°] and r varies from 1 to n.
Hence, with the help of above equations we oversample the synthetic points far from the cluster centroid for better learning. The problem of lack of density of instances that needs to be trained and small disjuncts created due to oversampling is removed as uniform distribution is focused for locating synthetic samples. Fig. 4(a) represents the instances of the classes of dataset and Fig. 4(b) exhibit zoomed image of some instances. Fig. 5(a) is showing that how classical SMOTE algorithm works and generates synthetic samples. It is prominent in Fig. 5(b) that synthetic samples of minority class lie in a single line, shown in red colour. Fig.5(c) shows the synthetic samples that are plotted for different instances for the complete dataset. It can be noted that small clusters are formed near to each other minority samples, which is a disadvantage of SMOTE algorithm.
Our proposed algorithm is represented in Fig. 5(d) where synthetic samples to be generated are located outside the two co-ordinates of minority samples, shown by dotted line. The two samples, that are considered, are shown by a solid blue line and the new minority sample that is plotted is shown by a dotted blue line. Fig. 5(e) shows the generated minority samples and it can be clearly seen that better distribution will lower the number of small clusters or small disjuncts. The new samples are plotted on the opposite side of considered minority samples using different coordinates, to overcome the problem of lack of density and small disjuncts by increasing the exploration area and not leading to form clusters near minority samples. Hence, representation in Fig. 5(f) shows the elimination of small disjuncts in the complete dataset. The representation in figures shows working of D_SMOTE on binary classes, similarly it works for minority instances in multiple classes. Algorithm for the proposed D_SMOTE is formalized in Fig. 6.

IV. Experimental Setup

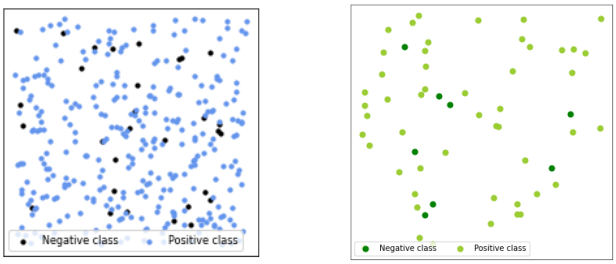
The proposed method is analyzed and compared on 17 multi-class datasets. The five techniques were compared for validating the working of the proposed models namely: No random oversampling (No ROS), SMOTE, BBO+SMOTE, M_BBO+SMOTE and M_BBO+D_SMOTE on four datasets. The four classifiers used for classification are:NB, RF, LR and SVM. The three performance measures namely; AUC, G-mean and F1-score are used to analyze the classification performance. The 10-fold cross validation and 50 runs are performed on each model. After the analysis of results, our approach outperformed the other techniques and can be used for balancing the dataset with feature selection. Ranks are calculated using statistical Friedman test on different performance measures, and our proposed technique ranked highest in all techniques.

A. Data Sets and Classifiers

The 17 multi-class dataset that are chosen for the experiment and analysis is based on the survey. Datasets are chosen from Openml [1], KEEL [2] and UCI [27]. Complete

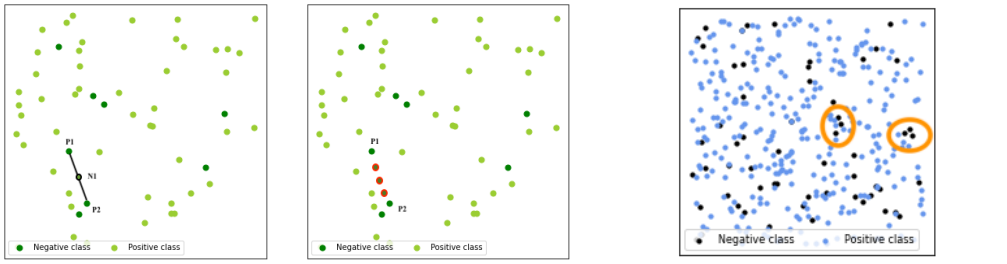
2691-4581 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 31,2022 at 08:55:19 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAI.2022.3144651, IEEE Transactions on Artificial Intelligence

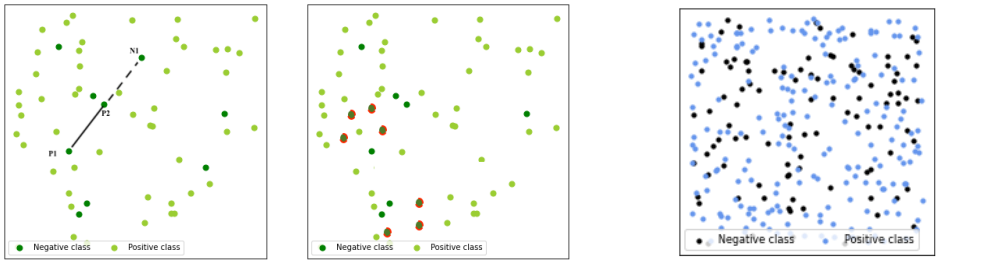


(a) Classes with all instances (b) Zoomed image of instances

Fig. 4: Figure of instances



(a) Classical SMOTE (b) Synthetic samples (c) Synthetic samples of all instances



(d) Proposed D_SMOTE (e) Synthetic samples (f) Synthetic samples for all instances

Fig. 5: Comparison of classical SMOTE with Distributed SMOTE

description of dataset is summarized in Table III. The 10-fold cross-validation is used to carry out the experiment on each dataset for the validation of results. We collected diversified dataset varying classes from 3 to 26, features ranging from 8 to 1586 and instances from 122 to 20,000. The Imbalanced Ratio(IR) is calculated as the ratio of total number of majority class instances to the total number of minority class instances. The ratio varies between 1.57 to 6.99. The brief description of our four high dimensional datasets is as follows:

SRBCT- This is a microarray dataset named as small round blue cell tumors. It mainly consist of four classes named as neuroblastoma (NB), Ewing's family of tumours (EWS), Burkitt's lymphoma (BL), and Rhabdomyosarcoma (RMS). Total number of attributes are 1586 with 122 instances.

Burczynski- Gene Expression omnibus is a super dataset of Burczynski. There are 22,823 attributes with 127 number of samples. It mainly consist of three classes normal, crohn's diseases and ulcerative crisis.

Glioma- This dataset is a textual reports of glioma cancer incidence for mortality analysis. It consist of comprehensive study of PLCO dataset. Nearly 4434 attributes are extracted and 80 instances are considered.

Bullinger- This dataset is a constituent of gene expression of 116 de novo Acute Myeloid Leukemia. It contains high number of features numbered 17,404. We have considered 4 type of classes of leukemia(monocytoid, general leukemia, myeloid and promyelocytic.)

We have shorten the name of dataset IIF_intensity_all_features_data_set_index—files as IIF_intensity in this paper. We have used four different classifiers for our model which are implemented in Python library scikit-learn [60]. The classifiers used for experimentation are: NB [54], RF [53], LR [21] and SVM [54]. All these classifiers are known to produce best results among other classifiers. NB classifier needs less training and converge quickly like LR. RF classifier estimate the importance of features in a reliable manner. SVM can clearly maintains the separation between the classes. SVM works better with high dimensional datasets.

B. Assessment measures

For the analysis of performance of imbalanced multi-class dataset, widely used performance measure is AUC [37], G-Mean [67] and F1-score [1]. Precision(P) [29] and Recall(R) [40] are used as a measure for binary class imbalance problem. Precision (P) and Recall (R) is the measure of exactness and completeness respectively [1]. As we have proposed

2691-4581 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 31,2022 at 08:55:19 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAI.2022.3144651, IEEE Transactions on Artificial Intelligence

D_SMOTE(N,d,k)
Begin
N - Number of minority class instances
d - Amount of SMOTE-synthetic samples
/* Condition : N+2d is less than the number of majority class samples/*
k - Number of nearest neighbors
/* The amount of oversampling depends on parameter d (SMOTE-synthetic samples) */
attrnum = Number of attributes
arrnn = To store k nearest neighbors of each minority class instance
Sample: Matrix of minority class instances coordinates
Synlower: Matrix to store synthesized minority class instances coordinates for lower point
Synhigher: Matrix to store synthesized minority class instances coordinates for higher point
for *i= 1 to T* **do**
  Calculate k nearest neighbor of i^{th} minority class instance and store their indices in arrnn.
  Populate [N,i,arrnn,k]
end
for *synindex = 1 to d* **do**
  rand = random index between 1 and k
  /* Choosing a random neighbor from arrnn nearest neigbor matrix set */
  for *j = 1 to attrnum* **do**
    δ = a random number between [0,1]
    θ = random angle between [0^o , 360^o]
    diff = Sample(arrnn(rand),j) - Sample(i,j)
    Synlower(synindex,j) = Sample(i,j) - δ*diff at angle (θ)
    Synhigher(synindex,j) = Sample(arrnn(rand),j) + δ*diff at angle (θ)
    /* Considering Sample(arrnn(rand),j) to be higher coordinated point than Sample(i,j) such that 2 new points are generated outside these original points */
  end
end
End

Fig. 6: D_SMOTE

model for multi-class imbalance problem, we have measured performance using F1-score, G-mean and AUC. F1-score computes value using Precision and Recall, to balance the trade-off between the measures [29]. The formula is described in Equation 9

$$F1 - score = 2 \times ((R \times P)/(R + P)) \tag{9}$$

G-mean indicates how well a classifier can balance the recognition among different classes. High value of G-mean means that all the classes are considered without the ignorance of a single class. G-mean [67] is calculated as geometric mean of the Recall over all classes. Consider a problem for c-class:

$$G - mean = (\prod_{i=1}^{c} R_i)^{1/c} \tag{10}$$

Receiver Operating Characteristics (ROC) curve is created by plotting True Positive Rate (TPR) and False Positive Rate (FPR). Area Under Curve (AUC) is referred to the area lies under the ROC curve [1].

TABLE III: Description of multi-class dataset

Dataset	class	instances	Attributes	IR
SRBCT	4	122	1586	6.55
Burczynski	3	127	22,823	3.88
Glioma	4	80	4434	6.14
Bullinger	4	170	17,404	11.25
Faults	7	1941	27	14.05
Abalone	18	4139	10	45.93
wine-quality-red	6	1599	11	68.1
Breast tissue	6	1006	9	1.57
Glass	6	214	10	8.44
yeast	10	1484	8	28.1
letter	26	20000	16	1.11
Plates_faults7	7	1941	27	61.11
housing10	10	506	13	152.35
IIF_intensity	3	600	57	4.31
gas_batch1	6	445	128	2.19
Penbased_10an_nn	10	10,992	16	1.95
dataset32pendigits	10	10,992	16	1.16

V. Results and Analysis

A. Performance of Distributed SMOTE with Modified BBO

This section analyses and investigates the result of the proposed approach. The results are compared with other approaches as well as the other models proposed by different authors. The comparison and analysis increase the validation of results for the future use by improving the efficiency of text classification. It can be seen through results that proposed D_SMOTE approach over-samples the minority class samples without modifying the originality of minority class. Particularly, in multi-class cases the minority class regions can be wrongly classified into majority class regions. D_SMOTE is performing oversampling along with increasing the density area of minority data instances. In other words, D_SMOTE overcomes the problem of class disjuncts and lack of density by increasing the area to be explored for minority samples. The more instances can be explored and trained, and hence generalization of the results for optimized classification can be achieved. The comparison of D_SMOTE with other six oversampling techniques proved the efficacy of our technique. D_SMOTE performed best in three datasets compared with other techniques and in fourth dataset it performed nearly equal to the best performance by ADASYN. The use of feature selection technique, M_BBO, enhances the classification performance. To analyze the performance of our model, M_BBO+D_SMOTE, with other models, we perform comparison on three performance measures. We have also performed comparison with the state-of-the-art techniques proposed by other authors on seventeen datasets, for AUC

2691-4581 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 31,2022 at 08:55:19 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAI.2022.3144651, IEEE Transactions on Artificial Intelligence

performance measure, using SVM classifier only as we find the best results with the SVM classifier for our approach in most of the cases. Our model outperformed other techniques for fourteen datasets. The calculated ranks of different samples using Friedman test also signifies statistically that proposed model produces the best results among the other approaches.

1) Comparison with SMOTE variations--Table IV shows the comparison with other techniques used to handle imbalance problem. The comparison is made using AUC scores for different variants of oversampling techniques exist in literature Random-over sampling [71], SMOTE [9], SMOTE-B [36], ADASYN [39] and SMOTE-SL [6]. No RS signifies the scores for no random sampling and ROS signifies the score for random oversampling. Our proposed technique D_SMOTE performed best in three datasets. ADASYN performed best in Bullinger dataset with 91.8%. The rise with 8% to 25% is seen when compared to best values of different variants. In some cases, SMOTE-B and ADASYN performed nearly good to our technique. Due to space constarint, we have used short form of our datasets as SRBCT (S), Burczynski (B), Glioma (G) and Bullinger (Bu). The following observations are analyzed:-

1) In SRBCT, Burczynski and Glioma dataset, our proposed approach of D_SMOTE performed best when compared to other techniques with scores 93.4%, 90.3% and 93.9% respectively. For Bullinger dataset ADASYN performed best with 91.8% and D_SMOTE also performed with 91.6%.

2) The SMOTE-SL performed worst in Burczynski dataset with 45.4% score. ROS performed worst in Bullinger dataset with 60% and in Glioma dataset ADASYN technique performed worst with 50.9%. SMOTE also performed poor for Glioma dataset. No RS performed worst for SRBCT dataset with 46.3%.

3) In Glioma dataset the performance of ROS and SMOTE-SL performed good with scores 91.6% and 86.9%. The performance of ROS with LR classifier achieved best score when compared among other techniques.

4) In Bullinger dataset our approach produces equal score with SMOTE combined with SVM classifier. RF classifier performed worst with our technique for this dataset when compared to other techniques.

Our technique D_SMOTE did not work satisfactorily for three datasets with NB classifier and for two datasets with RF classifier. It is due to the presence of noise instances and overlapping of classes, which lead to a decrease in the classification scores. D_SMOTE works best with NB with one dataset which have medium imbalanced ratio. The other two classifiers LR and SVM performed best for all the datasets with D_SMOTE. Hence, D_SMOTE is best for all types of datasets when combined with LR and SVM.

2) AUC Scores- In Table V, comparison of AUC measures is performed for all datasets with other techniques that can be applied to improve the text classification performance. AUC is the correct measure of classes separability and works on probability. The efficiency of a model in terms of classification or separability of classes is measured by AUC. Higher values of AUC signifies that the model classifies the different classes more accurately. From these values, we can analyze that our approach of M_BBO+D_SMOTE has outperformed over all other techniques. SVM performed best for our model for all the datasets. SVM classifier follows kernel trick which increases the feature space size which cannot be performed by other classifiers. SVM generally performed well with small and medium sized feature space. The better density distribution produced by the proposed D_SMOTE is clearly fabricating best results for the datasets in Table V namely: SRBCT, Bullinger, Burczynski and Glioma. M_BBO+D_SMOTE is best in all datasets, hence it generalizes to use the approach for balancing in multi-class datasets.

The other observations made from Table V are:

1) In SRBCT dataset, AUC best scores for NB classifier is 85.6%, while with SVM it is 96.2%, with LR is 89.4% and with RF the best scores obtained is 83.7%. The SVM score rise around 17.4% when compared to the worst score. The dataset performed worst with BBO+SMOTE model with three classifiers and with NB, M_BBO+SMOTE is producing lowest results.

2) In Burczynski dataset, the best scores are attained by our technique combined with SVM classifier with score 99.9% which is 17.2% more than the worst score. Our technique M_BBO+D_SMOTE are better than all the other techniques with SVM classifier. In this dataset, the worst score with NB classifier is attained by BBO+D_SMOTE with 76.5%. Rest all three classifiers produced worst score with BBO+SMOTE.

3) In Glioma dataset, the best score of 100 is attained by our proposed technique with SVM classifier, rise with 8.8% when compared to worst score. With LR classifier the best scores attained by our technique with 92.2% and also M_BBO+SMOTE performed well with 92%. In this dataset, BBO+SMOTE performed worst with all classifiers.

4) In Bullinger dataset, the best score is attained by our approach with SVM classifier with score 99.10%, with rise of 7.2% when compared to worst score. Our model outperformed with all the classifiers. The worst score with NB and RF classifiers was produced by BBO+D_SMOTE, and rest two classifiers that is LR and SVM performed worst with BBO+SMOTE.

Our proposed model M_BBO+D_SMOTE performed best in all the datasets with all the classifiers. It is analyzed that BBO+D_SMOTE did not performed well in two datasets with NB and RF in one dataset. Although, we have improved D_SMOTE such that it can understand the dataset more precisely. The classification performance degraded due to the problem of poor exploitation that exists in BBO. The assumption of independent predictor features of NB doesn't combine with BBO+D_SMOTE.

3) G-mean scores- In Table VI, G-mean for all techniques is calculated for optimal text classification. G-mean indicates how well a classifier can balance the recognition among different classes. High value of G-mean means that all the classes are considered without the ignorance of a single class. We can clearly conclude from the table that our proposed model outperforms among other techniques with difference

2691-4581 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 31,2022 at 08:55:19 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAI.2022.3144651, IEEE Transactions on Artificial Intelligence

TABLE IV: AUC score for various oversampling methods

D	No RS				ROS				SMOTE				SMOTE-B				ADASYN				SMOTE-SL				D_SMOTE			
	NB	RF	LR	SVM	NB	RF	LR	SVM	NB	RF	LR	SVM	NB	RF	LR	SVM	NB	RF	LR	SVM	NB	RF	LR	SVM	NB	RF	LR	SVM
S	71.7	73.5	46.3	73.6	79.5	59.9	52.5	78.4	54.0	70.7	59.1	73.0	73.5	67.6	64.4	52.1	72.3	69.2	51.6	78.7	76.5	47.6	65.8	77.4	72.2	75.5	76.3	93.4
B	62.4	69.2	59.7	80.3	72.2	63.3	50.2	78.8	59.1	65.1	53.3	81.5	67.4	72.8	78.9	82.2	62.6	53.3	60.1	75.6	60.1	58.2	45.4	74.5	76.1	83.8	88.6	90.3
G	69.3	84.9	80.6	89.1	90.0	70.6	73.8	91.6	52.6	66.2	68.5	88.7	85.1	75.3	80.8	88.6	82.4	50.9	81.5	79.9	70.1	69.3	77.7	86.9	82.6	70.4	82.8	93.9
Bu	75.1	70.8	69.4	88.7	83.4	60.0	75.1	81.9	79.6	74.1	72.5	90.8	63.2	69.5	65.3	87.4	81.6	61.7	69.1	91.8	79.4	79.2	61.3	83.2	70.8	65.6	75.4	91.6

TABLE V: AUC score

Dataset	BBO+SMOTE				BBO+D_SMOTE				M_BBO+SMOTE				M_BBO+D_SMOTE			
	NB	RF	LR	SVM	NB	RF	LR	SVM	NB	RF	LR	SVM	NB	RF	LR	SVM
SRBCT	79.9	68.4	77.3	78.8	81.6	78.2	81.0	86.7	75.1	79.9	84.5	88.8	85.6	83.7	89.4	96.2
Burczynski	78.7	76.6	79.1	82.7	76.5	82.6	86.4	91	85.4	83.8	87.4	90.9	88.2	97.5	94.3	99.9
Glioma	75.5	74.1	89.4	91.2	76.0	75.8	91.3	93.2	80.6	79.6	92.0	95.5	96.7	86.8	92.2	100
Bullinger	76.2	75.5	71.0	91.9	71.4	73.1	76.5	92.1	74.7	77.3	84.1	93.3	89.9	88.6	92.5	99.10

of 5% to 30%. The value of G-mean signifies the balance measure, and calculated values signifies that it is correctly predicting the majority and minority class. It can be seen that the best values for most of the datasets are obtained using SVM classifier. The difference between the G-mean values and AUC values lies in the fact that AUC is a performance metric that equals to the probability that a classifier will rank a randomly chosen positive instance higher than a randomly chosen negative one. G-mean is a performance metric that combines true negative rate and true positive rate at 'one' specific threshold - where both the errors are considered equal. Hence, G-mean works on absolute values.

The various observation made from VI:

1) In SRBCT dataset, the best scores are obtained by M_BBO+D_SMOTE technique. For NB classifier score achieved is 86.4, while with SVM is 95.9, with LR it is 91.1 and with RF it is 82.9%. BBO+D_SMOTE with SVM also performed good with 91.5%. The SVM score with our technique rises around 15% when compared to the worst score. The dataset performed worst with BBO+SMOTE model.

2) In Burczynski dataset the best scores are attained by our technique combined with SVM classifier with score 97.7% which is 17.2% more than the worst score. Our technique M_BBO+D_SMOTE are better than all other techniques. In this dataset, the worst scores attained with BBO+SMOTE with two classifiers, NB and SVM with 74.1% and 80.5% respectively. The two classifiers RF and LR performed worst with BBO+D_SMOTE with 62.4% and 72.8% respectively.

3) In Glioma and Bullinger dataset, our proposed technique best with SVM classifier with 98.6% and 97.5% respectively. The worst scores are produced with BBO+SMOTE technique in both the datasets with all classifiers, except in Bullinger dataset with LR classifier. M_BBO+SMOTE performed worst for LR with 62.3%.

Our model performs best with all the classifiers in all the datasets. The best performance is achieved by SVM, although the other classifiers also performed well. M_BBO+SMOTE does not work well for one dataset with LR classifier which is due to the lack of influential outliers. The other reason for degraded performance is the presence of small disjuncts and noise instances. Also, BBO+D_SMOTE does not perform well with RF and LR for Burczynski.

4) F1-score values- The values are shown in Table VII. It can be clearly stated that higher value of F1-score shows the balancing of dataset, and our proposed model is generating highest values when compared to other approaches. The value of F1-score gives the information of the balance between the Precision and Recall. So, the value of Precision or Recall affects the value of F1-score, hence any zero value will generate zero. Fairly balanced dataset will generate higher F1-score results and hence our model, which is balancing the multi-class dataset is outperforming other techniques.

The observations made from Table VII:

1) In SRBCT dataset, our proposed approach performed best with all classifiers. The best score produced is with SVM classifier with 94.2%, with rise of 14.7% compared to worst score. The worst scores are produced by BBO+SMOTE with all classifiers.

2) In Burczynski dataset, the best score is attained is 96.5% with our approach combined with SVM classifier. The worst score of 42.3% for RF is obtained by BBO+D_SMOTE. The other three classifiers performed worst with BBO+SMOTE.

3) In Glioma dataset, our approach performed best with 99.8% score with SVM classifier, and also performed better with all other classifiers. M_BBO+SMOTE produces good results with LR classifier and worst result with NB classifier. Although NB classifier performed worst with BBO+D_SMOTE. All other three classifiers performed worst with BBO+SMOTE.

4) In Bullinger dataset, the best scores are produced by our approach with all classifiers. The best score with SVM classifier is 97.1%, which is 37.0% value higher than the worst value which is attained by BBO+SMOTE. All classifiers performed worst with BBO+SMOTE.

BBO+D_SMOTE is not recommended with NB and RF classifiers, due to the mixed classification performance values. It is producing good results with LR and SVM classifiers. Our proposed approach can be applied with all the classifiers and to all kind of datasets varying from low dimensional to high dimensional, and low imbalanced to highly imbalanced. The comparison and analysis for AUC measure with other models proposed by authors namely; KNU [64], SMOTE-SF [54], MDO [1], SMOM [77] and DECOC [5], validate our model as it performed better in fourteen datasets. All results in Table VIII, are generalized by considering the average value of 50 runs. By seeing towards the results, our model outperformed with values in the range from 4% to 25%.

2691-4581 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 31,2022 at 08:55:19 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAI.2022.3144651, IEEE Transactions on Artificial Intelligence

TABLE VI: G-mean score

Dataset	BBO+SMOTE				BBO+D_SMOTE				M_BBO+SMOTE				M_BBO+D_SMOTE			
	NB	RF	LR	SVM	NB	RF	LR	SVM	NB	RF	LR	SVM	NB	RF	LR	SVM
SRBCT	63.4	65.6	54.5	80.9	73.3	72.2	64.3	91.5	83.6	70.6	72.5	75.2	86.4	82.9	91.1	95.9
Burczynski	74.1	66.5	74.6	80.5	76.1	62.4	72.8	83.6	78.8	75.5	86.4	89.9	90.2	91.6	95.5	97.7
Glioma	65.4	73.3	79.3	86.2	72.5	80.8	84.7	90.2	68.1	76.4	85.1	92.7	96.1	88.5	93.4	98.6
Bullinger	61.2	55.8	66.3	52.7	69.6	63.0	76.1	65.4	72.5	65.1	62.3	78.6	89.3	88.6	92.3	97.5

TABLE VII: F1-score values

Dataset	BBO+SMOTE				BBO+D_SMOTE				M_BBO+SMOTE				M_BBO+D_SMOTE			
	NB	RF	LR	SVM	NB	RF	LR	SVM	NB	RF	LR	SVM	NB	RF	LR	SVM
SRBCT	65.2	63.3	69.1	79.5	70.1	72.5	69.6	82.5	82.4	78.4	82.1	86.7	88.6	83.7	89.4	94.2
Burczynski	68.4	56.6	59.6	63.4	69.2	42.3	60.5	75.7	80.6	85.7	82.1	85.8	82.1	94.5	94.3	96.5
Glioma	64.5	65.1	58.3	65.7	60.1	68.6	77.4	74.8	64.9	81.9	91.3	93.2	98.3	86.8	92.2	99.8
Bullinger	69.6	68.5	52	60.1	79.6	74.1	72.9	63.6	85.1	88.3	62.2	89.10	90.9	88.6	92.5	97.10

TABLE VIII: Comparison of performance measure (AUC X 100) with different approaches

Dataset	KNU	SMOTE-SF	MDO	SMOM	DECOC	our approach
SRBCT	70.1	76.1	74.2	82.1	71.3	93.5
Burczynski	62.1	82.1	72.1	71.3	69.2	97.2
Glioma	78.5	89.1	69.3	78.6	74.4	99.4
Bullinger	82.3	92.5	79.10	65.9	72.3	98.3
Faults	81.2	78.3	75.4	82.5	71	90.1
Abalone	79.2	85.5	93.9	86.6	82.1	98.7
wine-quality-red	85.6	91.4	95.3	85.2	90.3	98.5
Breast tissue	86.3	83.7	92.5	80.1	91.5	90.4
Glass	89.2	92.3	93.1	82.3	86.3	96.3
yeast	85.6	81.3	92	90.1	86.3	98.6
letter	93.5	90.3	91.4	83.2	87.2	99
Plates_faults	72.4	62.1	59.9	69.9	78.2	92.3
housing10	77.3	86	63.2	76.2	78.4	96.2
IIF_intensity	86.2	73.17	75.2	82.6	85.4	93.1
gas_batch1	81.3	86.3	73.2	96.3	89.2	99
Penbased_10an_nn	75.2	68.2	76.9	83.3	95	92.2
dataset32pendigits	79.8	87.1	86.2	78.1	97.1	90.8

The comparison with BERT [18] for Accuracy measure [47] is shown in Table IX for four datasets. The implementation of BERT is performed using tensorflow. We have considered small datasets to large datasets. We found that our model outperformed in small datasets but performed equally well in large datasets, when compared with BERT. The training of BERT on small datasets results in over-fitting which affects the classification performance. The pre-training and fine-tuning of BERT is considered similar to [18].

TABLE IX: Comparison of performance measure (ACC X 100) with BERT

Dataset	BERT	Our model
SRBCT	93.3	95.6
Glioma	86.1	97.9
Plates_faults7	92.7	96.2
dataset32pendigits	93.8	93.5

B. Statistical Analysis

The statistical test helps in determining whether some given hypothesis is unlikely to occur with an actual observed data. It helps in bridging the gap between numbers and methodologies and forming a quantitative picture of the given process.Friedman test is chosen to compare different techniques and check whether there is any significant difference between them or not. Friedman test [12] is a non-parametric test, and it is preferred over the parametric test as for a parametric test we have to take some data assumptions. It assumes k-different experiment-based different techniques and one variable value is determined multiple times. Each cell value is ranked exclusively, columns symbolize treatments and a row indicates the observed values over the dependent variable and later the ranks of each treatment are added up. The large deviation between the sum value indicates a lower p-value indicating the techniques are different. The number of techniques is 6, therefore the authors decided to go with Friedman test.

Friedman test is a highly effective way to prove the technique score by means of a ranking method. Table X shows the results of Friedman test on AUC for all comparative oversampling approaches. The conclusion was made after analysis that the D_SMOTE over rules all the other techniques. We have computed the results with degree of freedom as 5 which refers to the number of correlated samples minus 1 significant value (α=0.05). The obtained p-value is less than 0.001, that is 5.456957e-07, which signifies the result is significant for the proposed technique. Higher value obtained after applying the test signifies the higher rank of the technique. With the help of rank, we can find out the significant difference between the techniques.

The omnibus p-value is below the respectable critical threshold of 0.05, so post-hoc pairwise multiple comparison tests are conducted to discern which of the pairs have significant differences. These possible post-hoc tests are conducted: the methods of (1) Conover and (2) Nemenyi.

For the (1) Conover method, the p-value is adjusted in two ways, first according to the family-wide error rate (FWER) procedure of Holm, and next by the false discovery rate (FDR) procedure of Benjaminyi-Hochberg. The values indicate that the proposed approach is significantly different than the previously discovered techniques.

The obtained Friedman chi-square statistic is 37.201681. The statistical value of chi-square indicates the deviation over mean ranks of the treatments (different techniques). It's zero signifies that the mean of all treatments is the same, and as this numeric value increases, it indicates the deviation between means of different techniques. We have considered the following hypothesis:-

Technique 1:- Rank of model proposed by KNU [64].
Technique 2:- Rank of model proposed by SMOTE-SF [54].
Technique 3:- Rank of model proposed by MDO [1].
Technique 4:- Rank of model proposed by SMOM [77].

2691-4581 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 31,2022 at 08:55:19 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAI.2022.3144651, IEEE Transactions on Artificial Intelligence

Technique 5:- Rank of model proposed by DECOC [5]..
Technique 6:- The proposed approach.
Null Hypothesis:- There is no significant difference between the techniques proposed by [64], [54], [1], [77], [5] and the proposed technique.
Alternate hypothesis:- There is a significant difference between the algorithms [64], [54], [1], [77], [5] and the proposed technique.
Table X shows the average ranks calculated using Friedman test.
TABLE X: Average rank by Friedman Test results for AUC measure

Techniques	Name	AUC (Ranks)
Technique 1	KNU	1.26
Technique 2	SMOTE-SF	2.38
Technique 3	MDO	2.68
Technique 4	SMOM	3.9
Technique 5	DECOC	4.1
Technique 6	Our approach	5

VI. Conclusion

In this research, the problem of high dimensionality and problems caused due to imbalanced dataset is solved. The technique to handle imbalance problem avoids the problem of small disjuncts and lack of density by creating samples widely. Another limitation of improving classification performance is high number of features, which we handled by modifying mutation operator. The comparison with different variants of oversampling techniques confirmed the model efficiency. The average mean of best values of AUC is 98.8, G-mean is 97.4 and F-score is 96.9, which is satisfactory performance of the proposed model. It was observed that our model performed best with SVM classifier, when compared to other techniques. To validate our model, we compared the proposed approach with other four state-of-the-art algorithms on seventeen datasets. The experiments are carried out on low-dimensional as well as high-dimensional datasets and our model outperformed other techniques for fourteen datasets. Hence, our model can be generalized for low-dimensional as well as high-dimensional datasets. The comparison with the latest classification algorithm BERT on four datasets, where our model outperformed on three datasets and almost performed equally well on fourth dataset. It made our model more promising for improving accuracy with our new techniques. After observation and statistical analysis, it can be said that the classification performance is improved with the proposed feature selection and over-sampling technique. The proposed over-sampling method with a suitable classifier can be used for other applications. Future work will focus on working with other over-sampling techniques and reducing the overlapping problem which our technique does not handle. Other area we can focus is on the optimization of data with other meta-heuristic hybrid algorithms.

References

[1] Lida Abdi and Sattar Hashemi. To combat multi-class imbalanced problems by means of over-sampling techniques. *IEEE transactions on Knowledge and Data Engineering*, 28(1):238–251, 2015.

[2] Jesús Alcalá-Fdez, Alberto Fernández, Julián Luengo, Joaquín Derrac, Salvador García, Luciano Sánchez, and Francisco Herrera. Keel data-mining software tool: data set repository, integration of algorithms and experimental analysis framework. *Journal of Multiple-Valued Logic & Soft Computing*, 17, 2011.

[3] Adnan Amin, Sajid Anwar, Awais Adnan, Muhammad Nawaz, Newton Howard, Junaid Qadir, Ahmad Hawalah, and Amir Hussain. Comparing oversampling techniques to handle the class imbalance problem: A customer churn prediction case study. *IEEE Access*, 4:7940–7957, 2016.

[4] Aniruddha Bhattacharya and Pranab Kumar Chattopadhyay. Solving complex economic load dispatch problems using biogeography-based optimization. *Expert Systems with Applications*, 37(5):3605–3615, 2010.

[5] Jingjun Bi and Chongsheng Zhang. An empirical comparison on state-of-the-art multi-class imbalance learning algorithms and a new diversified ensemble learning scheme. *Knowledge-Based Systems*, 158:81–93, 2018.

[6] Chumphol Bunkhumpornpat, Krung Sinapiromsaran, and Chidchanok Lursinsap. Safe-level-smote: Safe-level-synthetic minority over-sampling technique for handling the class imbalanced problem. In *Pacific-Asia conference on knowledge discovery and data mining*, pages 475–482. Springer, 2009.

[7] Chumphol Bunkhumpornpat, Krung Sinapiromsaran, and Chidchanok Lursinsap. Dbsmote: density-based synthetic minority over-sampling technique. *Applied Intelligence*, 36(3):664–684, 2012.

[8] Cagatay Catal and Suat Guldan. Product review management software based on multiple classifiers. *Iet Software*, 11(3):89–92, 2017.

[9] Nitesh V Chawla, Kevin W Bowyer, Lawrence O Hall, and W Philip Kegelmeyer. Smote: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, 16:321–357, 2002.

[10] Nitesh V Chawla, Nathalie Japkowicz, and Aleksander Kotcz. Special issue on learning from imbalanced data sets. *ACM SIGKDD explorations newsletter*, 6(1):1–6, 2004.

[11] Nitesh V Chawla, Aleksandar Lazarevic, Lawrence O Hall, and Kevin W Bowyer. Smoteboost: Improving prediction of the minority class in boosting. In *European conference on principles of data mining and knowledge discovery*, pages 107–119. Springer, 2003.

[12] Sheng Chen, Haibo He, and Edwardo A Garcia. Ramoboost: ranked minority oversampling in boosting. *IEEE Transactions on Neural Networks*, 21(10):1624–1642, 2010.

[13] Si Chen, Gong-De Guo, and Li-Fei Chen. Clustering ensembles based classification method for imbalanced data sets. *Pattern Recognition and Artificial Intelligence*, 6, 2010.

[14] Fanyong Cheng, Jing Zhang, and Cuihong Wen. Cost-sensitive large margin distribution machine for classification of imbalanced data. *Pattern Recognition Letters*, 80:107–112, 2016.

[15] Behrouz Zamani Dadaneh, Hossein Yeganeh Markid, and Ali Zakerolhosseini. Unsupervised probabilistic feature selection using ant colony optimization. *Expert Systems with Applications*, 53:27–42, 2016.

[16] M Dolores Del Castillo and José Ignacio Serrano. A multistrategy approach for digital text categorization from imbalanced documents. *ACM SIGKDD Explorations Newsletter*, 6(1):70–79, 2004.

[17] Liliya Demidova, Evgeny Nikulchev, and Yu Sokolova. The svm classifier based on the modified particle swarm optimization. *arXiv preprint arXiv:1603.08296*, 2016.

[18] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018.

[19] Diab M Diab and Khalil M El Hindi. Using differential evolution for fine tuning naïve bayesian classifiers and its application for text classification. *Applied Soft Computing*, 54:183–199, 2017.

[20] José F Díez-Pastor, Juan J Rodríguez, César García-Osorio, and Ludmila I Kuncheva. Random balance: ensembles of variable priors classifiers for imbalanced data. *Knowledge-Based Systems*, 85:96–111, 2015.

[21] Georgios Douzas, Fernando Bacao, and Felix Last. Improving imbalanced learning through a heuristic oversampling method based on k-means and smote. *Information Sciences*, 465:1–20, 2018.

[22] Duygu Yilmaz Eroglu and Kemal Kilic. A novel hybrid genetic local search algorithm for feature selection and weighting with an application in strategic decision making in innovation management. *Information Sciences*, 405:18–32, 2017.

[23] Hatem A Fayed and Amir F Atiya. A novel template reduction approach for the k-nearest neighbor method. *IEEE Transactions on Neural Networks*, 20(5):890–896, 2009.

[24] Alberto Fernández, Salvador Garcia, Francisco Herrera, and Nitesh V Chawla. Smote for learning from imbalanced data: progress and chal-

2691-4581 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 31,2022 at 08:55:19 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAI.2022.3144651, IEEE Transactions on Artificial Intelligence

lenges, marking the 15-year anniversary. *Journal of artificial intelligence research*, 61:863–905, 2018.

[25] Alberto FernáNdez, Victoria LóPez, Mikel Galar, MaríA José Del Jesus, and Francisco Herrera. Analysing the classification of imbalanced data-sets with multiple classes: Binarization techniques and ad-hoc approaches. *Knowledge-based systems*, 42:97–110, 2013.

[26] Simon Fong, Raymond Wong, and Athanasios Vasilakos. Accelerated pso swarm search feature selection for data stream mining big data. *IEEE transactions on services computing*, (1):1–1, 2016.

[27] Andrew Frank, Arthur Asuncion, et al. Uci machine learning repository, 2010. *URL http://archive. ics. uci. edu/ml*, 15:22, 2011.

[28] Mikel Galar, Alberto Fernandez, Edurne Barrenechea, Humberto Bustince, and Francisco Herrera. A review on ensembles for the class imbalance problem: bagging-, boosting-, and hybrid-based approaches. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 42(4):463–484, 2011.

[29] Salvador García and Francisco Herrera. Evolutionary undersampling for classification with imbalanced datasets: Proposals and taxonomy. *Evolutionary computation*, 17(3):275–306, 2009.

[30] Vicente García, José Salvador Sánchez, and Ramón Alberto Mollineda. On the effectiveness of preprocessing methods when dealing with different levels of class imbalance. *Knowledge-Based Systems*, 25(1):13–21, 2012.

[31] Wenyin Gong, Zhihua Cai, and Charles X Ling. De/bbo: a hybrid differential evolution with biogeography-based optimization for global numerical optimization. *Soft Computing*, 15(4):645–665, 2010.

[32] Santiago González-Carvajal and Eduardo C Garrido-Merchán. Comparing bert against traditional machine learning text classification. *arXiv preprint arXiv:2005.13012*, 2020.

[33] K Gowda and G Krishna. The condensed nearest neighbor rule using the concept of mutual nearest neighborhood (corresp.). *IEEE Transactions on Information Theory*, 25(4):488–490, 1979.

[34] Javier Grande, María del Rosario Suárez, and José Ramón Villar. A feature selection method using a fuzzy mutual information measure. In *Innovations in Hybrid Intelligent Systems*, pages 56–63. Springer, 2007.

[35] Guo Haixiang, Li Yijing, Jennifer Shang, Gu Mingyun, Huang Yuanyue, and Gong Bing. Learning from class-imbalanced data: Review of methods and applications. *Expert Systems with Applications*, 73:220–239, 2017.

[36] Hui Han, Wen-Yuan Wang, and Bing-Huan Mao. Borderline-smote: a new over-sampling method in imbalanced data sets learning. In *International conference on intelligent computing*, pages 878–887. Springer, 2005.

[37] David J Hand and Robert J Till. A simple generalisation of the area under the roc curve for multiple class classification problems. *Machine learning*, 45(2):171–186, 2001.

[38] Trevor Hastie and Robert Tibshirani. Classification by pairwise coupling. In *Advances in neural information processing systems*, pages 507–513, 1998.

[39] Haibo He, Yang Bai, Edwardo A Garcia, and Shutao Li. Adasyn: Adaptive synthetic sampling approach for imbalanced learning. In *2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*, pages 1322–1328. IEEE, 2008.

[40] Haibo He and Edwardo A Garcia. Learning from imbalanced data. *IEEE Transactions on Knowledge & Data Engineering*, (9):1263–1284, 2008.

[41] T Ryan Hoens, Qi Qian, Nitesh V Chawla, and Zhi-Hua Zhou. Building decision trees for the multi-class imbalance problem. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, pages 122–134. Springer, 2012.

[42] Kun Jiang, Jing Lu, and Kuiliang Xia. A novel algorithm for imbalance data classification based on genetic algorithm improved smote. *Arabian journal for science and engineering*, 41(8):3255–3266, 2016.

[43] Mingyang Jiang, Yanchun Liang, Xiaoyue Feng, Xiaojing Fan, Zhili Pei, Yu Xue, and Renchu Guan. Text classification based on deep belief network and softmax regression. *Neural Computing and Applications*, 29(1):61–70, 2018.

[44] P John. Van de geer,"some aspects of minkowski distance", department of data theory, leiden university. Technical report, RR-95-03.

[45] M Karnan and K Thangavel. Automatic detection of the breast border and nipple position on digital mammograms using genetic algorithm for asymmetry approach to detection of microcalcifications. *Computer methods and programs in biomedicine*, 87(1):12–20, 2007.

[46] Baljit Singh Khehra and Amar Partap Singh Pharwaha. Comparison of genetic algorithm, particle swarm optimization and biogeography-based optimization for feature selection to classify clusters of microcalcifications. *Journal of The Institution of Engineers (India): Series B*, 98(2):189–202, 2017.

[47] Anshu Khurana and Om Prakash Verma. Novel approach with nature-inspired and ensemble techniques for optimal text classification. *Multimedia Tools and Applications*, pages 1–28, 2020.

[48] Miroslav Kubat, Stan Matwin, et al. Addressing the curse of imbalanced training sets: one-sided selection. In *Icml*, volume 97, pages 179–186. Nashville, USA, 1997.

[49] Liang Liu, Yudong Cai, Wencong Lu, Kaiyan Feng, Chunrong Peng, and Bing Niu. Prediction of protein–protein interactions based on pseaa composition and hybrid feature selection. *Biochemical and biophysical research communications*, 380(2):318–322, 2009.

[50] Yi Liu, Yanzhen Wang, Xiaoguang Ren, Hao Zhou, and Xingchun Diao. A classification method based on feature selection for imbalanced data. *IEEE Access*, 2019.

[51] Victoria LóPez, Alberto FernáNdez, MaríA José Del Jesus, and Francisco Herrera. A hierarchical genetic fuzzy system based on genetic programming for addressing classification with highly imbalanced and borderline data-sets. *Knowledge-Based Systems*, 38:85–104, 2013.

[52] Victoria López, Alberto Fernández, Salvador García, Vasile Palade, and Francisco Herrera. An insight into classification with imbalanced data: Empirical results and current trends on using data intrinsic characteristics. *Information sciences*, 250:113–141, 2013.

[53] Li Ma and Suohai Fan. Cure-smote algorithm and hybrid algorithm for feature selection and parameter optimization based on random forests. *BMC bioinformatics*, 18(1):169, 2017.

[54] Sebastián Maldonado, Julio López, and Carla Vairetti. An alternative smote oversampling strategy for high-dimensional datasets. *Applied Soft Computing*, 76:380–389, 2019.

[55] Parham Moradi and Mozhgan Gholampour. A hybrid particle swarm optimization for feature subset selection by integrating a novel local search strategy. *Applied Soft Computing*, 43:117–130, 2016.

[56] Heinrich Niemann and R Goppert. An efficient branch-and-bound nearest neighbour classifier. *Pattern Recognition Letters*, 7(2):67–72, 1988.

[57] Aytuğ Onan, Serdar Korukoğlu, and Hasan Bulut. Ensemble of keyword extraction methods and classifiers in text classification. *Expert Systems with Applications*, 57:232–247, 2016.

[58] V Panchal, Harish Kundra, and Amanpreet Kaur. An integrated approach to biogeography based optimization with case based reasoning for retrieving groundwater possibility. *International Journal of Computer Applications*, 1(8):975–8887, 2010.

[59] VK Panchal, Parminder Singh, Navdeep Kaur, and Harish Kundra. Biogeography based satellite image classification. *arXiv preprint arXiv:0912.1009*, 2009.

[60] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al. Scikit-learn: Machine learning in python. *Journal of machine learning research*, 12(Oct):2825–2830, 2011.

[61] Clifton Phua, Damminda Alahakoon, and Vincent Lee. Minority report in fraud detection: classification of skewed data. *Acm sigkdd explorations newsletter*, 6(1):50–59, 2004.

[62] Ronaldo C Prati, Gustavo EAPA Batista, and Diego F Silva. Class imbalance revisited: a new experimental setup to assess the performance of treatment methods. *Knowledge and Information Systems*, 45(1):247–270, 2015.

[63] Ryan Rifkin and Aldebaro Klautau. In defense of one-vs-all classification. *Journal of machine learning research*, 5(Jan):101–141, 2004.

[64] Anandarup Roy, Rafael MO Cruz, Robert Sabourin, and George DC Cavalcanti. A study on combining dynamic selection and data preprocessing for imbalance learning. *Neurocomputing*, 286:179–192, 2018.

[65] Dan Simon. Biogeography-based optimization. *IEEE transactions on evolutionary computation*, 12(6):702–713, 2008.

[66] Dan Simon. A probabilistic analysis of a simplified biogeography-based optimization algorithm. *Evolutionary computation*, 19(2):167–188, 2011.

[67] Yanmin Sun, Mohamed S Kamel, and Yang Wang. Boosting for learning multiple classes with imbalanced class distribution. In *Sixth International Conference on Data Mining (ICDM'06)*, pages 592–602. IEEE, 2006.

[68] Bo Tang, Haibo He, Paul M Baggenstoss, and Steven Kay. A bayesian classification approach using class-specific features for text categorization. *IEEE Transactions on Knowledge and Data Engineering*, 28(6):1602–1606, 2016.

[69] Ivan Tomek. Two modifications of cnn. *IEEE Trans. Systems, Man and Cybernetics*, 6:769–772, 1976.

2691-4581 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 31,2022 at 08:55:19 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAI.2022.3144651, IEEE Transactions on Artificial Intelligence

[70] Alper Kursat Uysal and Serkan Gunal. Text classification using genetic algorithm oriented latent semantic features. *Expert Systems with Applications*, 41(13):5938–5947, 2014.
[71] Benjamin X Wang and Nathalie Japkowicz. Imbalanced data set learning with synthetic samples. In *Proc. IRIS Machine Learning Workshop*, volume 19, 2004.
[72] Hui Wang, Xinyu Zhou, Hui Sun, Xiang Yu, Jia Zhao, Hai Zhang, and Laizhong Cui. Firefly algorithm with adaptive control parameters. *Soft computing*, 21(17):5091–5102, 2017.
[73] Ke Yan and David Zhang. Feature selection and analysis on correlated gas sensor data with recursive feature elimination. *Sensors and Actuators B: Chemical*, 212:353–363, 2015.
[74] Bee Wah Yap, Khatijahhusna Abd Rani, Hezlin Aryani Abd Rahman, Simon Fong, Zuraida Khairudin, and Nik Nik Abdullah. An application of oversampling, undersampling, bagging and boosting in handling imbalanced datasets. In *Proceedings of the first international conference on advanced data and information engineering (DaEng-2013)*, pages 13–22. Springer, 2014.
[75] Yong Zhang, Dun-wei Gong, and Jian Cheng. Multi-objective particle swarm optimization approach for cost-based feature selection in classification. *IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB)*, 14(1):64–75, 2017.
[76] Yong Zhang, Bo Liu, Jing Cai, and Suhua Zhang. Ensemble weighted extreme learning machine for imbalanced data classification based on differential evolution. *Neural Computing and Applications*, 28(1):259–267, 2017.
[77] Tuanfei Zhu, Yaping Lin, and Yonghe Liu. Synthetic minority oversampling technique for multiclass imbalance problems. *Pattern Recognition*, 72:327–340, 2017.
[78] Ezgi Zorarpacı and Selma Ayşe Özel. A hybrid approach of differential evolution and artificial bee colony for feature selection. *Expert Systems with Applications*, 62:91–103, 2016.



Anshu Khurana is pursuing her doctoral degree in the Department of Computer Science and Engineering from Delhi Technological University. She is working as TRF in the Department of Computer Science and Engineering since 2017. She completed her M.Tech in Computer Science and Engineering from M.D.University in 2010. She served as a contractual faculty in the Department of Information Technology in Delhi Technological University. Her area of interest lies in the areas of Nature-Based Optimization, Text Classification, Machine Learning and Evolutionary Computing.



Om Prakash Verma received PhD degree from University of Delhi, India. From 1992 to 1998 he was with Department of Electronics & Communication Engineering, Malaviya National Institute of Technology, Jaipur, India. He joined Department of Electronics & Communication Engineering, Delhi Technological University, Delhi in 1998. He is currently serving as Professor in the department.
Dr.Verma has an excellent track record of quality teaching and research. He is the recipient of Premier Research Award for Excellence in Research. He is the author of more than 80 publications in International Journals and Conference Proceedings. He is listed in the list of top 2% Scientist in the world in the study performed by Stanford University. He was the Organizing Chair of IEEE India International Conference on Information Processing, 12-14 August 2016 and General chair, 4th International Conference on Computers and Management 22-23 December 2018. His present research interest includes: Data Engineering, Machine Learning, Evolutionary computing, Image Processing.

2691-4581 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 31,2022 at 08:55:19 UTC from IEEE Xplore. Restrictions apply.

Environmental Science and Pollution Research
https://doi.org/10.1007/s11356-021-18426-x

RESEARCH ARTICLE

Optimum techno-eco performance requisites for vacuum annulus tube collector–assisted double-slope solar desaltification unit integrated modified parabolic concentrator

Ashok Kumar Singh[1,2] · **Samsher Gautam**[1]

Received: 28 September 2021 / Accepted: 27 December 2021
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract

This work presents a novel approach for a double-slope solar desaltification system having parallel array of evacuated annular tube collectors with compound parabolic modified concentrators (DS-SDS-EATC-MCPC) that are observed for eco-design requisites for the optimum performance with environ-economic viabilities. The proposed scheme has been configured to obtain the utmost probable basin's medium hotness as 99.6 °C of having greater depth of water (0.16 m) for the East–West-faced direction of basin peak cover (30°) along with South-oriented evacuated annular tube (30°). The highest circulation pace (thermo syphon) is obtained ~ 55 kg/h. The generalized efficiencies (energy-exergy) of the system are 46.53% and 3.62%, correspondingly. The everyday distillate (16.94 kg) and its production cost (energy $0.007/kWh; exergy $0.013/kWh) at a titular selling cost ($0.07/l) maintains its goodness. The CO_2 mitigates (energy-exergy) and green earned credits are 139.74 and 77.30 tons, and $1396 and $772.24, in that order. The framework outlay of the scheme is quite stumpy at $200.79, and productivity of the model is set up > 100% that shows the scheme as appreciably realistic. The evident distillate at little operating cost, ecological returns, elevated alleviation, and short payoff period makes the arrangement sustainable, viable for reasonable collector areas, and optimum EATC with MCPC as eco-design requisites for the projected scheme.

Keywords Eco-design requisites · Modified compound parabolic · Double-slope solar desaltification · Energy-exergy · Environ-economic

Nomenclature

A_a	Accumulative opening area for MCPC
A_g	Top glaze area (m^2)
A_b	Basin liner area (m^2)
A_{rc}	Annular tube receiver area
C_w, C_f	Water mass specific heat capacity (J/kg K)
d	Interior dia of inner annular tube (m)
E_e	Energetic growth (kWh)
E_x	Exergetic growth (kWh)
F'	Annular tube efficiency aspect
G_r	Grashof number
h_{ba}	Net heat transmission factor from liner to ambiance (W/m^2 K)
h_{bw}	Heat transmission factor (convective) from liner to water (W/m^2 K)
h_{t-g}, h_o	Heat loss factor (convention) from top glass to ambiance (W/m^2 K)
h_{t-wg}	Net heat transmission factor from water surface to top glaze cover (W/m^2 K)
h_{c-wg}	Heat transmission coefficient (convective) of brackish water to top glass cover (W/m^2 K)
h_{e-wg}	Heat transmission factor (evaporative) of brackish water to top glass cover (W/m^2 K)
h_{r-wg}	Heat transmission factor (radiative) from brackish water to top glass cover (W/m^2 K)
h_{r-ev}	Thermal loss factor (radiative) of annular tube along vacuum region (W/m^2 K)
h_{saf}	Overall heat transmission factor from absorber to annular tube water (W/m^2 K)
$I_d(t)$	Solar irradiance (diffused and reflected) (W/m^2)

Responsible Editor: Philippe Garrigues

✉ Ashok Kumar Singh
agashok26@gmail.com

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

[2] Mechanical Engineering Department, Galgotias College of Engineering and Technology, 201306 Greater Noida, UP, India

Published online: 17 January 2022

$I_s(a)$	Solar energy engrossed at annular tube receiver (W/m^2)
$I_s(t)$	Radiative astral energy (instantaneous) at SDS unit (W/m^2)
$I_b(t)$	Irradiated solar beam radiation over EATC collectors (W/m^2)
K_g	Glass thermal conductivity (W/m K)
K_b	Basin liner thermal conductivity (W/m K)
K_w	Water thermal conductivity (W/m K)
L	Latent heat (water) for vaporization (J/kg)
L_{EATC}	Annular tube length (m)
$\dot{m}_{ewg}$	Yield hourly distillation (kg/h)
$\dot{m}_f$	Thermo syphon flow rate in annular tube (kg/s)
m_w	Quantity of brackish medium in still (kg)
m_f	Quantity of water in EATC (kg)
n_o	Air refractive index
n_g	Glaze refractive index
n_w	Water refractive index
N_u	Nusselt no.
N_{sh}	Sunshine hour numbers
P_r	Prandtl number
p_w	Vapor pressure (partial) at the surface of still's water (Pa, N/m^2)
p_{gi}	Vapor pressure (partial) at of top glaze interior face (Pa, N/m^2)
$\dot{Q}_{N-EATC}$	Rate of heat grow in SDS through paralleled N-EATC-MCPC arrangement (kJ/s)
$\dot{q}_{ewg}$	Heat transmission rate of evaporation from brackish water to glass cover up (kJ/s)
R_a	Rayleigh no.
R_e	Reynold no.
R_g	Glaze reflectivity
R_{i1}	Internal tube interior radius (m)
R_{i2}	External tube interior radius (m)
R_{o2}	External tube exterior radius (m)
R_{sc}	Selective absorber reflectivity
T_a	Atmosphere temperature (°C)
T_b	Temperature of basin liner (°C)
t_b	Liner depth (m)
t	Time (s)
t_g	Glaze thickness (m)
T_{sa}	Selective absorber temperature (°C)
T_f	Annular tube fluid (water medium) temperature (°C)
T_{fo1}	Exit water temperature from independent EATC (°C)
T_{N-fo}	Exit temperature of associative water medium impending through N-EATC (°C)
T_w	SDS basin water mass temperature (°C)
T_{go}	Top glaze temperature of external face (°C)
T_{gi}	Top glaze temperature of interior face (°C)
U_{saa}	Net heat transmission (loss) factor from external annular tube
V	Air velocity (average) (m/s

Abbreviations

ALN	Aluminum nitride
CCPC	Combined parabolic collector
CPC	Parabolic collector
DS	Slope
EATC	(Vacuum) annular tube collector
FP	Plate
MCPC	Combined parabolic collector (modified)
OCPC	Combined parabolic collector (oriented)
SDS	Solar desaltification system

Greek letters

ν	Kinematic viscosity water (m^2/s)
β'	Water volumetric thermal expansion coefficient (K^{-1})
α	Absorber absorptivity
μ	Dynamic viscosity of water ($N{\cdot}s/m^2$)
α_g	Glass absorptivity
α_w	Water absorptivity
α_b	Liner absorptivity
η_i	Instantaneous efficiency (%)
η_e	Energy efficiency (%)
η_x	Exergy efficiency (%)
τ	Transmissivity
σ	Constant of Stefan-Boltzmann ($W/m^2\ K^4$)
ρ	Density (kg/m^3)
ε_{eff}	Emissivity (effective)
ε_g	Glaze emissivity
ε_w	Water emissivity

Introduction

Despite the endless availability of solar energy, the need for eco-design requisites still remains. This is mainly due to the fact that solar still system is an economical and self-sustainable option (Tiwari et al. 1992). Although solar energy has become more prevalent, its heat loss, i.e., ~ 60% is still significant (Malik et al. 1982). This is the reason why it is necessary to improve the efficiency of the system through modified designs, additional components, etc. The advancements in this area are presented by Singh et al. in 2020 and show the improved results of implementing the various components of a solar desaltification system. Furthermore, the use of evacuated annular tube collectors (EATC) in solar desaltification systems (SDS) represent a superior technique to improve the overall performance of the distiller. Sato et al. (2012) discovered the annular tube to be a superior option in that thermal syphon for its inherent cycle of heated fluid medium with homogeneous inter amalgamation and

transmission of thermal energy to hot–cold streams of fluid medium. SDS incorporation with EATC has become one of the better options because EATC has minimum heat losses and does not require sun ray tracking due to its cylindrical form. SDS-EATC's setup is far less expensive, easier to transport, install, run, and maintain, and can also withstand bad weather surroundings (Kalogirou 2003; Morrison et al. 1984). Then a set of 21 annular tubes was examined for heat gain and assessed respective performance variables that determine the circulatory pace of fluid in evacuated annular tube for the management of a flat mirror reflector at a 45° angle positioned underneath the EATC (Morrison et al. 2005). The simulated results indicated a favorable appearance in the best interest of thermo syphon circulation of annular tube water, which not only has adequate flow to move the engrossed solar heat, but moreover creates a tumultuous consequence in basin, which disrupts scaling, and stratification, as well as other deposition in the framework. Then after, EATC was taken to investigate the efficient fluid medium heating under implicated criteria like the sun's energy, hotness of still's water, inclination, and aspect proportion, as well as dedicated observations of flow rate in annular tube under rectangular circuit of thermo syphon framework under the respective solar radiations, and found that the maximum thermo syphon rate attained was 25 kg/h (Budihardjo et al. 2007). Budihardjo and Morrison (2009) went on to expand the previous study similarly having 30 nos. of annular tubes and relate the results to flattened collectors. The findings revealed that optical and heat loss features have a significant effect in EATC's solar thermal heating performance. Koffi et al. (2008) theoretically and experimentally analyzed the thermo syphon method in a solar water heater, with very good accordance of results for ordinal temperature differences between EATC-SDS, which are usually needed for the commencement of thermo syphon work practices. Dev and Tiwari (2012) used a successful combination of EATC and a passive still with an extra heat chamber. The results were 1.9 l/day yield for 1-m^2 operating area at 30.1% energy efficiency with a large number of unmanaged beneath regions of EATCs (24). And, in an experiment with a high number of EATCs (15 nos.) coupled with passively operated SDS, Sampathkumar et al. (2013) observed a yield of 3.9 l a day for a 1-m^2 area for 43% efficiency with the uncontrolled layer of EATC. Then a hypothetical analysis was conducted for passively operated solar still with annular tube system with ten numbers of array and discovered a day yield of 3.8 l/m^2 with 54.4% and 8.25% energetic and exergetic efficiencies, respectively (Singh et al. 2013). Further, the SDS-EATC as an active function was driven for liquid mass movement, instead of fluid flow or thermo syphon operating circumstances, results in a significant drop in the total performance of the system (Kumar et al. 2014). The entire research is focused on Singh et al.'s (2013) prior model with a modification in working principle, namely, thermal syphon method to forceful circulation type. In relation to the past findings of the same design, a negative result was observed, with significant decreases in energetic-exergetic efficiencies, and yields of 37.9%, 68.5%, and 8.7%, respectively. These findings also support and encourage the application of the thermo syphon operating principle in the SDS-EATC system, particularly for water treatment. As a consequence, Yari et al. (2016) investigated SDS-EATC using the same criteria as Singh's model, with a few amendments such as 30 annular tubes against 10 and compelled mode of water flow, and found a disappointing performance (58% decrease in yield), as predicted and faced by Kumar et al. (2014) in their findings. Then, in the favorable atmospheric conditions of West Texas, Issa and Chang (2017) systematically investigated a solar desalination system with EATC (5 nos. of horizontal orientation). And the unused underneath of annular tube and its horizontal position slow the thermo syphon circulation and mingling of hot–cold waterways inside the SDS-EATC. Solar still performs with 19% efficiency and generates 3.6-l yield a day despite the favorable climatic conditions all around it and a reduced number of EATC. Also, a variety of solar still with annular tube setups was studied to purify seawater and found better outcomes of 4.05-l yield for 1-m^2 area per day by 28.23% effectiveness for processing of highly saline water and may increase the entire study's validity by appropriately utilizing EATC to capture radiation from the sun via all over its periphery (Patel et al. 2019). To improve the productivity of SDS schemes, Singh (2020) disclosed several more potential of unique combinations and dexterous integration of diverse materials used in solar-operated desaltification technique. Moreover, many scholars have examined the feasibility of several critical issues and executions of fiscal, environmental, and energetic matrix views in still geometries in recent years. The use of clean power in SDS produces a tiny amount of harmful gases as a consequence into the environment, which can cause a variety of major health difficulties, human disease, farm damage, global warming, heat island effect, and other issues. With these considerations in mind, Tiwari et al. (2015) examined the solar still system's exergy-environment-economics and found a good exergy cost of $6.29/year for $14.5/ton of respective CO_2. Afterwards, Reddy and Sharon (2017) examined vacuum annular movement incorporated with still coordination for environmental and fiscal assertions and discovered a financial basis with emissions of 221.8, 1.6, and 0.7 tons for CO_2, sulfur dioxide, and NO that was praised in pecuniary stipulations based on the international competitive market rate. Correspondingly, the CO_2 mitigates as well as other environ-economic inferences for slanted SDS were analyzed and the results showed conspicuous output (~4.9 kg/day) at a cost of 1.74/l with a 5% consumer price index rate and alleviates 17.65 tons of

carbon expulsion from the scheme for 20-year existence of the system (Sharon et al. 2017). Yousef et al. (2019) analyzed the various passive SDS using various criteria such as exergo-environment-economic, and variations of these, and discovered maximum feasible carbon alleviates ~ 41.6 tons. Furthermore, a research of circular still shows substantially minor yielding charges ($0.04) per liter for 30 years of the system's operating life at a 5% interest rate, and a $4.42 per annum exergo-environmental-economic expense (Bait 2019). The result, such as factors (including interest rate and longevity), play a crucial role in predicting the system's feasibility and long-term viability. Singh and Samsher (2021a) investigated solar desalination still with vacuum annular tubes (10 Nos.) for environmental and energetic matrix observations, and found payout time (0.68 yrs) with better yield output (3.8 kg/m^2 per day), greater mitigates (> 77 tones), and credits ($771.23) for a 30 years life of the system.

The SDS-assisted EATC using thermo syphon operating concept is clearly one of the finest practices in solar desaltification technique, as evidenced by the aforementioned study. It also necessitates a less number of EATC since one element is capable of raising the temperature to a major extent by applying appropriate design and skilled integration of the subsidiary elements. However, there is a gap in the research about the best way to use EATC for absorbing uniform radiation from the sun (Budihardjo and Morrison 2009; Yari et al. 2016; Sadeghi et al. 2021; Barbosa et al. 2020). The preceding EATC lessons merely offer a rectangular thermal syphon framework through solar irradiation, typically with compelled circulation within the large quantity of vacuum annular tubes (Yari et al. 2016; Dubey et al. 2021; Xu et al. 2019) of symmetric connections followed by Xu et al. (2019) as little substantial gains or lowered recitals as depicted by Dwivedi and Tiwari (2008); Rashidi et al. (2018); Yousef and Hassan (2019) and Dubey et al. (2021) and given in Table 3. Also suggested, double-trail astral-controlled desalination model operates under natural thermo syphon cycle into the best possible and small numbered EATCs with totally enclosed periphery using oriented (OCPC) and cusp (CCPC) parabolic concentrators (i.e., MCPC) to maintain a uniform thermo syphon circuit framework. The new adaptive integration of SDS-EATC-MCPC, which offers a consistent control of directed sunlight irradiative heat toward the recipient, enables the system to perform at its full potential in provisions of effectiveness, wealth, lifespan cost alteration, and drinkable water production efficiency to continue serving the living being community cost-effectively that is one way to achieve the above-said operational conditions. Furthermore, the previous research provides a rectangular thermo syphon framework instead of a unified thermo syphon loop design, which is easy with the support of the proposed method, i.e. DS-SDS-EATC-MCPC, which undertakes a conical thermo syphon cycle with an improved flow rate and outcomes well enough in terms of overall performance and drinkable water production efficiency in an environmentally friendly manner.

Need of current work with gap of research

Some researchers (Xu et al. 2019) have demonstrated the usefulness of EATC in SDS units for thermo syphon operation conditions with definite constraints (expensive due to massive quantity of annular collector and elevated amount of energy; low heat exploitation due more unallocated sides of EATCs, low water depth occupies lower sensible heat that results lower nocturnal yield, water pump includes overall cost and maintenance, etc.) that impact badly on environment also. The majority of the flaws are being addressed in the proposed approach of tech-eco configuration of association by the positive and effective initiatives (i.e., lower cost due to optimal EATC; utilization of all sides of EATC through MCPC agreement; enhanced thermo syphon).

As a result of the potential research gap (substantial proportion of EATC, rectangle thermo syphon, unequal use of EATC), the current work corrects it by using MCPC, a pair of novel configuration of parabolic reflectors, i.e., OCPC in combination with CCPC. Furthermore, it has been discovered that no study has been done on the single-basin dual-slope desalination unit including customized compound parabolic converter and evacuated annular tube collectors. The gap in the literature stimulates to push further in this field to the point where one can set a benchmark in solar-powered water desaltification technologies. The study relied on a performance of the proposed scheme (DS-SDS-EATC-MCPC) that is explicitly or implicitly based on potable water productivity, energy, exergetic efficiencies, energetic-exergetic metrics, economic analyses, etc., and also the impact of the system's high-tech ecological footprint. The suggested system's performance is compared to that of previous studies, as well as the effect of various controlling factors. As indicated in Fig. 1, the mentioned objectives are used for the analytical analysis of the proposed unit.

i. Create an innovative technology for producing potable water that meets eco-design criteria.
ii. Thermal modeling of said scheme to assess performance of the system using energy, exergetic, performance, matrices, and various economic studies.

System depiction and requisites

Figure 1 depicts a labeled schematic of a one slope desalination unit with similar *N* numbers of paralleled array EATCs enhanced by a modernized compound parabola concentrator.

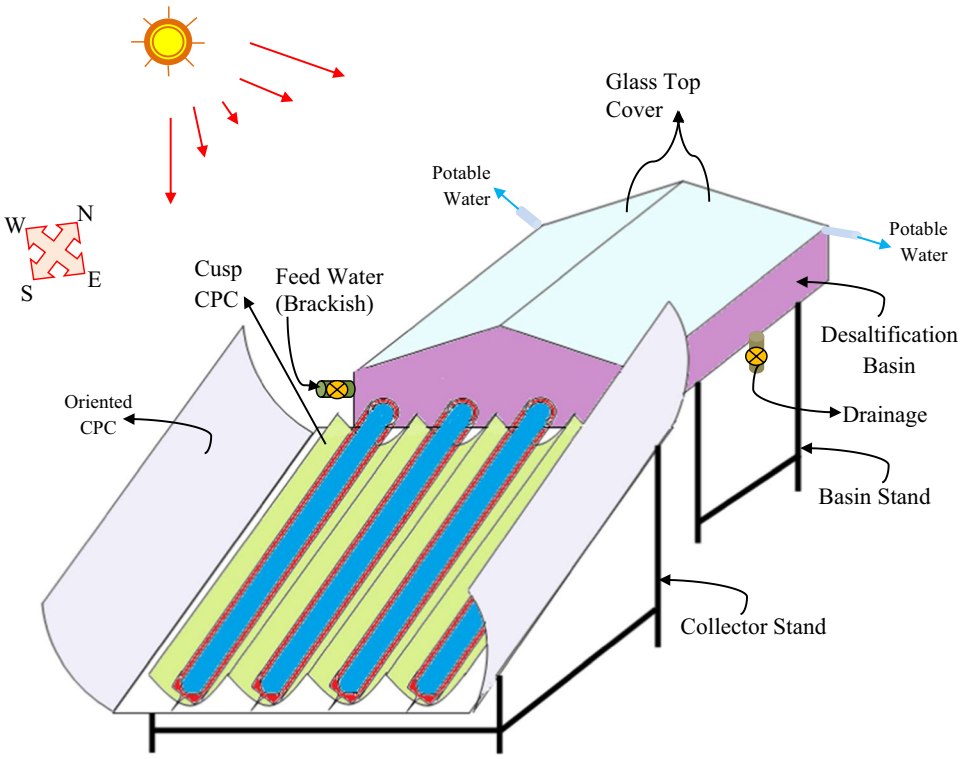


Fig. 1 Graphical representation of suggested model (solar desaltification unit of double slope typed vacuum annular tube collector with combined parabolic modified concentrator (DS-SDU-N-EATC-MCPC))

Table 1 shows the proposed system's specifications, as well as differentiating aspects and parameters connected to its elements of the DS-SDS aided by EATC most likely as depicted by Singh and Samsher (2022). Under passively working circumstances, the SDS unit is constructed of fiber-reinforced polymer (FRP) and connected with EATC-MCPC. The still pool is situated in an East–West direction. The window glass of the DS-SDS chamber has 0.78 W/m K heat conductibility and positioned by 30 degrees due south by Tiwari and Tiwari (2007), and is the ideal viewpoint of these installations (Northern hemisphere) and obtains full energy throughout the day. The SDS pool liner is put beneath the surface water and has a greater absorptive capability (0.97) owing to the matt black dye on top, ensuring that the most solar thermal energy is captured. Furthermore, the basin's inner side walls (immersed region) are covered with black dye up to the water head (0.16 m), which aids in the absorption of solar irradiation over the immersed side walls. The remaining of the wall surface is coated with a white inset paint that bounces the incoming sun's radiation back into the water volume, helping to raise the water temperature at the same time. A key solar collector's module, i.e., evacuated annular tube catcher coupled to the DS-SDS device, is positioned in the parallel pattern of the same nature to increase the quality of the system, as shown in Fig. 2. The annular tube is a capable combination of concentrically unified dual cylindrical glaze ducting (borosilicate 3.3 with an end open and another adiabatically sealed) with preferential absorber coating of 3 layers (copper + steel + ALN) over the outer layer of the interior glaze in a vacuum environment. With a vacuum and specific receiver, the annular dual-glazed circular tube absorbs 95% of energy and preserves the short wavelength sun spectrum with reduced radiation losses. EATCs are spaced 0.314 m apart in the middle and 0.079 m over the CCPC, and the entire configuration is similarly integrated well with DS-SDS framework. Furthermore, the CCPCs are arranged in such a manner that solar radiation falls on each side of the EATC and is adequately reflected back by CCPC to EATC's lower half, as illustrated in Fig. 3. The installation of solar collector has a 0.314 m opening and a leverage ratio of 2.1, as well as an optimum concentration range of 2. Furthermore, no research has been done on the side segmental use of EATC, and this research problem has been filled by utilizing a novel intended seamless set of four equivalent and associative parts of common CPC, which has been established in such a way that its equivalent sector of CPC focuses the incoming solar irradiance toward the side of EATC also. The orient combined parabolic collector (OCPC), placed over each face of the frame for exploiting flanks together of the annular tubes reasonably, is the name given to the integrated segmental CPC alliance. Furthermore, the OCPC-CCPC dimensions have been scrutinized to the point where they cannot obstruct the receipt of focused astral irradiance above the corresponding faces together of annular tubes (Fig. 4). As a result, the suggested OCPC-CCPC combination named MCPC displays a revolutionary

Table 1 Stipulation and computational constraints of the projected novel model (DS-SDS-EATC-MCPC)

Solar desaltification double slope system (DS-SDS)			
Scheme consideration	Measurement	Scheme consideration	Measurement
Still basin length	1 m	Framework substance	Mild steel
Still basin width	1 m	Still basin direction	East–West
East–West height of wall	0.260 m	Water depth	0.16 m
Basin center wall stature	0.394 m	Thickness of basin liner	0.05 m
Top glass area	1.155m^2	Thickness of insulation	0.05 m
Inclination (top glass)	30°	m_w	160 kg
Material of basin	FRP	α_g	0.05
Insulating material	Glass wool	α_b	0.97
Cover material	Glass	α_w	0.34
Cover thickness	0.004 m	n_o	1
σ	5.6697×10^{-8} W/m^2 K^4	n_g	1.5
ε_g	0.9	n_w	1.3
ε_w	0.9	K_g	0.78 W/m K
C	0.54	n	0.25
K_b	0.33 W/m K	$K_{i(\text{glasswool})}$	0.03 W/m K
Evacuated annular tubular collector (EATC)		MCPC: OCPC-CCPC	
Scheme consideration	Measurement	Scheme consideration	Measurement
Length of annular tube	1.8 m	OCPC length	1.8 m
Density of CPC Al-foil	2710 kg/m^3	OCPC aperture (left–right)	0.935 m
Inner annular tube ID	0.047 m	OCPC overall aperture	2.87 m
Inner annular tube OD	0.049 m	OCPC-EATC effective concentration ratio	2.5
Outer annular tube ID	0.056 m	CCPC length	1.8 m
Outer annular tube OD	0.058 m	CCPC aperture opening	0.314 m
Concentric tube spacing (vacuum zone)	0.007 m	CCPC-EATC effective concentration ratio	2
Thickness of glaze	0.002 m	OCPC orientation to EATC	39.89° ~ 40°
Selective absorber area	0.0034 m^2	OCPC angle to horizontal	45°
EATC material	High borosilicate 3.3 glass	CCPC orientation	45°
Center distance	0.314 m	CPC (Al-foil) thickness	0.3 mm
F'	0.968	Reflectivity	0.86
R_{sc}	0.05	CPC material	1060 Al alloy
α	0.95	CPC color	Silver
τ	0.95	CCPC-EATC gap	0.079 m
h_{saf}	100 W/m^2 K	MCPC-EATC effective concentration ratio	2

techno-environ-design related to solar energy systems which accompanied solar-operated desaltification techniques

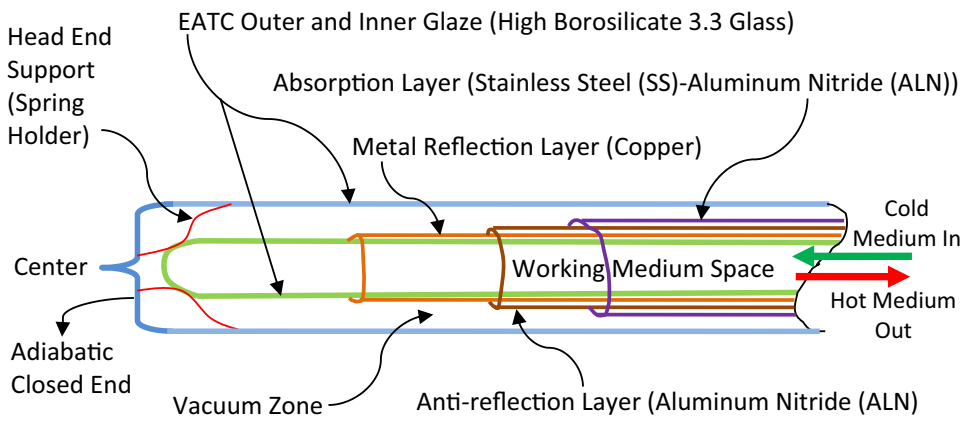


Fig. 2 Line diagram of vacuum annular tube collector (EATC)

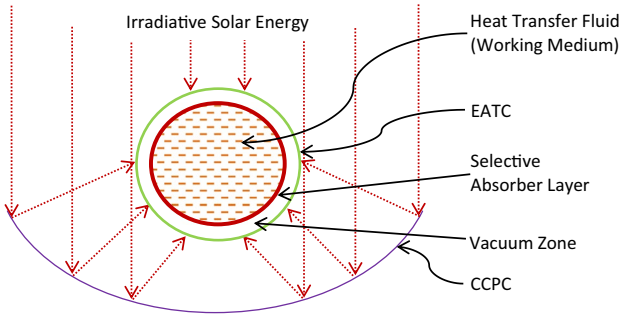


Fig. 3 Elevated representation of EATC-CCPC casing at bottom half of EATC

including effective concentration range of 2 for intended assembly. The MCPC's most significant benefits for total solar heat transmission and related flow rates through EATC to SDS basin are as follows:

i. Owing to CCPC-MCPC, the entire circumferential region of EATC is evenly lighted, and the consistent heat gain via the perimeter of EATC is approximated.
ii. A higher discharge rate is achieved due to descending of the narrowed heat transport boundary between hot stream and cold streamed water as colder water have more gravity, which provides extra push to the heated water, allowing for a greater flow speed of elevated temperature water movement to the watershed. It happens rapidly and homogeneously to heat uptake into the SDS chamber, as well as de-stratification and descaling consequences.
iii. For blending of hot–cold water channels, a circumferential heat exchange into EATC features an axially conical circuit and a cross-sectional annular loop.
iv. The suggested scheme does not require a water motor; it eliminates electrical power and also eliminates the need for continual system supervision, which immediately benefits in monetary terms and additional labor.

MCPC geometrical association

Through the use of CCPC-OCPC, the broad use of EATC to its full capacity has been preserved. As illustrated in Fig. 3, the topological illustration of equivalent vacuum annular tube and CCPC under the impact of sensitive solar irradiance is investigated. In addition, the whole assemblage (Fig. 4) of vacuum annular tube with modified combined parabolic convertor is being constructed with extensive examination of OCPC aligned CCPC. In consideration to the incoming radiations from the sun, the CCPC-OCPC has been properly parabolied to cover the all-round peripheral region of EATC (roughly uniform). The generated CCPC-OCPC expressions are provided in such a fashion that all solar energy between the opening regions is efficiently steered toward EATC (inside the intense tangents) with no solar ray losses, as written:

a) $x = 17.7\sqrt{y}$ (CCPC Profile)
b) $x(x - 2.8) = 4.7y + 4696.3$ (OCPC Profile)

Solar irradiance: east–west-inclined surface

The Indian Meteorological Department, Pune, India, conveniently provides accessible solar irradiation energy on the horizontal surface, and transformed so to obtain the relevant irradiation across the tending plane using the MATLAB analytical tool. To compute the solar incident response across the inclined plane using Liu and Jordan (1960), first obtain the cosine part of the tilt angle (θ_i or East inclination θ_{iE}, and West inclination θ_{iW}) of radiated solar energy as depicted,

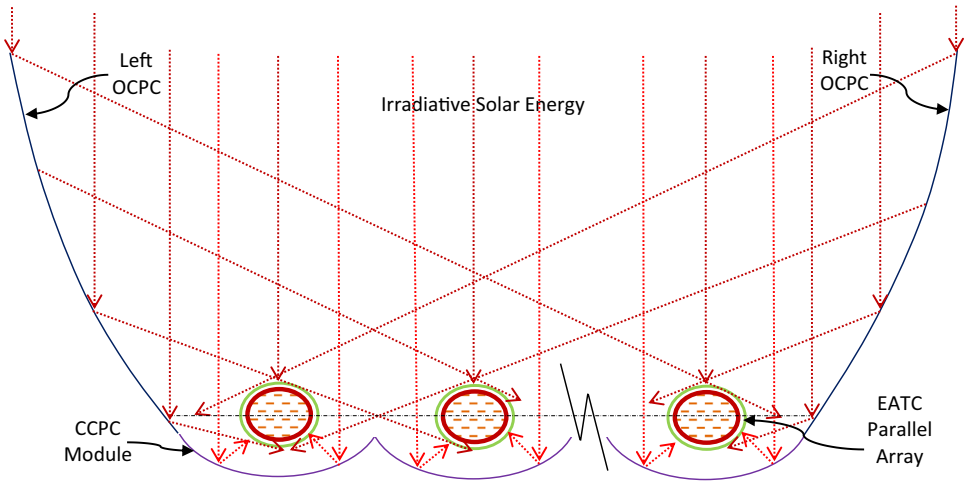


Fig. 4 Elevated representation of N-EATC-MCPC casing all sides of EATC

$$\mathrm{Cos}\theta_i = (\mathrm{Cos}\phi \times \mathrm{Cos}\beta + \mathrm{Sin}\phi \times \mathrm{Sin}\beta \times \mathrm{Cos}\gamma).\mathrm{Cos}\delta \times \mathrm{Cos}\omega + \mathrm{Cos}\delta \times \mathrm{Sin}\omega.\mathrm{Sin}\beta \times \mathrm{Sin}\gamma + \mathrm{Sin}\delta.(\mathrm{Sin}\phi.\mathrm{Cos}\beta - \mathrm{Cos}\phi.\mathrm{Sin}\beta.\mathrm{Cos}\gamma)$$

The current proposal (DS-SDS-EATC-MCPC) is aligned due East–West, so the azimuth degree (γ) and inclined face slope (β) would be − 90° and + 90°, and − 30° and + 30°, correspondingly. For that geometry, the zenith degree (θ_z) could be represented as,

$$\mathrm{Cos}\theta_z = \mathrm{Cos}\phi.\mathrm{Cos}\delta.\mathrm{Cos}\omega + \mathrm{Sin}\delta.\mathrm{Sin}\phi$$

where the angular values, i.e., ϕ, δ, and ω, are referred to as the latitude angle, declination angle, and hour angle, correspondingly. Also, following the relative expression to evaluate declination angle as articulated (Cooper 1969),

$$\delta = 23.45 \times Sin\left[\left(\frac{360}{365}\right).(284 + n)\right]$$

where $\omega(= (ST - 12).15°)$ along with n signifies hour angle in addition to respective day (n) of the time for which the angle is being depicted. ST denotes the solar time and 15° is a multiplication factor that represents the Earth's 15° hourly rotating angle to its axis. Furthermore, by considering the Liu and Jordan (1962), the total sum of equivalent solar irradiance across the inclined plane (East–West oriented) may be stated as,

$$I_s(t) = \left[\rho'.R_r.\left(I_d(t) + I_b(t)\right)\right] + \left[I_d(t).R_d + I_b(t).R_b\right]$$

The requisites, $R_b\,(Cos\theta_i/Cos\theta_z)$; $R_d\,((1 + Cos\beta)/2)$; and $R_r\,((1 - Cos\beta)/2)$, indicate the proportion of astral beam energy above leaning to flat face, proportion of solar diffused energy for slanted to the flat face, and reflected energy multiplier part from the earth, correspondingly.

Thermal model analysis

The thermodynamic equations are supported by the thermodynamic laws in the building of the thermal model for the said present scheme. Furthermore, the created thermal model and associated balanced expressions for the various subsystems have been reorganized to get extended expressions of random variables in the form of defined properties for a certain time interval. During the analysis, the following assumptions were taken into account:

i. The arrangement is airtight, compact, and leak-proof of quasi-steady-state (Badran and Abu-Khader 2007).
ii. Temperature drop is negligible to the component thickness since all of the model's components are properly integrated (Badran and Abu-Khader 2007). The ability of glass, insulator, and sidewalls to absorb heat is insignificant.
iii. The cold-hot water connection interface in the EATC is preferably a conical one, but for ease of computation, it is assumed to be a circular loop form of circular path profile across the EATC length.
iv. The EATC absorber array is uniform, and there is no contact among neighboring tubes dismantled at off-light period. Only homogeneous and laminar blending of hot–cold water flows allows for unidirectional and circumferential heat transfer in the EATC.
v. The average solar exergy fraction and wind velocity have been considered.

Parallel association: EATC-MCPC

When MCPC and EATC are combined, incoming solar radiation is increased over through the EATC. The whole sunlight (beam and focused) strikes the EATC's all-around contoured contact area, and then after expressing a negligibly tiny portion of incoming solar radiation, it travels toward the next location, the preferential absorber lining at the external surface of the internal coaxial glass pipe, at which the highest incident solar radiation is captured and then transferred to the liquid medium of EATC.

Energy balance equation for EATC absorber (outer surface of inner glaze concentric tube)

The selected absorbent (a sandwich of micro-thin Cu-ALN-SS layers) is the core of EATC, where solar irradiation energy is captured and transmitted to the neighboring surrounding fluid with little energy loss to the environment. An energy equation may be used to represent the same thing as,

$$I_s(a) = \left(F' \times h_{saf}\right) \times \left(T_{sa} - T_f\right) + U_{saa}.(T_{sa} - T_a) \quad (1)$$

where U_{saa} is the total heat flow (loss) factor of the EATC's outer glaze tubing, vacuum area, and internal glaze circumferential cannula. Equation (1) may also be adjusted to obtain the temperature of the preferential absorbent (T_{sa}) as follows:

$$T_{sa} = \frac{(\propto \tau)_{eff}.I_b(t) + F'h_{saf}T_f + U_{saa}T_a}{F'h_{saf} + U_{saa}} \quad (2)$$

The Appendix follows the various terms in the preceding equation.

EATC working fluid energy equation

Because the EATC is totally full of water, liquid obtains thermal gain from the selective absorbent concurrently owing to close interaction of cooling water with the highest temperature portion of the EATC (i.e., next to the

preferential absorbent). In addition, the heat balance equation may be written as follows:

$$\dot{m}_f C_f \frac{dT_f}{dx} dx = F^{'}.h_{saf}\left(T_{sa} - T_f\right).\left(2\pi R_{i1}\right)dx \tag{3}$$

Equation (2) is rearranged when T_f was subtracted from both sides, and thus the equation was obtained,

$$\left(h_{saf}.F^{'}\right) \times \left(T_{sa} - T_f\right) = F_1 I_b(t) \times (\propto \tau)_{eff} - U_L \times (T_f - T_a) \tag{4}$$

Using Eqs. (3)–(4) as a preliminary step, the rearranged version may be written as,

$$\dot{m}_f C_f \frac{dT_f}{dx} = \left(2\pi R_{i1}\right)\left[F_1 I_b(t).(\propto \tau)_{eff} - U_L(T_f - T_a)\right] \tag{5}$$

Furthermore, by solving Eq. (5) with the following boundary constraints,

$$T_f = \begin{cases} T_{fi}, x = 0 \\ T_{fo1}, x = L_{EATC} \end{cases} \tag{6}$$

Under the aforesaid boundary constraints, the response of Eq. (5) in the sense of working media for representing hotness of first vacuum annular tube is represented as

$$T_{fo1} = \left[\left(A_{rc}.F_r\right)/\left(\dot{m}_f C_f\right)\right] \times \left[F_1(\propto \tau)_{eff} I_b(t) + U_L T_a\right] + F_2 T_{fi} \tag{7}$$

where the preferential absorbent area (A_{rc}) coating over EATC inner tube is computed using the ($2\pi R_{o1} L_{EATC}$) calculation. Furthermore, due to the same incorporation of EATC to SDS basin that produces a thermal syphon circuit into the corresponding EATC duct, and a similar manner that is happening into the neighboring glass tubes as well, every annular tube will respond accordingly as in Eq. (7). For the instantaneous sun irradiations, the formula for T_{foN} for the N^{th} tube is

$$T_{foN} = \left[\left(A_{rc}.F_r\right)/\left(\dot{m}_f C_f\right)\right] \times \left[F_1(\propto \tau)_{eff} I_b(t) + U_L T_a\right] + F_2 T_{fi} \tag{8}$$

The gross temperature (T_{N-fo}) attained by the corresponding arrangement (N identical EATC with MCPC) terminals may also be estimated and expressed using the equations below.

$$\frac{1}{T_{N-fo}} = \frac{1}{T_{fo1}} + \frac{1}{T_{fo2}} + \ldots \frac{1}{T_{foN}} \tag{9}$$

The proposed parallel array combination has almost the same output water temperature flowing out of the EATC ($T_{fo1} = T_{fo2} = \ldots T_{foN}$). So, using the suggested criterion, Eq. (9) can be evaluated to obtain the gross temperature of liquid-out (T_{N-fo}) from paralleled arrangement of N-EATC-MCPC, so one may obtain,

$$T_{N-fo} = \frac{\left(A_{rc}.F_r\right)}{N.\dot{m}_f C_f}\left\{F_1(\propto \tau)_{eff} I_b(t) + U_L T_a\right\} + F_2 T_{fi} \tag{10}$$

Furthermore, because the paralleled arrangement of N identical EATC with MCPC is linked to still chamber, the still receives an extra quantity of heat intake ($\dot{Q}_{N-EATC}$), which is measured by adding the similar heat inputs as shown for the 1st EATC ($\dot{Q}_1$).

$$\dot{Q}_1 = \dot{m}_f C_f (T_{N-fo} - T_{fi}) \tag{11}$$

Similarly, the heat flux from the Nth EATC tube may be stated as follows:

$$\dot{Q}_N = \dot{m}_f C_f (T_{N-fo} - T_{fi}) \tag{12}$$

After combining the equations and using Eq. (10), the overall amount of radiant heat ($\dot{Q}_{N-EATC}$) into SDS reservoir via the N-EATC-MCPC may be calculated as

$$\dot{Q}_{N-EATC} = N.\dot{m}_f C_f \left[\frac{\left(A_{rc}.F_r\right)}{N.\dot{m}_f C_f}\left\{F_1(\propto \tau)_{eff} I_b(t) + U_L T_a\right\} + F_2 T_{fi} - T_{fi}\right] \tag{13}$$

Furthermore, by rearranging Eq. (13), the net quantity of radiant heat ($\dot{Q}_{N-EATC}$) may be found in its physical context as expressed.

$$\dot{Q}_{N-EATC} = \left[I_b(t) F_1(\propto \tau)_{eff}\left(F_r.A_{rc}\right)\right] - \left[U_L\left(F_r.A_{rc}\right)\left(T_{fi} - T_a\right)\right] - \left[\dot{m}_f C_f.T_{fi}(N-1)\right] \tag{14}$$

Thermal model: DS-SDS unit

The thermodynamic equations comprising East–West top glazing (interior-exterior surfaces), basin fluid, and the underside basin lining of the DS-SDS device, as well as the heat energy provided by N-EATC-MCPC, are included in the thermal model study.

East–West ambient exposed, top cover glaze surface

The East-facing solar radiation strikes the half-top covering, and the attenuated energy reaches the basin after minimum reflection and absorption. The heat that is now collected at the internal glaze area is transmitted outwards and then convected to the environment. As a result, the energy equation may be established as follows:

$$\frac{K_g.A_{gE}}{t_g}\left(T_{giE} - T_{goE}\right) = h_{t-gE} A_{gE}(T_{goE} - T_a) \tag{15}$$

The external surface temperatures (T_{goE}) of the Eastern shore top glaze can also be calculated by rearranging the previous equation.

$$\left(T_{giE}-T_{goE}\right)\times\left(K_g.A_{gE}/t_g\right)=\left(T_{giE}-T_a\right)\times\left(U_{CE}A_{gE}\right) \tag{16}$$

$$T_{goE}=T_{giE}+\frac{t_g}{K_g}.U_{cE}.\left[T_a-T_{giE}\right] \tag{17}$$

Similarly, the external surface equations for the West top glazing may be expressed to determine the outer layer temperature (T_{goW})

$$\left(T_{giW}-T_{goW}\right)\times\left(K_g.A_{gW}/t_g\right)=\left(T_{goW}-T_a\right)\times\left(h_{t-gW}.A_{gW}\right) \tag{18}$$

$$T_{goW}=\left[U_{cW}.\frac{t_g}{K_g}\right]\times\left[T_a-T_{giW}\right]+T_{giW} \tag{19}$$

The parameters t_g, h_{t-gE}, and h_{t-gW} represent the thickness of the upper glaze layer and the net heat flow factor from the East–West inner glazes to the atmosphere, respectively.

East–West top cover glaze condensing surface

The Eastern shore top cover condenser area (inner) receives heat from two sources: first, solar incidence energy, and secondly, water surface heat transfer via convection, evaporates, and radiation. Furthermore, a simultaneous energy loss manifested itself in two ways: first, conductive heat leak from the Eastern edge inner glazing to ambient, and second, radiative heat leak from East side interior glass to the West side interior glass into the basin area. As described, the complete energy transfer process may be expressed in terms of an energy model as mentioned,

$$\begin{aligned}\alpha_g^A.A_{gE}.I_{sE}(t)+\frac{A_b}{2}.\left(T_w-T_{giE}\right).h_{t-wgE}&=\left(T_{giE}-T_{goE}\right)\\&\times\left(\frac{K_g.A_{gE}}{t_g}\right)+A_{gE}.\left(T_{giE}-T_{giW}\right).h_{r-EW}\end{aligned} \tag{20}$$

Similarly, the heat balance equation for the West side covering condensing surface (inner) is as follows:

$$\begin{aligned}\alpha_g^A.A_{gW}I_{sW}(t)+\frac{A_b}{2}.\left(T_w-T_{giW}\right).h_{t-wgW}&=\left(T_{giW}-T_{goW}\right)\\&\times\left(\frac{K_g.A_{gW}}{t_g}\right)+A_{gW}\left(T_{giW}-T_{giE}\right).h_{r-WE}\end{aligned} \tag{21}$$

in which $h_{r-EW}=h_{r-WE}$ is the radiant temperature coefficient for topmost glass surfaces on the East and West sides. The abbreviations h_{t-wgE} and h_{t-wgW} relate to the East–West heat transfer characteristics (radiation, convective, and evaporation) and may be found in the Appendix.

Energy equation: SDS basin water

The basin's water attains thermal energy in a variety of ways, including solar radiation energy (diminished) from the Solar radiation through all the East–West side of the upper glaze, thermal energy to the watershed lining via convective heat flow, and an enormous rate of heat energy directly from the sequential array configuration of the EATC-MCPC via thermo syphon mass flow. The related heat depreciation due to energy transmission (convection, evaporate, and irradiation) from the surface of water to East–West faces of top glaze covers, as well as sensible heat reception into the still's water, may be defined in terms of the following energy equation:

$$\begin{aligned}\alpha_w^A.\frac{A_g}{2}\left[I_{sE}(t)+I_{sW}(t)\right]+A_b.\left(T_b-T_w\right).h_{bw}+\dot{Q}_{N-EATC}&=\left(T_w-T_{giE}\right)\\.\frac{A_b}{2}.h_{\ t-wgE}+\left(T_w-T_{giW}\right).h_{t-wgW}.\frac{A_b}{2}+m_wC_w.\frac{dT_w}{dt}&\end{aligned} \tag{22}$$

The proportional solar flux captured by the SDS basin's mass of water is denoted by the notation α_w^A. Furthermore, the acronym h bw refers to the convective heat flow factor between the basin laminate lining and the water mass, which may be seen in the Appendix. The empirical relationships in Tables 2 and 3 can be used to calculate the numerous terms linked to the thermo-physical characteristics of water that are impacted by temperature.

Energy equation: SDS basin liner

The basin lining is the final stage where the remaining thermal energy is absorbed with concomitant energy loss to the surface water (convention) as well as ambient (marginal), which may be described as follows using the energy equation:

$$\left[I_{sE}(t)+I_{sW}(t)\right]\times\left(\alpha_b^A.\frac{A_b}{2}\right)=\left(T_b-T_w\right)\times\left(h_{bw}A_b\right)+\left(T_b-T_a\right)\times\left(h_{ba}A_b\right) \tag{23}$$

The expression $\alpha_b^A=\left\{\left(1-R_g\right)\left(1-\alpha_g\right)\left(1-R_w\right)\left(1-\alpha_w\right)\alpha_b\right\}$ pertains to the SDS basin laminate through which a partial solar flux was absorbed. The word h_{ba} shows the total heat flow factor from basin laminate to ambiance with conductive-convective-irradiative heat transfer gateways. After rearranging Eq. (23), the corresponding expression for T_b may be derived as

$$T_b=\left[\frac{1}{2}.\alpha_b^A\left\{I_{sE}(t)+I_{sW}(t)\right\}+\left(T_w.h_{bw}+T_a.h_{ba}\right)\right]/\left(h_{bw}+h_{ba}\right) \tag{24}$$

The rearranged version, in which the unknown terms are supplied in the Appendices, may be obtained by subtracting T_w then multiplying h_{bw} along both sides.

$$\left(T_b-T_w\right)\times h_{bw}=\frac{1}{2}.\alpha_{beff}\left\{I_{sE}(t)+I_{sW}(t)\right\}-\left(T_w-T_a\right)\times U_{bwa} \tag{25}$$

T_{giE} may be obtained as described by using Eqs. (15)-(25) and Eq. (16) in Eq. (20).

Table 2 Still basin as well as EATC water thermo-physical properties subjective to temperature deviations expressed in terms of the empirical relations (IAPWS 1996; Pissavi 1982; Koffi et al. 2008; IAPWS 2008)

Constraint	Empirical relation
Specific heat competence (J/kg K)	$C_w = 4226 - 3.244.T + 0.0575.T^2 - 0.0002656T^3$
Water medium density (kg/m^3)	$\rho = 1001 - 0.08832.T - 0.003417.T^2$
Thermal conductivity (W/m K)	$K_w = 0.557 + 0.002198T - 0.00000708T^2$
Volumetric thermal expansion factor (K^{-1})	$\beta' = \left(0.3 + 0.116T - 0.0004T^2\right).10^{-4}$
Kinematic viscosity (m^2/s)	$\nu = \left[(0.5155 + 0.0192T)^{-1} - 0.12\right].10^{-6}$
Dynamic viscosity (N·s/m^2)	$\mu = 4.2844 \times 10^{-5} + \left\{0.517(T + 64.993)^2 - 91.296\right\}^{-1}$
Vaporization Latent heat (J/kg)	$L = 2.506 \times 10^6 - 2.369 \times 10^3.T + 0.2678.T^2 - T^3 \times (8.103 \times 10^{-3})$

$$T_{giE} = \frac{\frac{A_b}{2}.T_w.h_{t-wgE} + A_{gE}.\left[\alpha_g^A.I_{sE}(t) + U_{cE}.T_a + T_{giW}.h_{r-EW}\right]}{\left[\frac{A_b}{2}.h_{t-wgE} + A_{gE}.\left(U_{cE} + h_{r-EW}\right)\right]} \tag{26a}$$

$$T_{giE} = h'_{1E}.\left[\alpha_g^A.I_{sE}(t) + U_{cE}.T_a + T_{giW}.h_{r-EW}\right] + h'_{2E}.T_w \tag{26b}$$

Equation (21) can also be rewritten to represent T_{giW} in the provided form.

$$T_{giW} = \frac{\frac{A_b}{2}.T_w.h_{t-wgW} + A_{gW}.\left[\alpha_g^A.I_{sW}(t) + U_{cW}.T_a + T_{giE}.h_{r-WE}\right]}{\left[\frac{A_b}{2}.h_{t-wgW} + A_{gW}.\left(U_{cW} + h_{r-WE}\right)\right]} \tag{27a}$$

$$T_{giW} = h'_{1W}.\left[\alpha_g^A.I_{sW}(t) + U_{cW}.T_a + T_{giE}.h_{r-WE}\right] + h'_{2W}.T_w \tag{27b}$$

Equations (26)–(27) can also be rewritten in a different way, with the unknown terms as listed in Appendix.

$$\begin{aligned} h_{t-wgE}.\left(T_w - T_{giE}\right) &= \left(T_w - T_a\right) \times U_{taE} + h'_{1E}.h_{r-EW} \\ &.\left(T_w - T_{giW}\right) - h'_{1E}.\alpha_g^A.I_{sE}(t) \end{aligned} \tag{28}$$

$$\begin{aligned} \left(T_w - T_{giW}\right) \times h_{t-wgW} &= U_{taW} \times \left(T_w - T_a\right) + h_{r-WE}.\left(T_w - T_{giW}\right) \\ &.h'_{1W} - h'_{1W}.\alpha_g^A.I_{sW}(t) \end{aligned} \tag{29}$$

After plugging in the values of $\left(T_b - T_w\right).h_{bw}$, $\dot{Q}_{N-EATC}$, $\left(T_w - T_{giE}\right).h_{t-wgE}$, $h_{t-wgW} \times \left(T_w - T_{giW}\right)$, and T_{fi} equivalent to T_w into Eq. (22), the reorganized version of the equation will be:

Table 3 Recital of yielding for diverse solar desaltification schemes in amalgamation with astral thermal receivers

Referenced study	SDS system	Solar coverage (D + L/R + U)*	Avg. yielding (kg/m^2·h)	Operating state
Tiwari et al. 2003	Basin chamber	D	0.12	Conventional and passive
Tiwari et al. 2003	Basin chamber and FP collector	D	0.26	Active and fluid pump
Dev and Tiwari 2012	Basin chamber and EATC (24)	D	0.41	Passive and thermo syphon
Sampathkumar et al. 2013	Basin chamber and annular tube with heat pipe (15)	D	0.54	Active and fluid pump
Singh et al. 2013	Basin chamber and EATC (10)	D + beneath mirror	0.56	Passive and thermo syphon
Kumar et al. 2014	Basin chamber and EATC (10)	D + beneath mirror	0.58	Active and fluid pump
Yari et al. 2016	Basin chamber and EATC (30)	D	0.48	Passive and thermo syphon
Issa and Chang 2017	Inclined basin chamber with EATC (5)	D	0.36	Active and fluid pump
Patel et al. 2019	Stepped SDS and EATC (5)	D	0.41	Passive and thermo syphon
Xu et al. 2019	Tubical still with heat pipe condenser, annular tube (24)	D	0.57	Active and fluid pump
Dubey et al. 2021	Basin chamber, EATC (10)	D + beneath mirror	0.44	Active and fluid pump
Projected System	Basin chamber, EATC, and MCPC (4)	D + L/R + U	0.71	Passive and Thermo syphon

*D direct irradiance, L/R radiation at left–right face, U beneath solar radiation

$$\frac{dT_w}{dt} + a.T_w = f(t) \tag{30}$$

For the undertaken boundary constraints at $t = 0, T_w = T_{wo}$, with $\left\{ C = T_{wo} - \frac{f(t)}{a} \right\}$, one may get the temperature of the water (T_w) by integrating Eq. (30) as

$$T_w = T_{wo}.e^{-at} + \frac{\bar{f}(t)}{a}\left(1 - e^{-at}\right) \tag{30a}$$

In the above expression, $\bar{f}(t)$ is the averaged value for $f(t)$ with respect to the given time period of interpretation.

Fluid flow examination in annular tube with MCPC: thermo syphon loom

The thermo syphon fluid loop flows concentrically through into EATC from basin section of SDS, and actively substitutes the lighter hot stream of water with the denser cold stream. The thermo syphon loop is obtained intensely into EATC due to the radial use of sunlight (straightforward sunlight at the upper edge of the EATC + distributed irradiance from oriented-CPC just at side surface of annular tube + distributed irradiance from cusp-CPC at the right beneath the annular tube) but has not been mentioned by any other researcher. Furthermore, the net thermo syphon and mass flow frequency into the annular tube are influenced by insolation (solar), circular heat allocation covering EATC, and still liquid temperature, most of which are directly related and segmented as a result of the proposed novel associations. Following the aforesaid equations and accompanying relations, one can calculate the thermal performance for unit surface area (q) of moving volume of water with identical EATC.

$$\dot{q} = \dot{Q}_{N-EATC} / \left(N.A_{rc}\right) \tag{31}$$

$$Re = \frac{(\rho.V^2)}{(\mu.V/X)} = \frac{4\dot{m}_f}{\pi d.\mu} \tag{32}$$

Equation (32) can also be rewritten as,

$$\dot{m}_f = 1.5708.Re.\mu.R_{i1} \tag{33}$$

The Reynolds number (Re) for the EATC under thermo syphon circuit under changeable temperature circumstances may alternatively be expressed as follows, according to Budihardjo and Morrison (2009):

$$Re = 0.1914\left\{ \frac{Nu.Gr}{Pr}\left(\frac{L_{EATC}}{2R_{i1}}\right)^{1.2}.Cos\theta \right\}^{0.4084} \tag{34}$$

Equation (33) may now be described in more details by writing the value of Eq. (34) as follows:

$$\dot{m}_f = 0.6643\left\{ \frac{\left(g.\beta'.\rho^2.\dot{q}\right)}{\left(C_f.\mu^{0.55}\right)}.\left(L_{EATC}\right)^{1.2}.{R_{i1}}^{5.25}.Cos\theta \right\}^{0.4084} \tag{35}$$

Finally, using the values of Eqs. (31) and (14) as shown below, the resultant formula for flow rate within annular tube for thermo syphon operating condition may be expressed as

The still water is entering into the EATC; hence, $T_{fi} = T_w$. The water temperature within the EATC, which is a cumulative effect of concentrated sunlight radiation at the upper half face of annular tube radial face, as well as focused sunlight energy at each side (owing to oriented-CPC) and below its half rounded face (owing to cusp-CPC) with same annular tube, is dependent solely on the parameters conveyed according to the above equation. Furthermore, under contemporaneous test circumstances, the resultant equation Eq. (36) of flowrate ($\dot{m}_f$) connects incoming radiation from the sun, EATC alignment angle to perpendicular, and annular tube aspect ratio.

$$\dot{m}_f = 1.52\left[\frac{\left[\left(A_{rc}.F_r\right)\left\{F_1(\propto \tau)_{eff} I_b(t) - U_L\left(T_{fi} - T_a\right)\right\} - \dot{m}_f C_f.T_{fi}(N-1)\right]}{N.A_{rc}}.\left(\frac{L_{EATC}}{d}\right)^{1.2}.\frac{g.\beta'.\rho^2.{R_{i1}}^{6.45}}{C_f.\mu^{0.55}}.Cos\theta \right]^{0.41} \tag{36}$$

Eco-design requisites and performance observations

The eco-design requirements and performance measurements are being carried out for an emblematic sunny day in June in the Indian capital of New Delhi.

Distillate production

The production of distillate ($\dot{m}_{ewg}$) expressed as given below.

$$\dot{m}_{ewg} = \frac{\left\{h_{e-wgE}.\frac{A_b}{2}.(T_w - T_{giE})\right\} + \left\{h_{e-wgW}.\frac{A_b}{2}.(T_w - T_{giW})\right\}}{L} \times 3600 \tag{37}$$

Equation (37) may also be stated for the everyday yield production ($\dot{M}_{ewg}$) and yearly yield production $\left(M_{ewg}\right)_a$ by adding the hourly output for 24 h and the entire year, accordingly.

Energetic analysis

It is possible to calculate the corresponding energy generated by the suggested system, which is directly proportional to the amount of drinkable water produced. The hourly energy output $(E_e)_{\mathrm{hr}}$ is calculated as follows:

$$(E_e)_{\mathrm{hr}} = (\dot{m}_{ewg}.L)/3600 \tag{38}$$

Daily energy output $(E_e)_{\mathrm{daily}}$ and yearly energy output $(E_e)_{\mathrm{annual}}$ may also be calculated by adding the hourly energy (daily and yearly) accordingly. The energy efficiency of suggested system for a specific instance (η_i) may now be stated as

$$\eta_i = \frac{\left[\frac{A_b}{2}.T_w\left(h_{e-wgE}+h_{e-wgW}\right)-\left(h_{e-wgE}.T_{giE}+h_{e-wgW}.T_{giW}\right)\right]}{\frac{A_b}{2}.\left\{I_{sE}(t)+I_{sW}(t)\right\}+I_b(t).A_a}\times 100 \tag{39}$$

Furthermore, the inclusion of heat energy gained as of paralleled arrangement (identical EATC with MCPC) configuration with still unit as stated, the hourly efficiency $\left(\eta_e\right)_{\mathrm{hr}}$ of the present scheme may be represented similarly.

$$\left(\eta_e\right)_{hr} = \frac{(E_e)_{hr}}{\frac{A_b}{2}.\left\{I_{sE}(t)+I_{sW}(t)\right\}+\dot{Q}_{N-EATC}}\times 100 \tag{40}$$

Similarly, the system's day energy efficiency $\left(\eta_e\right)_{\mathrm{daily}}$ and yearly energy efficiency $\left(\eta_e\right)_{\mathrm{annual}}$ may be stated by adding the values for each day and year accordingly.

Exergetic analysis

The response of the system may be analyzed as qualitative basis through the ideas of entropy concept as well as exergy to the unit, as this is a key concept to measure energy in regard to exergetic investigation. The system's exergetic balance equation may be articulated as,

$$(\dot{E}_x)_{\mathrm{in}} = (\dot{E}_x)_{\mathrm{out}} + (\dot{E}_x)_{\mathrm{des}} \tag{41}$$

where $(\dot{E}_x)_{\mathrm{des}}$ denotes the energy that has been destroyed owing to irreversibility or exergetic degradation. Furthermore, the hour exergy $(\dot{E}_x)_{\mathrm{hr-out}}$ and day exergy $(\dot{E}_x)_{\mathrm{daily-out}}$ gain of the suggested scheme have been reported using principles of thermodynamics and formulating the hour exergy gains as (Cengel and Boles 2013),

$$\begin{aligned}(\dot{E}_x)_{\mathrm{hr-out}} = {} & \frac{A_b}{2}\times\left\{\left(T_w-T_{giE}\right)-T_a.ln\left(\frac{T_w}{T_{giE}}\right)\right\} \\ & \times h_{e-wgE}+\frac{A_b}{2}\times\left\{\left(T_w-T_{giW}\right)\times h_{e-wgW}-T_a.ln\left(\frac{T_w}{T_{giW}}\right)\right\}\end{aligned} \tag{42}$$

Now, the gross hour exergetic input $(\dot{E}_x)_{\mathrm{hr-in}}$ to the unit that is a mix of exergetic input directly from solar irradiance $(\dot{E}_{x_{sun}})_{\mathrm{hr-in}}$ and exergetic input from N-EATC-MCPC $\left(\dot{E}_{x_{\mathrm{EATC}}}\right)_{\mathrm{hr-in}}$ may be expressed as (Singh and Samsher 2021a, b, c),

$$(\dot{E}_x)_{\mathrm{hr-in}} = (\dot{E}_{x_{sun}})_{\mathrm{hr-in}} + (\dot{E}_{x_{\mathrm{EATC}}})_{\mathrm{hr-in}} \tag{43}$$

where $(\dot{E}_{x_{sun}})_{\mathrm{hr-in}}$ denotes the quality of solar radiation that is easily available with significant exergy (Petela 1964). Solar desalination devices, on the other hand, are demonstrating their incompetence when it comes to the excessive use of solar exergy owing to low exergetic efficiency (Petela 2010a, b), and also cannot be fully utilized due to the practical difficulties of full conversion with one energy form into the other due to potential of excessive irrevocability and energy obliteration (Petela 2010a, b). (Petela 2010a). Furthermore, the quality of solar, i.e., exergy directed toward the still, is comparatively poor and depicted as (Petela 1964, 2003),

$$(\dot{E}_{x_{sun}})_{\mathrm{hr-in}} = \frac{A_b}{2}.\left\{I_{sE}(t)+I_{sW}(t)\right\}.\left\{1-\frac{4}{3}\left(\frac{T_a}{T_s}\right)+\frac{1}{3}\left(\frac{T_a}{T_s}\right)^4\right\} \tag{44}$$

$$(\dot{E}_{x_{sun}})_{\mathrm{hr-in}} = I_s A \eta_{X\mathrm{sun}} \tag{44a}$$

$$\eta_{X\mathrm{sun}} = 1+\frac{1}{3}\left(\frac{T_a}{T_s}\right)^4-\frac{4}{3}\left(\frac{T_a}{T_s}\right) \tag{44b}$$

T_a refers to the ambient temperature (300 K) and T_s refers to the hotness of solar irradiance, means sun's temperature (6000 K), resulting in $\left\{1-\frac{4}{3}\left(\frac{T_a}{T_s}\right)+\frac{1}{3}\left(\frac{T_a}{T_s}\right)^4\right\}\approx 0.933$. Similarly, exergetic input $(\dot{E}_{x_{\mathrm{EATC}}})_{\mathrm{hr-in}}$ to the SDS unit from a parallel arrangement of N-EATC-MCPC is,

$$(\dot{E}_{x_{EATC}})_{\mathrm{hr-in}} = \dot{m}_f C_f\left\{\left(T_{N-fo}-T_{fi}\right)-T_a.ln\left(\frac{T_{N-fo}}{T_{fi}}\right)\right\} \tag{45}$$

As a result, Eq. (43) may be expressed in its textual form:

$$(\dot{E}_x)_{\mathrm{hr-in}} = 0.933.\frac{A_b}{2}.\left\{I_{sE}(t)+I_{sW}(t)\right\}+\dot{m}_f C_f\left\{\left(T_{N-fo}-T_{fi}\right)-T_a.ln\left(\frac{T_{N-fo}}{T_{fi}}\right)\right\} \tag{46}$$

Furthermore, the everyday exergetic input to SDS basin enclosure $(\dot{E}_x)_{\mathrm{daily-in}}$ may be represented as,

$$(\dot{E}_x)_{\mathrm{daily-in}} = \sum_{t=1}^{t=10}\left[0.933.\frac{A_b}{2}.\left\{I_{sE}(t)+I_{sW}(t)\right\}+(\dot{E}_{x_{\mathrm{EATC}}})_{hr-in}\right] \tag{47}$$

Exergetic efficiency represents the employed performance efficiency from its maximal attainable capacity of

performance. As a result, hourly exergetic efficiency is written as

$$\left(\eta_x\right)_{\mathrm{hr}} = \frac{\left(\dot{E}_x\right)_{\mathrm{hr-out}}}{\left[\left(\dot{E}_{x_{sun}}\right)_{\mathrm{hr-in}} + \left(\dot{E}_{x_{\mathrm{EATC}}}\right)_{\mathrm{hr-in}}\right]} \times 100 \tag{48}$$

Daily exergetic effectiveness is calculated by dividing the everyday exergy yield by the number of sunshine hours (10 h) and is expressed as follows:

period. For better functioning system considerations, the payoff element $(\mathrm{EP})_F$ is set as high as practicable. The system is regarded beneficial if $(\mathrm{EP})_F \to 1, \text{for} (\mathrm{EP})_T = 1$.

Efficiency for cycle conversion

Depending on the energetic and exergetic viewpoint, the efficiency throughout the whole lifespan was regarding the transformation $(\eta)_{\mathrm{LCC}}$ in gross output versus the total incom-

$$\left(\eta_x\right)_{\mathrm{daily}} = \frac{\sum_{t=1}^{t=24}\left[h_{e-wgE}.\frac{A_b}{2}\left\{\left(T_w - T_{giE}\right) - T_a.ln\left(\frac{T_w}{T_{giE}}\right)\right\} + h_{e-wgW}.\frac{A_b}{2}\left\{\left(T_w - T_{giW}\right) - T_a.ln\left(\frac{T_w}{T_{giW}}\right)\right\}\right]}{\sum_{t=1}^{t=10}\left[0.933.\frac{A_b}{2}.\left\{I_{sE}(t) + I_{sW}(t)\right\} + \dot{m}_f C_f\left\{\left(T_{N-fo} - T_{fi}\right) - T_a.ln\left(\frac{T_{N-fo}}{T_{fi}}\right)\right\}\right]} \times 100 \tag{49}$$

Energy-exergy matrices observations

One may identify the development of renewable sources with its maximum used and applicable potential, which must be completed in a timely manner, and the scheme is referred as preferable, if it has a shorter payback time and increased conversion efficiency. The energy reward period, factor, and device's conversion efficiency are all considered in the study of energy matrices for provided life concerns of the projected model.

Payoff time

The compensated equivalent energy as well as exergy with respect to the used embodied energy of a unit and by the unit is represented by the payoff period of the systems. It is about as simple as possible, and the system will be deemed superior. The energy payoff period $\left(\mathrm{EP}_e\right)_T$ and exergetic payoff period $\left(\mathrm{EP}_x\right)_T$ as described can be used to indicate the payoff period $(\mathrm{EP})_T$.

$$\left(\mathrm{EP}_e\right)_T = E_{\mathrm{in}}/\left(E_e\right)_{\mathrm{yearly}} \tag{50a}$$

$$\left(\mathrm{EP}_x\right)_T = E_{\mathrm{in}}/\left(E_x\right)_{\mathrm{yearly-out}} \tag{50b}$$

"E_{in}" denotes the entire embodied energy of the organizational framework. Embodied type energy includes net energy used in the groundwork of each entity (Tiwari et al. 2016) that are incorporated to the suggested model, and the total embodied type energy represents the cumulative embodied form of energy (E_{in}) for the said framework (DS-SDS-EATC-MCPC), as shown in Table 4.

Payoff aspect

The proportional payoff aspect ($(\mathrm{EP})_F$) of the system reflects the reciprocal of the system's energy and exergy payoff ing solar radiation received across the present scheme itself (DS-SDS-N-EATC-MCPC). As a result, according to Singh and Samsher (2021a), the energy-based efficiency for product lifecycle transformation $\left(\eta_e\right)_{\mathrm{LCC}}$ may be written as

$$\left(\eta_e\right)_{\mathrm{LCC}} = \frac{\left(E_e\right)_{\mathrm{yearly}} \times n - E_{in}}{\left(E_e\right)_s} \tag{51a}$$

where $\left(E_e\right)_s$ is the total received solar energy throughout a lifespan ($n = 30$ yrs) and is represented as

$$\left(E_e\right)_s = n \times \sum_{i=1}^{i=365}\left[\frac{A_b}{2}.\left\{I_{sE}(t) + I_{sW}(t)\right\} + I_b(t).A_a\right] \tag{51b}$$

Similarly, product lifecycle transformation $\left(\eta_x\right)_{\mathrm{LCC}}$ for the exergy-based efficiency may be represented as,

$$\left(\eta_x\right)_{\mathrm{LCC}} = \frac{\left(E_e\right)_{\mathrm{yearly}} \times n - E_{in}}{\left(E_x\right)_s} \tag{51c}$$

where $\left(E_x\right)_s$ denotes the total solar irradiation exergy during the lifespan period as follows:

$$\left(E_x\right)_s = 0.933.n.\sum_{i=1}^{i=365}\left[\left\{I_{sE}(t) + I_{sW}(t)\right\} \times \frac{A_b}{2} + A_a \times I_b(t)\right] \tag{51d}$$

It is preferable to have a unit with a larger $(\eta)_{\mathrm{LCC}}$ and a much greater yield for a given lifetime.

Different economic observations

In this section, the many possible economic aspects (intended and unintended structure related to energy-exergy) have also been examined in order to comprehend the system's feasibility and long-term viability. The suggested system's environmental effect has also been studied in order to verify eco-friendly system integration as well as

Table 4 Divisions of diverse elements employed to construct the projected coordination (DS-SDS-EATC-MCPC) with bazaar value and energy density

Still elements	Specification	Cost ($)	E_{in}(kWh)
Top cover glaze	1.16 kg, 4 mm	12.56	51.6
FRP still basin	19 kg	126.34	487.16
EATC	4 nos	21.16	106.84
Mild steel stand	20 kg	17	190
Mild steel clamp	2.5 kg	2.13	23.75
Mild Stand frame	5 kg	4.3	47.5
Inlet–outlet nozzle	0.100 kg	2.2	2.13
Basin liner-paint	250 ml	2	6.3
Gasket	98.9 m, 2.1 kg	3	69.09
Parabolic reflector	2.8 kg	10.1	168.5
Net worth		200.79	1152.87

the monetary advantages to the country as carbon trading obtained standards.

System economic investigation

The existing economic assessment used for the current plan is totally considered for the current price technique and characterized by the many fiscal credits and are connected with various items and collectively as a system. The complete inspection was made during the course of 30-year operation of the model. The total yearly cost (TAC), which would be the resultant capital in the influence of constant annual charge (FAC) or, to put it another way, unified end of the year charge, yearly charged cost of maintenance (AMC), and restoration cost on a yearly basis are all part of the method (ASC).

Cost of distilled water

Furthermore, based on the current economic facts, the price of pure water (C) and the production price for energy (C_e) and exergy (C_x) may be stated as (Singh and Samsher 2021a; Reddy et al. 2018),

$$C = \mathrm{TAC}/\left(M_{ewg}\right)_a \tag{52a}$$

$$C_e = \frac{\mathrm{TAC}}{\left(E_e\right)_{\mathrm{yearly}}} \tag{52b}$$

$$C_x = \frac{\mathrm{TAC}}{\left(E_x\right)_{\mathrm{yearly-out}}} \tag{52c}$$

$$\mathrm{TAC} = \mathrm{FAC} + \mathrm{AMC} - \mathrm{ASC} \tag{52d}$$

where

$$\mathrm{FAC} = \mathrm{PCC} \times \mathrm{CRF}; \mathrm{ASC} = S \times \mathrm{SFF} \tag{52e}$$

For the whole analysis of economic measurements, the typical rate of interest (i) and appreciation are being considered 10% as focused on the trending movement of the market. Because the system is self-sustaining, just 5% of FAC was used for the report's running and maintenance costs. Furthermore, the salvage cost (S) or appreciating value after the system's estimated service life has been calculated as 20% of principal cash flow (PCC) or original investment cost, which includes individual unit price, making costs, commissioning costs, setting up costs, etc., as:

$$\mathrm{AMC} = 0.05 \times \mathrm{FAC} \tag{52f}$$

$$S = 0.2 \times \mathrm{SFF} \tag{52g}$$

where

$$\mathrm{SFF} = \frac{i}{(1+i)^n - 1}; \mathrm{CRF} = \frac{i.(1+i)^n}{(1+i)^n - 1} \tag{52 h}$$

where the capitalization recovery factor (CRF) gives a compounding incremental appraisal part to the current scheme over a specific length of instance of fundamental initial investment and net value, also known as flat yearly charge, which is a necessary requisite (yearly) to keep the arrangement running. In addition, safeguarding costs are a requirement in order to keep the system running properly. Furthermore, the system's salvage cost (S) at the end of its useful life generates extra returns and is determined by the graded sinking fund fraction (SFF) applied to the specified interest rate (i).

Productivity of scheme

The measuring of a scheme's productivity is a critical metric in determining its long-term viability and sustainability. It establishes the link between the output of the production and the basic variables necessary to attain that output. Its primary goal is to obtain the highest possible return with the least amount of resources invested. Furthermore, the suggested system's productivity $\left(\eta_P\right)_{\mathrm{yearly}}$ may be stated as follows: (Cox 1951; Benson 1952),

$$\left(\eta_P\right)_{\mathrm{yearly}} = \frac{\left[\left(M_{ewg}\right)_{\mathrm{yearly}} \times C_s\right]}{\mathrm{TAC}} \times 100 \tag{53}$$

where C_s and $\left(M_{ewg}\right)_{\text{yearly}}$ denote the annual yield production output of the proposed model ($5/l or $0.07/l) and the sale price of freshwater ($0.07/l), respectively.

Exergetic economic investigation

The exergy-influenced financial examination combines traditional fiscal study with exergetic study to determine the most cost-effective constraint for a particular configuration (Tsatsaronis et al. 1993; Tsatsaronis and Winhold 1985). As a result, scholars will be able to increase the establishment's performance in more suitable and gainful way (Tsatsaronis and Park 2002). Furthermore, as shown, the exergy-economy component $\left(E_x\right)_{\text{EF}}$ is described as the proportion of yearly exergy return to net annual cost written as

$$\left(E_{\text{x}}\right)_{\text{EF}} = \left(E_{\text{x}}\right)_{\text{yearly-out}} / TAC \tag{54}$$

Environmental-energetic-exergetic-economic investigation

The environment-economical examination is more reliable and valuable methodologies for encouraging investigators and clients of green technologies. It reveals the unnoticed and considerable returns in net carbon credentials while causing no harm to the environment through toxic gas emissions.

Correspondent emission evaluation of noxious gases

This segment relates emissions of hazardous gas to the environment, particularly CO_2, sulfur dioxide, and NO, that may be created in the same way during the generation of drinkable water by its lifelong utilization. It has been established that coal-fired power plants emit a significant level of 0.98 kg, 0.008 kg, and 0.003 kg of CO_2, sulfur dioxide, and NO emissions for 1kWh of electrical power output, respectively (Tiwari and Mishra 2012; Kalshian 2008; Watt et al. 1998). Furthermore, after accounting for transmitting (40%) and distributing (20%) losses, the effective values of such pollutants are 1.58 kg, 0.012 kg, and 0.005 kg, accordingly, and are only relevant to Indian regime conventional thermal power stations. Likewise, the corresponding emissions for $(\text{EEq})_{\text{CO}_2}$, $(\text{EEq})_{\text{SO}_2}$, and $(\text{EEq})_{\text{NO}}$ for the present scheme (DS-SDS-EATC-MCPC) may be articulated as (EEq) (NO) (Singh and Samsher 2021a; Reddy et al. 2018),

$$(\text{EEq})_{\text{CO}_2} = E_{\text{in}} \times 1.58 \tag{55a}$$

$$(\text{EEq})_{\text{SO}_2} = E_{\text{in}} \times 0.012 \tag{55b}$$

$$(\text{EEq})_{\text{NO}} = E_{in} \times 0.005 \tag{55c}$$

Energetic-exergetic mitigation of noxious gases

Various atmosphere pollution mitigations over the lifespan of the developed framework may be described as (Singh and Samsher 2021b; Mittal et al. 2014),

$$\left(\text{MEq}_{\text{e}}\right)_{\text{CO}_2} = 1.58 \times \left\{ \left(E_{\text{e}}\right)_{\text{yearly}} \times n - E_{\text{in}} \right\} \tag{56a}$$

$$\left(\text{MEq}_{\text{e}}\right)_{\text{SO}_2} = 0.012 \times \left\{ \left(E_{\text{e}}\right)_{\text{yearly}} \times n - E_{\text{in}} \right\} \tag{56b}$$

$$\left(\text{MEq}_{\text{e}}\right)_{\text{NO}} = 0.005 \times \left\{ \left(E_{\text{e}}\right)_{\text{yearly}} \times n - E_{\text{in}} \right\} \tag{56c}$$

Similarly, the environment pollutant mitigations based on environment, energy, and economy may be depicted throughout the lifespan of the proposed model.

Energy-exergy-based carbon credit revenue

Pollution mitigations from the mechanism can be disintegrated and referred to as carbon credits obtained through the worldwide market's price structure. These values change over time, but according to Singh and Samsher (2022), and Roome (2019), and these were acknowledged at $9.99/ton of carbon dioxide. Further, the obtained fiscal worth is recognized as the meandering income (i.e., carbon credit collected) for the proposed still. As a result, the energetic green gain $\left(E_{\text{ce}}\right)_{\text{CO}_2}$ and exergetic environmental gain $\left(E_{\text{cx}}\right)_{\text{CO}_2}$ of carbon mitigates to the proposed system's lifespan may be represented as,

$$\left(E_{\text{ce}}\right)_{\text{CO}_2} = \left[1.58 \times \left\{ \left(E_{\text{e}}\right)_{\text{yearly}} \times n - E_{\text{in}} \right\}\right] \times 9.99 \tag{57a}$$

$$\left(E_{\text{cx}}\right)_{\text{CO}_2} = \left[1.58 \times \left\{ \left(E_{\text{x}}\right)_{\text{yearly-out}} \times n - E_{\text{in}} \right\}\right] \times 9.99 \tag{57b}$$

Analytical methodology

Following the methodological procedures, the resultant expression is sequentially followed up on to compute various resulting parameters. The stages involved in the operating approach are outlined as follows:

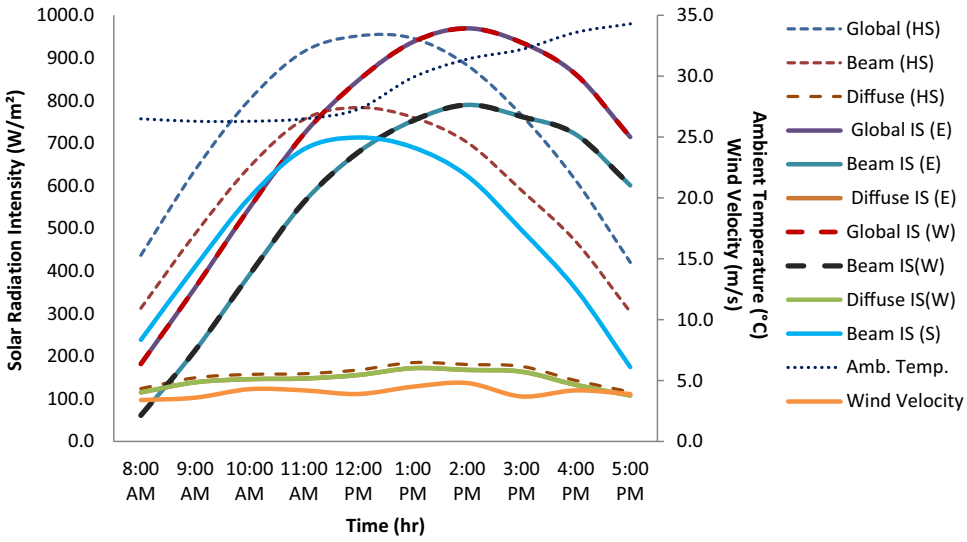


Fig. 5 Representation of hourly variations of different solar irradiations at horizontal along with inclined surfaces with air-stream velocity and atmospheric temperatures

First, the influence of solar irradiance over the sloped surface of solar receiver is investigated using data (IMD, Pune, India) at flat surface to the ground, which is then transformed for the solar response over incline plane using Liu and Jordan's method (1960).

Then resultant equations are solved by expressing water and condensation surface temperatures in context of known parameters such as solar flux and coefficients using the MATLAB computing tool. Through Eq. (37), yields are determined for the suggested solar still using T_{giE} and T_{giW} (Eqs. (26)–(27)) and T_w (Eq. (30a)). Equation (36) is used to calculate the influence of solar irradiation on thermo syphon flow rate in EATC owing to MCPC.

The system's hour energy (Eq. (38)), efficiencies (Eqs. (40) and (48)), and exergy (Eq. (42)) are examined in the third stage. Then, for daily energy-exergy, sums of hourly energy-exergy for 24 h are employed. Monthly energy-exergy is calculated by multiplying everyday energy-exergy by frequency of a sunny day (monthly) recorded. The sum of monthly energy-exergy for a period of 12-month yields represents the yearly energy-exergy accordingly.

Energetic and exergy matrix (Eq. 50) for the planned solar still are being completed. The suggested system has also undergone multiple economic analyses (economic evaluation, exergy economy, and environment economy). CO_2 emissions, CO_2 remediation, overall carbon credits, and environmental values are all evaluated in the environmental economic analysis (Eq. 57). In addition, a comparison analysis for the suggested solar desalination system combinations is offered.

Result investigation and discussion

The present section entails the numerical equations concerned to the proposed system's thermal form in order to simulate the findings and the demonstration of the research findings and predicts fundamental evaluation: feasibility, self sufficiency, eco-friendliness, cost-effectiveness, and competent performance. The projected analysis is being evaluated for 0.16-m working medium (i.e., depth of water) into the basin, as well as the water-filled EATCs (parallel arrangement with four optimal vacuum annular tubes) of 30° angulations for DS-SDS and EATC together to receive most of the solar rays during the day. The intact set of observations was evaluated using meteorological data from IMD, Pune, for a typical clear day in June in the New Delhi area of India.

Figure 5 depicts the hourly variability of various solar irradiative elements (e.g., global, diffused, beam, and earth reflected rays) for flat surface to the ground (IMD Pune, solar radiation foundation data), as well as corresponding solar irradiance over the oriented surface and atmospheric temperature for the corresponding evident day. Owing to the unavailability of ground reflective radiant energy and other minimal reduction due to high altitude of unit location and environs, the results obtained of astral irradiation at tending plane are reduced as compared to received radiations for flat plane, which make a meaningful impact over inclined plane radiant energy. However, the oriented plane at a latitudinal angle has an excellent solar exposure to strike the plane in perpendicular direction and > 95% of incoming sun rays that is conveyed to the DS-SDS compartment and an analogous (//) arrangement of vacuum annular tubes, which would be impossible to achieve with horizontal coverings. In the meanwhile, the air temperature ranges from 24 to 35 °C, with an average wind speed of 4.11 m per second, favoring the desalination process during the daylight hours.

Thermal validation: relative perspective

In conjunction with cusp-oriented CPCs and solely coupled by DS-SDS unit, the suggested method offers a unique way to combine for parallel arranged EATCs. This sort of innovative design agreement has not yet been mentioned, though, to affirm this scheme, the obtained energy formulations are now implemented, which include still's water hotness, upper glass cover interior face hotness, gain, and thermo syphon liquid flowrate, which are relatively similar to the governing model parameters presented by Dubey et al. (2021) and Singh et al. (2013), as given in Fig. 6. Further, the conclusion remarks are in strong harmony, wherever MCPC and depth of water have minor effects. The influence of MCPC boosts the circulatory thermo syphon circuit in the course of EATC (predominantly at higher sun's ray instant); consequently, sink fluid heats up quickly, and increased water medium deepness (0.16 m) (greater sensible heat absorptive competence) greatly holds such alteration. For framework mentioned by Dubey et al. (2021), and Singh et al. (2013), the current proposal holds commendable results and demonstrates substantial improvement in yields owing to corresponding increases in brackish water quantity (433 percent), basin temperature of the water (6.2%, and 1.2%), potable water (334.4% and 153%), and mass flow (27.6%, and 36.2%). The suggested system and thermal simulation justify the precision based on the above validated observations of the findings, and may be used to assess performance of the unit under the provided geometrical constraints of the system. And throughout the day, Fig. 7 depicts the hourly fluctuations in mass flow and yielding output.

Proposed model optimization

Influence of sun irradiance over thermo syphon Because of the EATC-MCPC combination's techno-specialty, the assembly's effective ratio of concentration (CR) is 2. In the integrated assortment of EATC-MCPC, the combined use of OCPC-CCPC allows a radiation gain of 65.63% towards the segmental edges, and 68.75% at the beneath halves of annular tubes during solar irradiance (direct and diffused). As shown in Fig. 4, the innovative design presents a sophisticated set with DS-SDS-N-EATC-OCPC-CCPC that indicates the efficient use of total solar energy via EATC to SDS. The enhanced thermo syphon circulating circuit into EATC is aided by uniform solar irradiance around the perimeter, and the increased heat transfer raises the temperature of reservoir volume of water proportionally.

Importance of SDS available water thickness and amount of EATCs Due to the constraints of thermal model, which is not applicable for the boiling point of water or well beyond, the suggested model necessitates the optimal usage of parallel sequence EATC. As a result, the analysis is being carried out for varying nos. of EATC under various basin water levels for the given SDS system, and it is observed that the 4 nos. of annular tubes of 0.16 m water depth achieve 99.6 °C hotness of brackish water into the basin, as shown in Fig. 8, and thus considered as an optimized design parameter. The overall model parameters which are considered for the said suggested scheme (DS-SDS-N-EATC-CCPC-OCPC) and pointed as superior than other associated arrangements

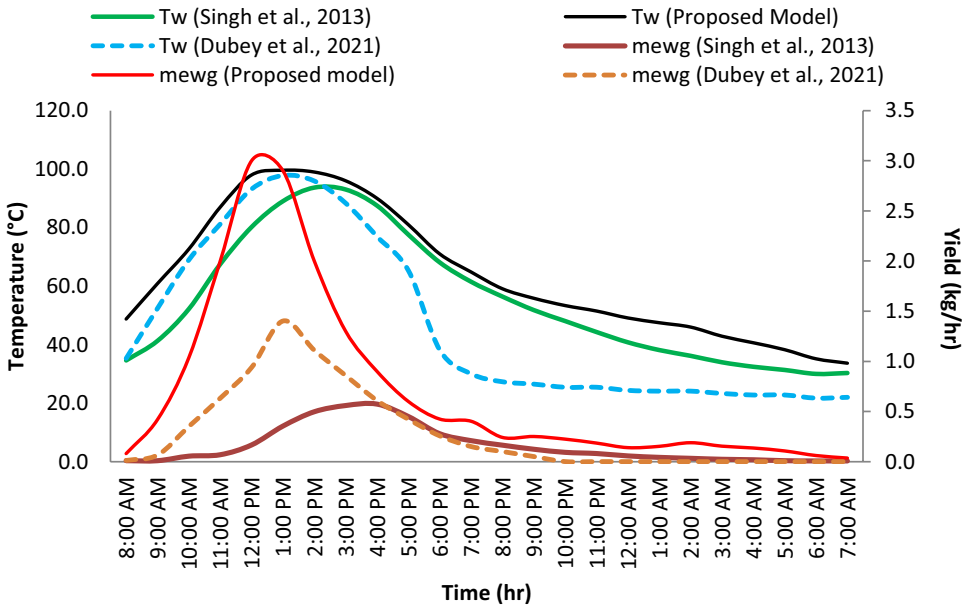


Fig. 6 Relative observation with validation of the thermal model of projected scheme for hourly deviation of basin's brackish water temperature and distillate production

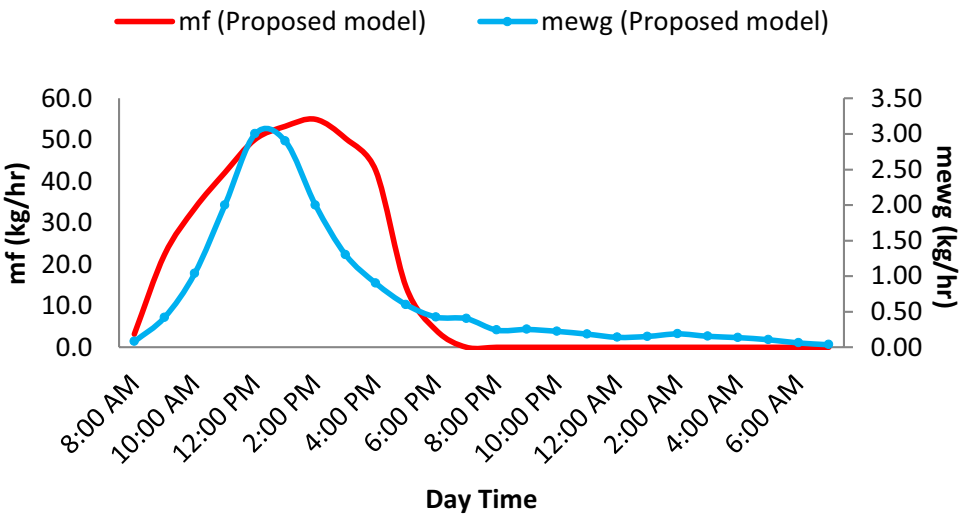


Fig. 7 Hourly deviation of thermo syphon flow rate and yielding throughout the day

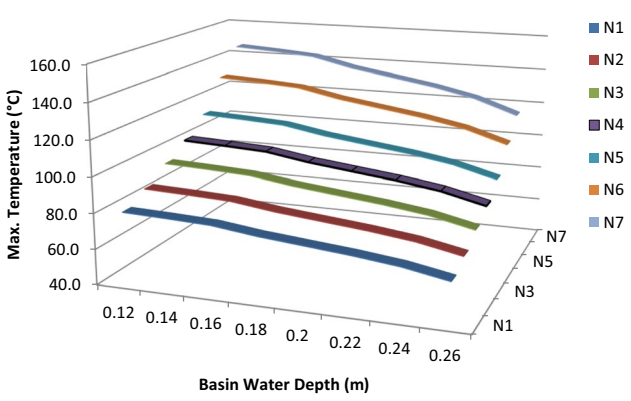


Fig. 8 Demonstration of still's water highest temperatures relative to deviated water deepness conditions and invariable numbers of annular tube collectors in the present scheme

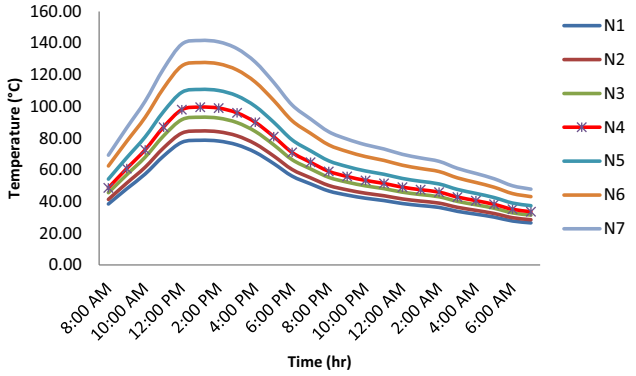


Fig. 9 Depiction of still's brackish water temperatures for changeable EATC under similar water depth conditions in a typical apparent day meteorological conditions of June

are then justified in Fig. 9. The optimal use of components allows for a basin temperature of approximately to 99.6 °C, as well as a substantial volume of water absorbing a larger quantity of warmth (sensitive heat) in addition to retain and increase temperature meant for a prolonged timeframe, which is beneficial for nocturnal output generation.

Performance examination: energy-exergy contemplation

For a prototypical bright day in June, Fig. 10 depicts temperature differences in different areas of the system. The EATC water temperature effectively regulates the basin's fluid temperature, with a quick rise toward 1 pm and wider temperature fluctuations, leading to a high flow velocity. The thermo syphon rotations into EATC are caused by the apparent temperature differential, which is accessible approximately until 7 p.m. Also, up to 2 p.m., a large temperature difference, $T_w - T_{giE}$, $T_w - T_{giW}$, occurs, and afterwards, minimal temperature variations persist until late at night, providing diurnal and nocturnal supplies, respectively. In addition, Fig. 11 depicts the corresponding fluctuations in heat transmission factors in the basin area. It demonstrates that East–West dehumidification heat transfer imparts a significant role in heat transport from the surface of water to upper wrap glass, whereas radiative heat flow has the least effect owing to substantially lower temperature of the water (i.e., less than 100 °C) and performs admirably at elevated temperature (i.e., upwards of 1000 °C). The most substantial benefits may be seen in the middle of the day, when the sun intensity is at its highest. The amount of vapor formation in the basin affects the convection heat transmission factor, which is higher in the middle of the day. As illustrated in Fig. 11, the fractional radiations transported from East to West or West to East are quite comparable in both circumstances.

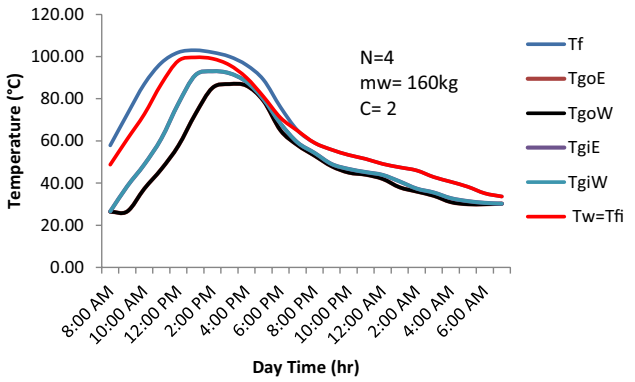


Fig. 10 Temperature differences at various sections in the model for a representative sunny day in June month

Figures 12–13 show the yield production with the matching thermo syphon at various basins' medium temperatures (T_w) and temperature changes during the day. As demonstrated in Fig. 13, yield production follows the mass flow and ($T_w - T_{giE}, T_w - T_{giW}$). However, the highest yield (3 kg) popped at 1:00 p.m. owing to utmost $T_w - T_{giE}$, as well as $T_w - T_{giW}$ (26 °C); even so, the mass run (55 kg/hr) appeared at 2:00 p.m. because optimum temperature variation was reached at that time, rapidly increasing basin fluid temperature, resulting in increased evaporation that was not applied due to excessive vapor production, resulting in faintly lower yielding at around that moment. The total distillate production pace for the suggested system is 0.08 kg/°C of thermal gradient ($T_w - T_{giE}, T_w - T_{giW}$) for such an iconic sunny day in June. In addition, when compared to other schemes of solar still desaltification units for alike regions of experience, the suggested system's average hourly yield, 0.71 kg/m^2, is much greater, i.e., 0.71 kg/m^2.

The highest benefit of 1.91 kWh as well as 1.48 kWh is obtained during prime hours of astral energies having lower most differences of $T_w - T_{giE}, T_w - T_{giW}$ which appeared as shown in Figs. 14–15, under the corresponding clear day circumstances, as calculated by Eqs. (38) and (42) accordingly. Furthermore, the energetic and exergetic efficiencies are sensitive metrics that are affected by both the SDS basin's crucial integral capabilities as well as the EATC, as illustrated in Fig. 16. As shown in Eq. (39), the energy efficiency is affected by temperature differences, incoming radiation from the sun, heat transmission factor, and transfer of energy from EATC → DS-SDS. High thermal differences (22.15–5.98 °C) result in reduced energy efficiency; nevertheless, as the temperature gradient increases, efficiency increases (7.65–73.99%), which is evident up to noontime. As the temperature gradient, and the sun intensity, decreases in the evening, efficiencies are much more aggressively moving higher.

Exergy efficiency, on the other hand, shows a contrasting picture, with exergy efficiency decreasing (3.01–1.94%) when sunlight exergy and temperature disparities grow owing to more irreversible phenomena, particularly up to solar prime times, as represented in Fig. 16. Likewise, due to denominator impact as depicted in Eq. (48), the afternoon has the highest exergy efficiency (14.6%) owing to the smallest temperature variations. Then, owing to evaporative exergy transmission, exergy efficiency declines (14.6–0.09%) under the impact of irrevocable exergy

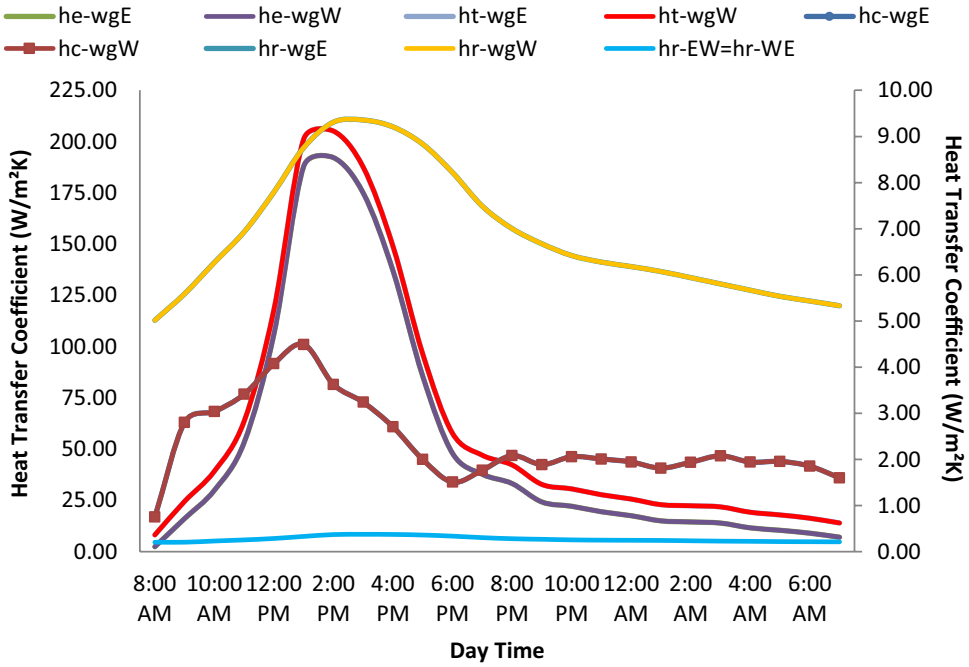


Fig. 11 Variations in coefficient of heat transfer at diverse segments in the model for $N=4$ in an archetypal sunny day of June

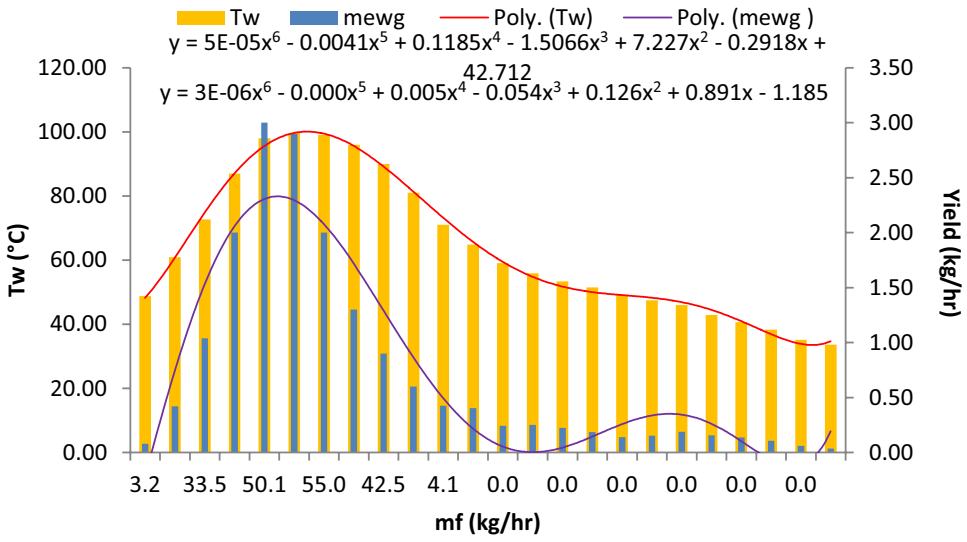


Fig. 12 Yield output representation at different basin temperatures and corresponding mass flow rate

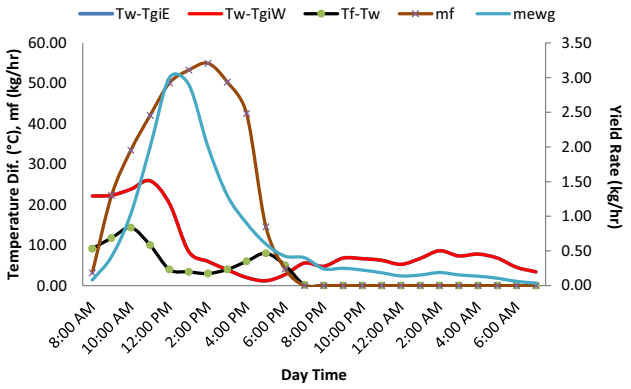


Fig. 13 Hourly variations of mass flow rate, yield output, and temperature differences

affiliation of respective components. Furthermore, due to a reduction in sunlight exergy along with temperature variations, exergy continues to rise until 6 p.m.

Energy matrix possibilities

Table 4 shows the divisions of several elements, as well as respective parameters, embodied based energy, and existing market value, which are used to evaluate various features of energetic-exergetic matrix. Table 5 shows the results of the said proposed system's energy-exergy matrix, specifically for addressing 275 typical nos. of bright days over 30 years work experience life deliberations, with an average daily solar strength of 401.8 kW.

Figure 17 shows the proposed model's energy-exergy matrix during the course of its recommended operating life of 30 years. The system's energy recovery period is

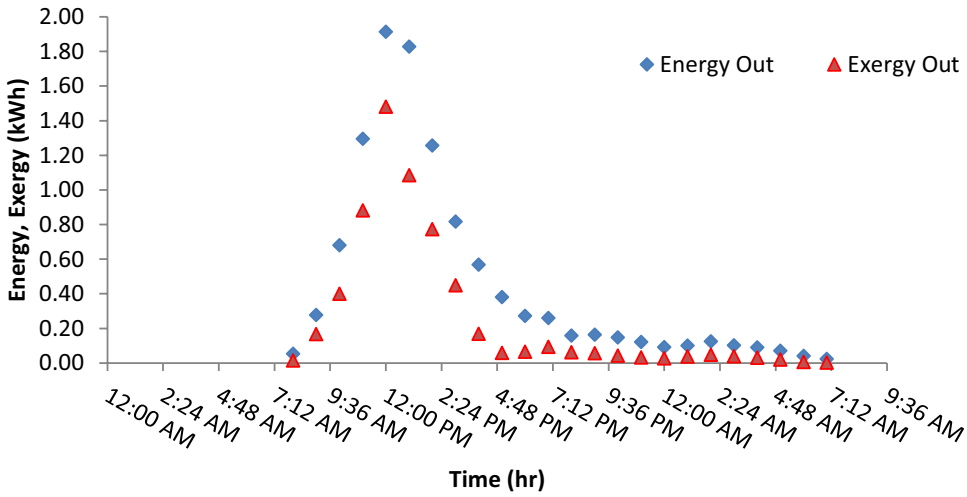


Fig. 14 Energetic-exergetic output of present scheme for a particular apparent day circumstances in June

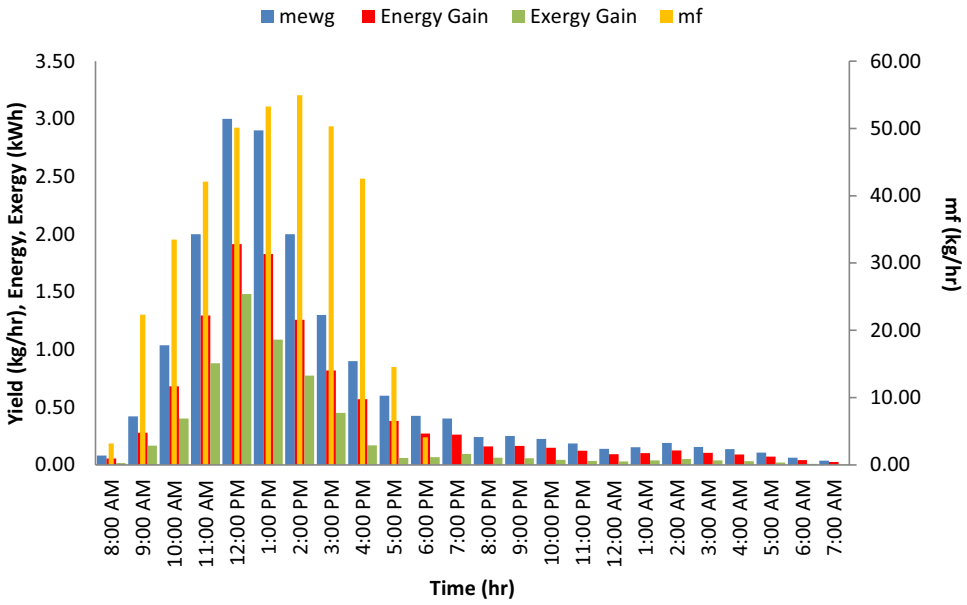


Fig. 15 Disparity of yield, exergy, energy on hourly basis for the projected model

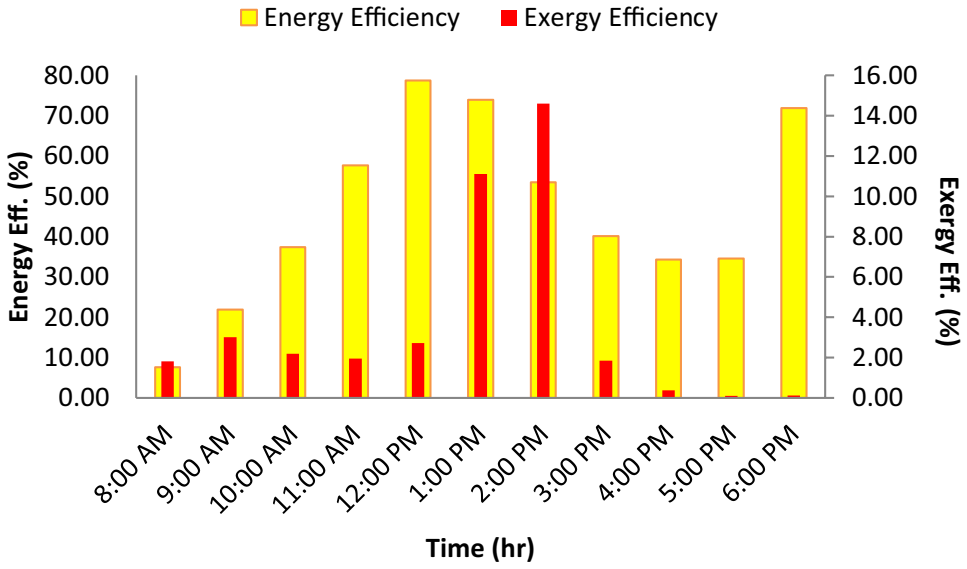


Fig. 16 Depiction of energetic-exergetic efficiency deviations of the model in June for $N=4$

0.39 years, but the exergy payoff duration is a bit longer, at 0.69 years. Because of larger number of exergy destructions involved, the exergy payoff period is taking longer. The related energetic-exergetic payoff values seem to be 2.59 and 1.45, correspondingly, reflecting the system's capabilities and competences, with the larger payoff factor representing the system's goodness. These outcomes are primarily determined by the system's embodied energy and yearly yield energetic-exergetic production, and are expressed as a longer energetic-exergetic payoff duration or multiplier for greater yearly production, or a lower embodied energy per unit. Figure 18 shows the system's energy-exergy-based lifespan conversion, which is considered superior if the corresponding efficiencies are much greater (tending to 1), like in the developed framework. Under varying system life operating circumstances, the energy-exergy service life conversion efficiency patterns are roughly identical. It improves with the length of a system's life, as seen by the 0.57 and 0.34 energetic and exergetic conversion efficiencies discovered over the first 30 years of the system's operation.

Technical-economical-environmental-design: rudiments and capabilities

Figure 19 depicts the yearly production output and exergo-economic aspect fluctuations of the proposed model under the effect of varying rate of interest and yield sales price. The system's productivity and exergy-based economic factor falls

Table 5 Payback period, payback factor, and efficiency of conversion with respective findings of the planned solar desaltification arrangement for 30 years of lifespan

Considerations	Results	
	Energy	Exergy
Yearly yield (kg)	4658.5	
Yield of energy day (kg/day)	16.94	
Lifespan yield (kg)	139,755	
Scheme running existence (years)	30	
E_{in}(kWh)	1152.87	
Average each day solar strength	401.8 kW	
Net out energy during life (kWh)	89,595	50,077.5
Life time solar energy (kWh)	154,192.5	143,861.6
Yearly solar energy (kWh)	5139.75	4795.39
yearly out energy from still (kWh)	2986.5	1669.25
Payoff time (years)	0.39	0.69
Payoff part	2.59	1.45
Life cycle conversion efficiency	0.57	0.34

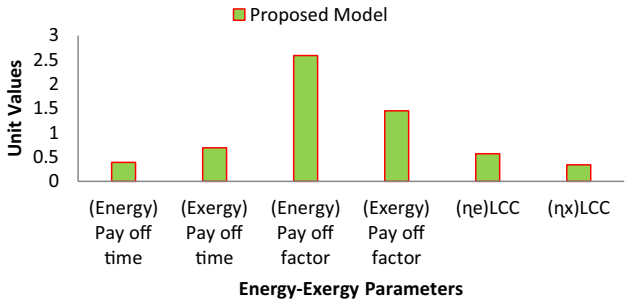


Fig. 17 Allocation of energy matrix of the presented scheme for 30 years working life

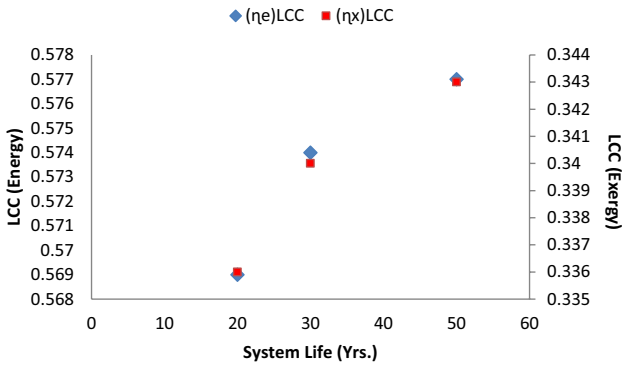


Fig. 18 Deviation in conversion efficiency for life cycle over the age of the system's existence

when interest rates rise, because a rise in interest rates raises the denominator amount in the subsequent Eqs. (53)-(54), lowering the total productivity and exergy-based economic

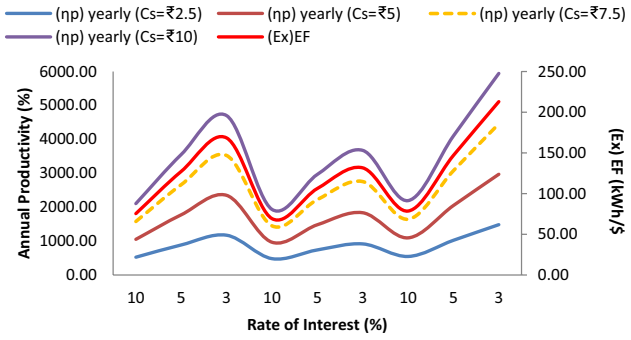


Fig. 19 Deviations of yearly productivity and exergy-based economic aspect under the control of diverse interest rates of anticipated scheme

component of the model. The system is significantly practicable for its continued survival into highly spirited environment to generate drinkable water with yearly productivity greater than 100%. For a 10% inflating rate of interest and a \$0.07/l sale price of yield, the findings discovered for such productivity and exergy-based economic aspects are far more to 100% and 75.46, correspondingly, for the suggested system analysis.

Furthermore, the current proposal is capable for effectively mitigating pollutants (carbon dioxide, SO_2, and NO) over its total service life due to significant yield production, with corresponding energetic-exergetic mitigates of 139.74 and 77.30 tons referred to carbon dioxide, 1.16 and 0.59 tons referred to sulfur dioxide, and 0.44 and 0.24 tons referred to NO, correspondingly. And, according to energy-exergy measurements, the suggested system's environmental cost generated as carbon credit income (indirect cost) over a 30-year life equals \$1396 and \$772.24, correspondingly. Figure 20 shows a generalized comparison of several enviro-economic characteristics and yields of the present scheme with those of previous solar desalination models. To observe an output result on a similar trajectory of multiple

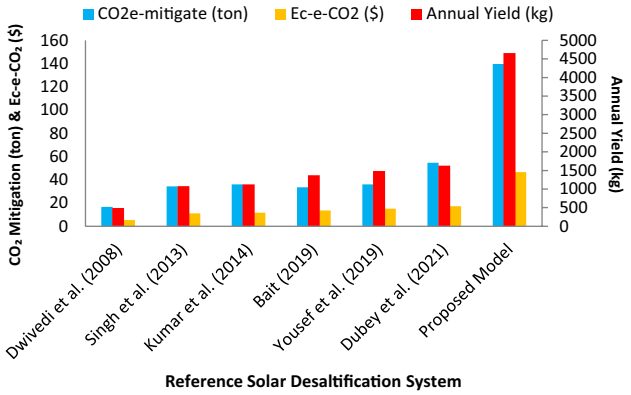


Fig. 20 Relative interpretation of environ-economic perceptions for diverse solar desaltification models in comparison to the suggested scheme of representation

systems, the findings reported by various researchers are comprehended according to a frequently referenced existence (30 years), carbon remediation (1.58 kg carbon dioxide/kWh emissions or mitigation), and carbon trading received a total at the rate of $9.99/ton of carbon dioxide. Regarding the comparative assessments, the suggested system has a greater effect and is preferred in terms of many environmental economic factors and yielding; hence, the current geometric association is superior to the other designs studied.

Conclusion

The developed framework (DS-SDS-N-EATC-MCPC) is being established via thermal model to estimate solar irradiation responses and eco-design necessities depending on optimum-environment-economic dependability over the performances, energetic-exergetic matrices, and different economic reflections under the application of MCPC implications. The following results concerning the optimal number of EATC relationships are based on the current study as framed:

i. The combination of MCPC and four optimal numbers of EATCs combined with DS-SDS performs adequately, reaching a record of 99.6 °C basin's water temperature during the day's peak times of solar irradiation. Owing to sensible heat usage for nocturnal distillation, the unique combination of MCPC-EATC delivers nearly consistent solar irradiance surrounding the EATC for concentration range 2 that enhances the temperature of the water having greater water depths. With a daily output of 16.94 kg, the combined diurnal and nocturnal yielding output is pretty impressive.
ii. The daily energetic-exergetic efficiencies are 46.53%, and 3.62%, accordingly, which fluctuate during the daylight and display its furthermost outcomes for respective temperature (min–max) variances.
iii. Corresponding interest rates, yielding cost, green revenue, etc. influence the cost of production. The suggested system's energetic-exergetic output generation costs are highly noticeable at $0.007 and $0.013/kWh respectively over a 30-year life cycle. For 0.39 years and 0.69 years, respectively, the energetic-exergetic payoff period is adequately provided. The suggested system's energetic-energetic conversion efficiency over time is claimed to be 0.57 and 0.34, respectfully.
iv. For carbon dioxide, the energy-exergy-based mitigates are 139.74 and 77.30 tons; for sulfur dioxide, 1.06 and 0.59 tons; and for NO, 0.44 and 0.24 tons, respectively. The carbon credit generates $1396 and $772.24 in environmental revenue, respectively, depending on energy-exergy. The scheme is significantly practical for its continued survival in the aggressive marketplace in favor of generating clean water, with an exergy-based economic factor of 75.46kWh/$ and productivity of more than 100%.
v. The system is sustainable and viable for suitable collector areas and optimal EATC using MCPC as eco-design necessary conditions for the suggested system because of the visible output at relatively low production price, environmental income, and substantial mitigation, with low payoff duration.

Appendix

Relations were utilized to solve Eq. (1) following Singh and Samsher (2020), Duffie and Beckman (2006), and Dunkle (1961) as stated,

$$U_{saa} = \left(\frac{R_{o2}\ln\left(\frac{R_{i2}}{R_{i1}}\right)}{K_g} + \frac{1}{h_{r-v}} + \frac{R_{o2}\ln\left(\frac{R_{o2}}{R_{o1}}\right)}{K_g} + \frac{1}{h_o} \right)^{-1}$$

$$I_s(a) = I_b(t).(\propto \tau)_{eff}$$

$$(\propto \tau)_{eff} = R_{sc}\alpha\tau^2\left(A_a/A_{rc}\right)$$

$$h_o = h_{c-wg} + h_{r-wg} = 5.7 + 3.8V$$

$$h_{r-v} = \varepsilon_{eff}.\sigma\left[\left(T_f + 273.15\right)^2 + \left(T_{sa} + 273.15\right)^2\right] \times \left[T_f + T_{sa} + 546.30\right]$$

$$\varepsilon_{eff} = \left(1/\varepsilon_g + 1/\varepsilon_w - 1\right)^{-1}$$

Expressions were utilized to solve Eqs. (7), (10), (17), and (19) as mentioned,

$$F_1 = \frac{F'.h_{saf}}{(F'h_{saf} + U_{saa})}; F_2 = 1 - \frac{\left(A_{rc}.F_r\right)U_L}{\dot{m}_f C_f}; U_L = \frac{F'.h_{saf}.U_{saa}}{(F'h_{saf} + U_{saa})}$$

$$F_r = \frac{\dot{m}_f C_f}{A_{rc}.U_L}\left\{1 - exp\left(-\frac{2\pi R_{i1}L_{EATC}.U_L}{\dot{m}_f C_f}\right)\right\}$$

$$h_{t-gE} = h_{t-gW} = 5.7 + 3.8V$$

$$U_{cE} = h_{t-gE}/\left(1 + \frac{h_{t-gE}}{K_g/t_g}\right); U_{cW} = h_{t-gW}/\left(1 + \frac{h_{t-gW}}{K_g/t_g}\right)$$

Relations were utilized to solve Eqs. (20)-(21) as mentioned (Cooper 1973; Dunkle 1961; Singh and Samsher 2021a, b, c; Singh and Samsher 2022),

$$\alpha_g^A = \left(1 - R_g\right)\alpha_g; R_g = 1 - \left[4n_o n_g / \left\{\left(n_o + n_g^2\right) \times (1 + n_o)\right\}\right]$$

$$h_{t-wgE} = h_{e-wgE} + h_{r-wgE} + h_{c-wgE}$$

$$h_{t-wgW} = h_{c-wgW} + h_{e-wgW} + h_{r-wgW}$$

$$h_{r-wgE} = \varepsilon_{eff}.\sigma\left\{\left(T_w + 273.15\right)^2 + \left(T_{giE} + 273.15\right)^2\right\}\left\{T_w + T_{giE} + 546.30\right\}$$

$$h_{r-wgW} = \varepsilon_{eff}.\sigma\left\{\left(T_w + 273.15\right)^2 + \left(T_{giW} + 273.15\right)^2\right\}\left\{T_w + T_{giW} + 546.30\right\}$$

$$h_{c-wgW} = 0.88 \times \left\{\left(T_w - T_{giW}\right) + \frac{\left(T_w + 273.15\right) \times \left(P_w - P_{giW}\right)}{\left(268.9 \times 10^3 - P_w\right)}\right\}^{(1/3)}$$

$$h_{c-wgE} = 0.884\left\{\left(T_w - T_{giE}\right) + \frac{\left(P_w - P_{giE}\right)\left(T_w + 273.15\right)}{268.9 \times 10^3 - P_w}\right\}^{1/3}$$

$$h_{e-wgE} = 16.273 \times 10^{-3} h_{c-wgE}\left\{\frac{P_w - P_{giE}}{T_w - T_{giE}}\right\}$$

$$h_{e-wgW} = 16.273 \times 10^{-3} h_{c-wgW}\left\{\frac{P_w - P_{giW}}{T_w - T_{giW}}\right\}$$

$$h_{r-EW} = h_{r-WE} = 0.034 \times \sigma\left\{\left(T_{giE} + 273.15\right)^2 + \left(T_{giW} + 273.15\right)^2\right\}\left\{T_{giE} + T_{giW} + 546.3\right\}$$

$$P_w = \exp[25.317 - \frac{5144}{(T_w + 273.15)}]$$

$$P_{giE} = exp[25.317 - \frac{5144}{T_{giE} + 273.15}]$$

$$P_{giW} = exp\left(25.317 - \frac{5144}{T_{giW} + 273.15}\right)$$

Equations were utilized to solve Eq. (22) as mentioned (Tiwari 2014; Singh and Samsher 2021a, b, c),

$$h_{bw} = \left(K_w / t_b\right)\left\{C\left(R_a\right)^{1/n}\right\}$$

$$\alpha_w^A = \left\{\left(1 - R_g\right)\left(1 - \alpha_g\right)(1 - R_w)\alpha_w\right\}$$

$$R_w = \left[1 - \left(4n_o n_w\right) / \left\{\left(n_o + n_w^2\right)(1 + n_o)\right\}\right]$$

$$R_a = \frac{g\beta' \rho^2 X^3 C_w \Delta T}{\mu K_w}; G_r = \frac{g\beta' \rho^2 X^3 \Delta T}{\mu^2}; P_r = \frac{\mu C_w}{K_w}$$

$$X = \left(L_o - B_o\right)/2(\text{Forrectangularhorizontalsurface})$$

$$X = \frac{\text{Area}(A)}{\text{Perimeter}(P)}(\text{Forothersurfaces})$$

Expressions were utilized to solve Eq. (25) as mentioned (Tiwari 2014),

$$h_{ba} = \left\{\left(t_b / K_b\right) + 0.357\right\}^{-1}$$

$$\alpha_{beff} = \left(\alpha_b^A.h_{bw}\right) / \left(h_{ba} + h_{bw}\right)$$

$$U_{bwa} = \left(h_{bw} \times h_{ba}\right) / \left(h_{ba} + h_{bw}\right)$$

Relations utilized to solve Eqs. (26)-(29), (31) as mentioned,

$$U_{taE} = \left(U_{cE} \times A_{gE} \times h_{t-wgE}\right) / \left[\frac{A_b}{2} \times h_{t-wgE} + A_{gE} \times \left(h_{r-EW} + U_{cE}\right)\right]$$

$$U_{taW} = \left(U_{cW} \times h_{t-wgW} \times A_{gW}\right) / \left[h_{t-wgW}.\frac{A_b}{2} + A_{gW}.\left(h_{r-WE} + U_{cW}\right)\right]$$

$$h'_{1E} = \left(A_{gE}.h_{t-wgE}\right) / \left[\frac{A_b}{2}.h_{t-wgE} + A_{gE}.\left(h_{r-EW} + U_{cE}\right)\right]$$

$$h'_{1W} = \left(h_{t-wgW}.A_{gW}\right) / \left[\frac{A_b}{2}.h_{t-wgW} + A_{gW}.\left(h_{r-WE} + U_{cW}\right)\right]$$

$$h'_{2E} = \left(\frac{A_b}{2}.h_{t-wgE}\right) / \left[h_{t-wgE}.\frac{A_b}{2} + A_{gE} \times \left(h_{r-EW} + U_{cE}\right)\right]$$

$$h'_{2W} = \left(\frac{A_b}{2}.h_{t-wgW}\right) / \left[h_{t-wgW}.\frac{A_b}{2} + A_{gW} \times \left(h_{r-WE} + U_{cW}\right)\right]$$

$$a = \frac{1}{m_w C_w} \times \left[U_{eff}^{DS} + h_{r-EW}.\frac{A_b}{2}.\left(h'_{1E} + h'_{1W}\right) + \dot{m}_f C_f.(N - 1)\right]$$

$$U_{eff}^{DS} = \left(\frac{U_{taE}}{2} + \frac{U_{taW}}{2} + U_{bwa}\right) \times A_b + U_L \times \left(A_{rc} \times F_r\right)$$

$$f(t)=\left[U_{eff}^{DS}\times T_{a}+\left\{I_{sE}(t)+I_{sW}(t)\right\}\times\left(\alpha_{beff}\times\frac{A_b}{2}+\alpha_{w}^{A}\times\frac{A_g}{2}\right)\right.$$
$$+\left(A_{rc}.F_r\right).F_1.(\propto\tau)_{eff}.I_b(t)$$
$$\left.+\frac{A_b}{2}\left\{T_{giW}.h_{r-EW}(h'_{1E}+h'_{1W})+\alpha_g^A\left(h'_{1E}.I_{sE}(t)+h'_{1W}.I_{sW}(t)\right)\right\}\right]/\left(m_wC_w\right)$$

$$(Nu.Gr)/Pr=\left(g.\beta'.d^4.\dot{q}\right)/\left(\nu^2.\mu.C_f\right)$$

$$Nu=\frac{h.X}{K_w};\nu=\mu/\rho$$

Author contribution Ashok Kumar Singh: conceptualization, data curation, formal analysis, validation, investigation, methodology, software, writing—original draft.

Samsher Gautam: supervision, resources, visualization, project administration.

Data availability The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethical approval Not applicable.

Consent to participate Not applicable.

Consent to publish Not applicable.

Competing interests The authors declare no competing interests.

References

Badran OO, Abu-Khader MM (2007) Evaluating thermal performance of a single slope solar still. Heat Mass Transfer 43:985–995

Bait O (2019) Exergy, environeeconomic and economic analyses of a tubular solar water heater assisted solar still. J Cleaner Production 212:630–646

Barbosa EG, Martins MA, Araujo MEV, Renato NS, Zolnier S, Pereira EG, Resende MO (2020) Experimental evaluation of a stationary parabolic trough solar collector: influence of the concentrator and heat transfer fluid. Journal of Cleaner Production 276:124174

Benson, F., 1952. Further notes on the productivity of machines requiring attention at random intervals. J. Roy. Stat. Soc. B XIV, 200–210.

Budihardjo I, Morrison GL (2009) Performance of water-in-glass evacuated tube solar water heaters. Sol Energy 83:49–56

Budihardjo I, Morrison GL, Behnia M (2007) Natural circulation flow through water-in- lass evacuated tube solar collectors. Sol Energy 81:1460–1472

Cengel XA, Boles MA (2013) Thermodynamics, an engineering approach. McGraw-Hill Education Pvt. Ltd., New York

Cooper PI (1973) The maximum efficiency of single-effect solar stills. Sol Energy 15:205–217

Cooper PI (1969) The absorption of solar energy radiation in solar stills. Sol Energy 12:133

Cox DR (1951) The productivity of machines requiring attention at random intervals. J. Roy. Stat. Soc. B XIII, 65–82

Dev R, Tiwari GN (2012) Annual performance of evacuated tubular collector integrated solar still. Desalination Water Treatment 41:204–223

Dubey A, Kumar S, Arora A (2021) Enviro-energy-exergo-economic analysis of ETC augmented double slope solar still with 'n' parallel tubes under forced mode: environmental and economic feasibility. J Clean Prod. https://doi.org/10.1016/j.jclepro.2020.123859

Duffie JA, Beckman WA (2006) Solar engineering of thermal processes. Hoboken- Wiley, New Jersey

Dunkle RV (1961) Solar water distillation: the roof type solar still and a multiple effect diffusion still, International Developments in Heat Transfer ASME 895–902

Dwivedi VK, Tiwari GN (2008) Energy and exergy analysis of single and double slope passive solar still. Trends in Applied Sciences and Research 3(3):225–241

IAPWS (1996) International association for the properties of water and steam, for the thermodynamic properties of ordinary water substance for general and scientific use, 1996

IAPWS (2008) International association for the properties of water and steam, the viscosity of ordinary water substance, 2008

Issa RJ, Chang B (2017) Performance study on evacuated tubular collector coupled solar stillin west texas climate. Int J Green Energy. https://doi.org/10.1080/15435075.2017.1328422

Kalogirou S (2003) The potential of solar industrial process heat applications. Appl Energy 76:337–361

Kalshian R (2008) Energy versus emissions: the big challenge of the new millennium. Info Change News and Features, www.infochangeindia.org/agenda5_01.jsp, accessed 21 March 2008

Koffi PME, Andoh HY, Gbaha P, Toure S, Ado G (2008) Theoretical and experimental study of solar water heater with internal exchanger using thermosiphon system. Energy Convers Manage 49:2279–2290

Kumar S, Dubey A, Tiwari G (2014) A solar still augmented with an evacuated tube collector in forced mode. Desalination 347:15–24

Liu BYH, Jordan RC (1960) The interrelationship and characteristic distribution of direct, diffuse and total solar radiation, Solar Energy 4(3)

Liu BYH, Jordan RC (1962) Daily insolation on surfaces tilted towards equator. ASHRAE J 3(10):53

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, Oxford, U. K.

Mittal ML, Sharma C, Singh R (2014) Decadal emission estimates of carbon dioxide, sulphur dioxide and nitic oxide emissions from coal burning in electric power generation plants in India. Environ Monit Assess 186:6857–6866

Morrison G, Tran N, McKenzie D, Onley I, Harding G, Collins R (1984) Long term performance of evacuated tubular solar water heaters in Sydney, Australia. Sol Energy 32:785–791

Morrison GL, Budihardjoa I, Behnia M (2005) Measurement and simulation of flow rate in a water-in-glass evacuated tube solar water heater. Sol Energy 78:257–267

Patel J, Markam BK, Maiti S (2019) Potable water by solar thermal distillation in solar salt works and performance enhancement by integrating with evacuated tubes. Sol Energy 188:561–572

Petela R (1964) Exergy of heat radiation. Journal of Heat Transfer, ASME 2:187–192

Petela R (2003) Exergy of undiluted thermal radiation. Sol Energy 74:469–488

Petela R (2010a) 2010. McGraw Hill, Inc, Engineering thermodynamics of thermal radiation for solar power utilization

Petela R (2010b) Radiation spectra of surface. Int J Exergy 7:89–109

Pissavi P (1982) Modelling of the dynamic behavior of a tank for solar storage with internal exchanger. Re´v Ge´n Thermique 246–247, 521–35

Rashidi S, Karimi N, Mahian O, Esfahani JA (2018) A concise review on the role of nanoparticles upon the productivity of solar desalination systems. J Therm Anal Calorim 135:1145–1159

Reddy KS, Sharon H (2017) Energy–environment–economic investigations on evacuated active multiple stage series flow solar distillation unit for potable water production. Energy Convers Manage 151:259–285

Reddy KS, Sharon H, Krithika D, Philip L (2018) Performance, water quality and enviro-economic investigations on solar distillation treatment of reverse osmosis reject and sewage water. Sol Energy 173:160–172

Roome J (2019) State and Trends of Carbon Pricing (2019). World Bank Group, Washington DC

Sadeghi G, Pisello AL, Nazari S, Jowzi M, Shama F (2021) Empirical data-driven multi-layer perceptron and radial basis function techniques in predicting the performance of nanofluid-based modified tubular solar collectors. J Clean Prod 295:126409

Sampathkumar K, Arjunan T, Senthilkumar P (2013) The experimental investigation of a solar still coupled with an evacuated tube collector. Energy Sources, Part A 35:261–270

Sato AI, Scalon VL, Padilha A (2012) Numerical analysis of a modified evacuated tubes solar collector. International Conference on Renewable Energies and Power Quality (ICREPQ'12)

Sharon H, Reddy KS, Krithika D, Philip L (2017) Experimental performance investigation of tilted solar still with basin and wick for distillate quality and enviro-economic aspects. Desalination 410:30–54

Singh AK (2020) An inclusive study on new conceptual designs of passive solar desalting systems. Heliyon 7:e05793

Singh AK, Samsher, (2020) Analytical study of evacuated annulus tube collector assisted solar desaltification system: A review. Sol Energy 207:1404–1426

Singh AK, Samsher, (2021) A review study of solar desalting units with evacuated tube collectors. Journal of Cleaner Production 279:123542

Singh AK, Samsher, (2021b) Material conscious energy matrix and enviro-economic analysis of passive ETC solar still. Materials Today: Proceedings 38:1–5

Singh AK, Samsher (2021b) Tech-en-econ-energy-exergy-matrix (T4EM) observations of evacuated solar tube collector augmented solar desaltification unit: a modified design loom. Materials Today: Proceedings, https://doi.org/10.1016/j.matpr.2021.09.088

Singh AK, Yadav RK, Mishra D, Prasad R, Gupta LK, Kumar P (2020) Active solar distillation technology: a wide overview. Desalination 493:114652

Singh AK, Samsher, (2022) Techno-environ-economic-energy-exergy-matrices performance analysis of evacuated annulus tube with modified parabolic concentrator assisted single slope solar desalination system. Journal of Cleaner Production 332:129996

Singh RV, Kumar S, Hasan M, Khan ME, Tiwari G (2013) Performance of a solar still integrated with evacuated tube collector in natural mode. Desalination 318:25–33

Tiwari AK, Tiwari GN (2007) Annual performance analysis and thermal modelling of passive solar still for different inclination of condensing cover. Energy Res 31(4):1358–1382

Tiwari G, Shukla S, Singh I (2003) Computer modeling of passive/active solar stills by using inner glass temperature. Desalination 154:171–185

Tiwari GN (2014) Solar energy: fundamentals, design, modelling and applications. Narosa Publishing House, New Delhi

Tiwari GN, Mishra RK (2012) 2012. RSC Publishing Cambridge, UK, Advanced renewable energy sources

Tiwari GN, Raj K, Maheshwari KP, Sawhney RL (1992) Recent advances in solar distillation. In: International Journal of Solar Energy and Energy Conversion. Wiley Eastern, New Delhi, pp. 32e149 (Chapter 2)

Tiwari GN, Tiwari A, Shyam, (2016) Handbook of Solar Energy. Theory, Analysis and Applications, Springer Nature, Singapore

Tiwari GN, Yadav JK, Singh DB, Al–Helal, I.M., Abdel-Ghany, A.M., (2015) Exergoeconomic and enviroeconomic analyses of partially covered photovoltaic flat plate collector active solar distillation system. Desalination 367:186–196

Tsatsaronis G, Lin L, Pisa J (1993) Exergy costing in exergoeconomics. J. Energy Resour. – ASME 115, 9–16

Tsatsaronis G, Park H (2002) On avoidable and unavoidable exergy destructions and investment costs in thermal systems. Energy Convers Manage 43(9–12):1259–1270

Tsatsaronis G, Winhold M (1985) Exergoeconomic analysis and evaluation of energy conversion plants – a new general methodology. Energy 10:69–94

Watt M, Jonhnson A, Ellis M, Quthred N (1998) Prog Photovolt Res Appl 6(2):127–136

Xu L, Liu Z, Li S, Shao Z, Xia N (2019) Performance of solar mid-temperature evacuated tube collector for steam generation. Sol Energy 183:162–172

Yari M, Mazareh AE, Mehr AS (2016) A novel cogeneration system for sustainable water and power production by integration of a solar still and PV module. Desalination 398:1–11

Yousef MS, Hassan H, Sekiguchi H (2019) Energy, exergy, economic and enviro economic (4E) analyses of solar distillation system using different absorbing materials. Appl Therm Eng 150:30–41

Yousef MdS, Hassan H (2019) Assessment of different passive solar stills via exergoeconomic, exergoenvironmental, and exergoenviroeconomic approaches: a comparative study. Sol Energy 182:316–331

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Environmental Science and Pollution Research
https://doi.org/10.1007/s11356-021-17974-6

REVIEW ARTICLE

Plastic waste management practices pertaining to India with particular focus on emerging technologies

Sampriti Kataki[1] · Krithika Nityanand[1,2] · Soumya Chatterjee[1] · Sanjai K. Dwivedi[1] · Dev Vrat Kamboj[1]

Received: 13 August 2021 / Accepted: 2 December 2021
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

Abstract
Under the parent petrochemical industries, plastic industry is proliferating enormously over the past several years globally due to its advantages in terms of weight, robustness, expense, versatility, and durability. Due to the diversified consumer base representing varied climate zones, food habits, and standards of living, the generation and growth opportunities for the plastic industry in India are particularly distinct and humongous. The present work extensively reviews the Indian plastic industry with primary focus on the evolving technologies for plastic waste valorization encompassing their level of utilization, technology readiness, and progress achieved at R&D level. The study attempts to recognize different issues related to technology, recycling, policy, research, regulation that should be given attention to formulate an improved plastic waste management strategy in the region. Though significant shares of waste plastics in the country are processed by traditional practices, state-of-the-art technologies primarily plastic to oil conversion, in road making and in cement manufacturing, are being deployed at increasing rate. Action to tackle the problem of plastic contamination in India will need to adopt a pan India strategic consensus/concurrent approach for effective waste collection and segregation with active participation of urban local bodies, fixing the role of the informal sectors, investment for reliable technology adoption with skilled manpower for operation, adoption of circular economy schemes involving plastic waste co-processing, and providing support to work on R&D for better penetration of the proven plastic valorization options along with their environmental and social implications.

Keywords Plastic · Pollution · Waste · Management · India · Technology

Introduction

When compared to other materials, the plastics industry, which is part of the parent petrochemical industry, has grown tremendously in recent years. This is due to the advantages that plastics have in terms of weight, robustness, cost, versatility, profitability, durability, and ease of technological adaptation (OECD 2018; Kumar and Agarwal 2020). Due to the diversified consumer base representing varied climate zones, food habits, and standards of living, the generation and growth opportunities for the plastic industry in country like India are particularly distinct and humongous with a 16% annual growth in consumption (Lewis et al. 2018; Rajmohan et al. 2019). India is among the largest consumers of plastic products in the world and each year generates approximately 9.4 million tons of plastic waste (MoHUA 2019). Plastic industry of the country presently comprises 15 large polymer factories, nearly 30,000 manufacturing units, and over 7000 recycling units employing more than 4 million people (Jaganmohan 2020; IBEF 2021). According to the GDP 2020 ranking, India is the world's third largest economy in the world in terms of purchasing power parity, after the USA and China, which clearly calls for the plastic industry to boom (Economic Times, India 2020). India's massive boom in plastic consumption is attributed to fast demographic growth, perpetual growth in consumerism, industrial progress, rapid urbanization, new consumption trends aided by spread of supermarket chains, increasing inclination towards E-commerce, and food delivery applications (Pillai, 2021). Its large networks of unlicensed manufacturing units manufacture surplus quantities of low-quality,

Responsible Editor: Philippe Garrigues

✉ Soumya Chatterjee
drlsoumya@gmail.com

[1] Biodegradation Technology Division, Defence Research Laboratory, DRDO, Tezpur 784001, Assam, India

[2] Delhi Technological University, Shahbad, Daulatpur, Delhi 110042, India

Published online: 22 January 2022

cheap plastic bags, the key reason why India produces an enormous amount of plastic waste (Banerjee 2019).

Plastic is the biggest source of environmental contamination due to its widespread domination, uninterrupted demand, and substantial interference with the food chain and environmental components. Like much of the planet, India is struggling to formulate strategies for managing the growing volume of omnipresent plastic waste in an effective way (MoEFCC 2018). The issue of managing plastic wastes in India is particularly acute owing to the limited availability of structured waste management schemes and the leakage of plastic into the atmosphere at various points of manufacturing and use (Dhanshyam and Srivastava 2021). With the present development patterns, and as a nation that relies heavily on the informal waste industry, it is facing an important challenge in tackling domestic plastic pollution. According to a UNEP (United Nations Environment Programme) study in 2018, the country is among the few other countries that have initiated absolute or partial national bans in their jurisdictions (Srivastav 2019). As of 2019, around 18 states of the country have enforced a ban on plastics bag (Economic Times, India, 2019a). Despite attempts for addressing the problem of plastics, including introducing bans, levies, penalties, extended producer responsibility, and alternatives to plastics, the implementation of these measures seems to be minimal or ineffective at ground level (MoEFCC 2018). Although there is a positive attitude towards combating plastic pollution, the actual individual consumption behaviour in terms of plastic usage is quite poor due to the absence of outreach programmes about plastic pollution reduction (Singh and Mathur 2019). In view of these obstacles, India needs to ensure that any more rise in plastic use should take place within a circular economy and must not contribute to further waste and emissions. At the national and state (municipal) levels, the government is pushing trash segregation programmes through legislation and voluntary intervention, and numerous technologies for plastic disposal are being evaluated. A comprehensive and updated review of the management practices is very much necessary to understand present status of various technology utilization and future line of action, which has been done in the present study. The present work also reviews Indian plastic industry encompassing existing technologies in use for plastic waste management, current recycling status and major setbacks disrupting the recycling process, regulatory control on plastic waste management, and the issues that need attention to formulate an improved plastic waste management strategy in the country. To comprehend the current state of different technology usage and the future course of action, a thorough and up-to-date examination of current management practices is required, and this has been accomplished in the current research. The current work examines the plastic industry of the country, including the existing technologies in use for its management, the current recycling status and major setbacks that have disrupted the recycling process, and regulatory control over plastic waste management. Another primary focus of the work is to critically evaluate various emerging technology options for plastic waste valorization deployed in the country and to understand their level of utilization such as plastic waste conversion to oil, plastic as concrete component, its use in road construction, co-processing of plastic in cement kilns, and gasification of waste plastic and bioplastic production. Finally, an attempt is made to discuss the issues that require attention to formulate a better plastic waste management strategy for the nation along with a conceptual framework that highlights on the role of each stakeholder for improved plastic waste management strategy.

Status on plastic waste generation and traditional way of management

Plastic waste generation

Indian plastic industry comprises of approximately 2000 exporting companies, 150 plastic processing machinery producers, and over 30,000 processing units (MoEFCC 2018). Figure 1 gives the overall statistics related to plastic in India and Fig. 2 gives the overall structure of Indian plastic industry and plastic waste flow. The predominant polymer of the country is polyvinyl chloride (PVC). An analysis by Plastindia Foundation (www.plastindia.org) indicates the compound annual growth rate of the industry at a 10.5% during five financial year (FY 15 to FY 20 FY) reaching a yearly production of 22 million metric tons (Plastindia 2018). This increase may further be driven by different initiatives, being promoted by the government, like *Make in India*, *Skill India*, *Digital India*, and the *Swachh Bharat Abhiyan* (IINS 2016; Mohanty 2017). Presently, an Indian consumes approximately 8–11 kg of plastic per year, which is relatively low by global standards. Consumption levels are 38 in China, 50–65 in Europe, and 68–109 in the USA in kilogram per capita per year, with a world average of 28 kg (Bhattacharya et al. 2018a, b). According to the Central Pollution Control Board (CPCB), 8% of the gross solid waste comprises plastics (Bhattacharya et al. 2018a, b). The weighted average service life of plastic products in the country is around 8 years (Panda et al. 2010). This limited-service life reflects that a large share of the plastic consumed here is short-life goods such as packaging (Cleetus et al. 2013). Figure 3a gives the statistics related to plastic production capacity, actual production, export, import, demand, and demand growth of plastic in India and Fig. 3b shows the generation of plastic waste in the states and union territories during 2018–2019.

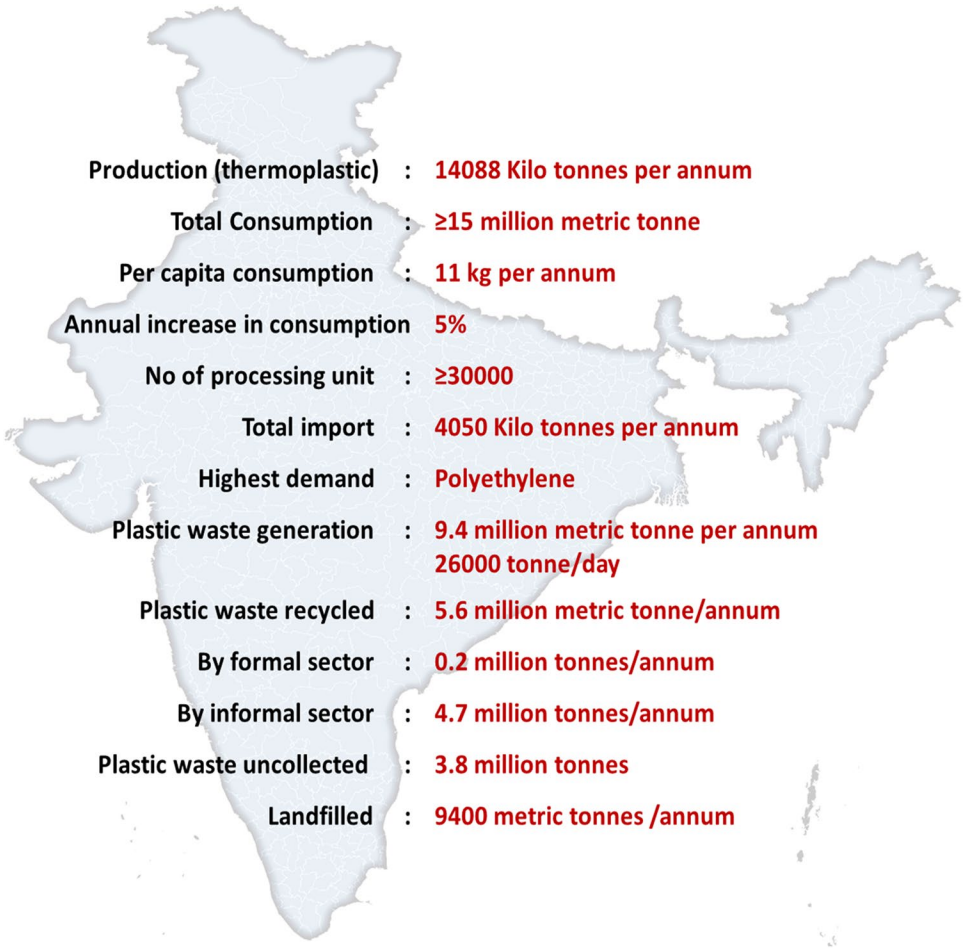


Fig. 1 Snapshot of statistics related to plastic in India

In 2011, the volume of plastic waste produced throughout the country was almost 7 million tons in a year, and by 2041, it is projected that more than 50 million tons per year waste will be generated (Jaganmohan 2020). As of 2017, according to the CPCB study, 60 major cities of India generate approximately 4059 tons of plastic waste per day, which gives an estimated total of 25,940 tons per day of plastic waste generation (CPCB 2018). However, a large share of it (~40%) remains uncollected (Times of India 2019). In India, the proportion of plastic waste ranges from 3.10% (Chandigarh) to 12.47% (Surat) in overall municipal solid waste (National Productivity Council report 2019). For plastic recovery, the informal sector plays the key role by processing an approximate 4.7 million tons of plastics each year, as opposed to just 0.2 million tons managed by the formal municipal waste management system (Lewis et al. 2018). In 2019, the CPCB found that 80% of the total plastic produced landed up in the garbage (Kumar and Agarwal 2020). Nearly half of the plastics produced are single-use plastics that eventually end up in landfills and drains and rivers flowing into the sea (FICCI 2019). In addition, Asian waterways particularly the Ganges and the Indus in India are two of the world's ten rivers, which collectively account for the 86% of the plastics that enter the ocean (Lebreton et al. 2017). Presently, 18 states/union territories (UTs) have enacted total prohibition on carry bags, 7 states have imposed the partial ban on plastic carry bags at religious/historical places, and 5 states/UTs have not imposed any ban on plastic carry bags except meeting thickness criteria (CPCB 2019).

Traditional management practices

In India, the primary disposal method of plastic waste is either by landfill or incineration. An engineered landfill is a bottom-lined pit where layers of solid garbage are piled high, compacted, and then covered for final disposal.

Landfill Landfill has been the primary waste disposal method in the world including plastic (Gopinath et al. 2020). The problems associated with a landfill site are well documented, which are further severe in developing economies like India (Taylor 1999; Mukherjee et al. 2015). However, the country often lacks what is called an ideal technically competent environmentally safe sanitary landfill sites equipped with baseliners, leachate, and collecting systems for gases (Singh and Ruj 2016; Swati et al. 2018;

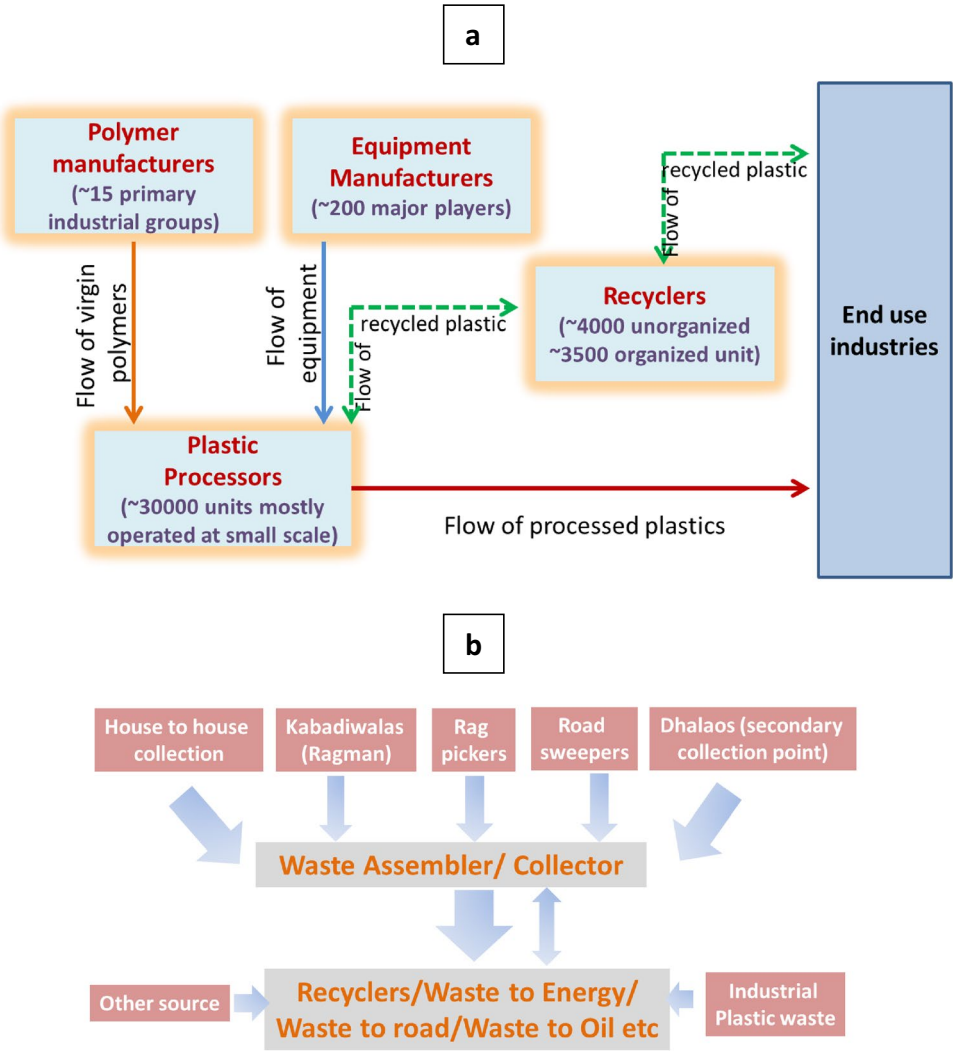


Fig. 2 **a** Structure of Indian plastic industry (Adapted from: MoEFCC Guideline on EPR, 2020). **b** Plastic waste flow diagram in India (Adapted from MoEFCC Guideline on EPR, 2020)

Bhattacharya et al. 2018a, b). Here, the landfill sites most of the times represent a site simply dumped up with waste pile in low-lying areas outside city conveniently overlooking the recommended requirements for a sanitary landfill or without taking any kind of sanitary measures (Kalyani and Pandey 2014; Bhattacharya et al. 2018a, b). The high population density with limited isolated free space and absence of proper land use planning are the other challenging factors, increasing the difficulty to identify a suitable landfill location (Dube et al. 2014; Kumar and Agarwal 2020). With the boundaries of cities expanding, numerous landfill sites are falling within the near proximity of human settlement. In most of the cases, no appropriate leachate or gas collection system, leakage prevention steps, such as compaction, waste levelling, and soil covering in landfills are found. As India lacks effective source segregation, the waste dumped at landfill sites are often mixed waste of all types; the indiscriminate burning of which poses threat by contributing to GHG emission (Verma et al. 2016).

Though there are many issues and problems related to landfilling in India, but still at present, landfill appears as the most obvious and extensively accepted practice for plastic waste disposal choice in the very foreseeable future (Swati et al. 2018; Cleetus et al. 2013). Many new landfill sites are being established, or the existing sites are being expanded, as most of these sites in the major metropolitan cities such as Delhi, Kolkata Mumbai, and Chennai already have exceeded their waste carrying capacity. However, it is gradually becoming a less attractive option because of the land costs and scarcity, production of obnoxious gasses, its high volume to weight ratio, and low biodegradability of widely used packaging polymers (Cleetus et al. 2013). Hence, is thus very much crucial that the rules of sanitation be practised through a scientifically engineered landfilling as much as possible to keep the harm minimum.

Incineration Presently, incinerators or waste-to-energy plants are ranked as a waste treatment modus operandi

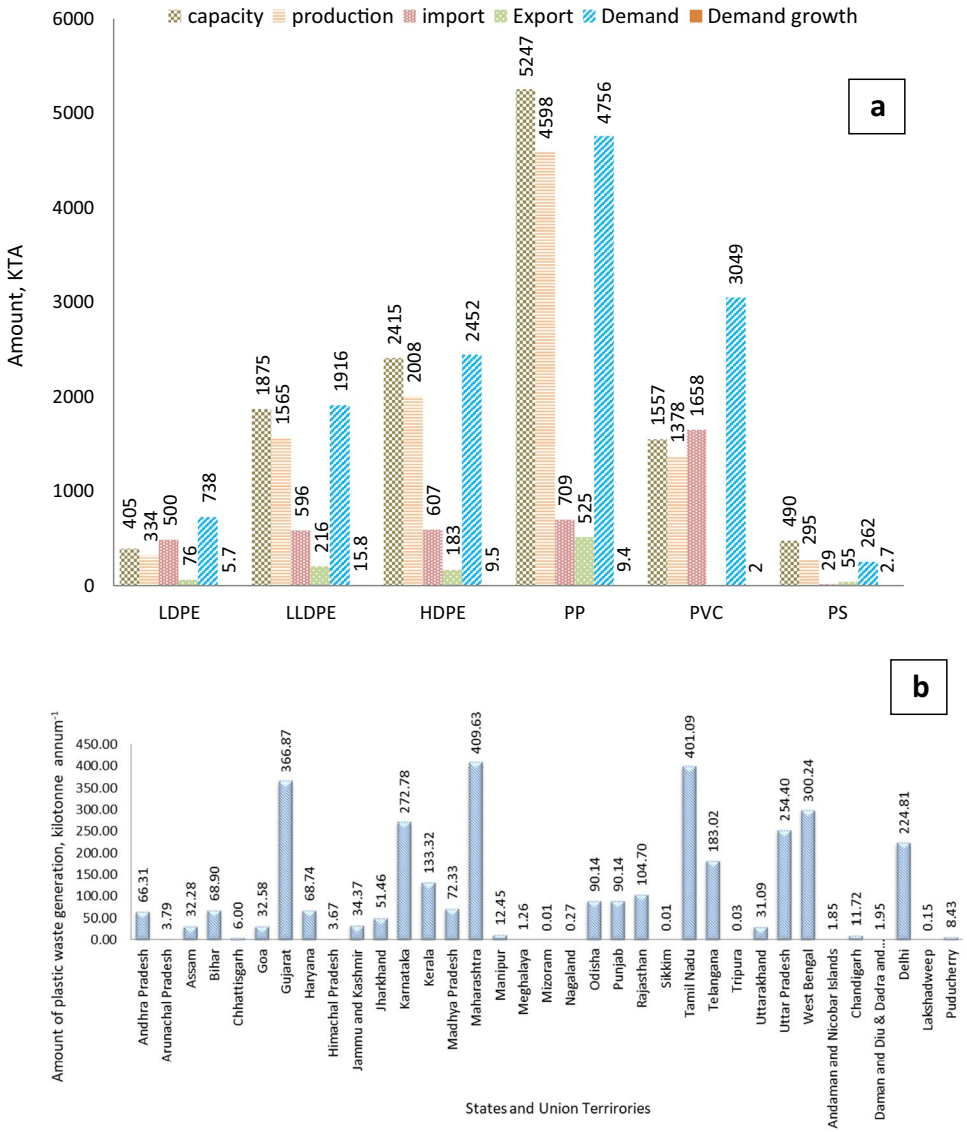


Fig. 3 **a** Statistics related to plastic capacity, production, export, import, demand, and demand growth of plastic in India. **b** Production of plastic waste in the states and union territories of India during 2018–2019 (Data source: CPCB annual report on PWM rules, 2019)

(Phanisankar et al. 2020). Incineration when applied for plastic wastes can also be an energy recovery method, as the plastics are being of petrochemical origin and possess high calorific value (36,500 kJ kg^{-1}) comparable with coal (28,500 kJ kg^{-1}) and fossil fuel (42,500 kJ kg^{-1}) (Banerjee et al. 2014). Incineration is thus becoming the alternative for landfills in different countries particularly in the advanced ones, a practice that is yet to be established extensively. However, the concept of waste-to-energy in incineration plant is being increasingly accepted as a suitable option for management of wastes. At present, utilization of only limited cases of successful co-processing of sorted plastic from municipal solid waste in cement kilns following CPCB regulations on common hazardous waste incineration is well documented. Its net contribution to the energy demand is yet rather low. Ministry of New and Renewable Energy, Government of India, estimated a potential contribution of approximately 0.42% by waste-to-energy projects to the current energy consumption by 2030 (Dube et al. 2014).

Although it is considered as one of the most utilized waste management solutions, waste incineration has some reservations and concerns about its appropriateness to waste characteristics pertaining to Indian conditions (Dube et al. 2014). Report says successful establishment of 199 urban, industrial, and agriculture waste-based waste-to-energy power projects in the country as of October 31, 2019 (Economictimes, India, 2019b). However, the current assessment shows hurdles and un-optimal performances of most of the non-biodegradable waste-based projects (The Hindu 2019). The underlying cause for the inefficiency or non-functionality of these plants is the mixed quality and composition of waste received by the plants. The quality of waste received is often unsegregated type with high inert content and low energy value and high-level moisture content causing the

plants to work non-optimally (Joshi and Ahmed 2016). Overall, incineration represents a more demanding waste management option in terms of construction, operation, and management, which makes the energy output of such system a relatively expensive energy source. As per latest assessment, the cost of power provided by thermal as well as solar sources is still far less costly (Indian Rupee ₹3–5 per unit) than the power output from waste to energy project (₹7 per unit) (The Hindu, 2019). Furthermore, the facility requires a steady supply of waste to make the overall process an economically feasible option. Though incineration can decrease the volume of plastic waste entering landfill significantly (by 80–95%), it comes with heavy environmental burden. Hence, the acceptance and success of incineration plants in developed countries show importance of transparency and data availability regarding parameters related to emission levels, environment and public health impacts, and economic and social outcomes. There is still a considerable scope for improvement on incineration-based plastic waste management in India; however, significant consideration in the quality of waste stream treated as well as filtration systems to minimize the pollutants from incinerators is required.

Environmental improvements in solid waste management are far from being satisfactory in most of the urban areas (Dhanshyam and Srivastava 2021). With over 26,000 tons of plastic generated every day, there are plenty of chances to turn plastic into a valuable resource provided adequate infrastructure is available. At present, India is promoting unlocking the capacity of waste-plastic as fuel and pilot such practice into a sustainable business model. Various processes have been utilized for the processing of resources from plastic wastes, such as pyrolysis, incineration, gasification, and palletisation. More details about the emerging plastic waste utilization and management techniques that are being explored in India are discussed under section 4.

Recycling of plastic in India: growing but many challenges to overcome

Plastic recycling is not only a waste management strategy, but it can also be seen as a way to implement the concept of industrial ecology, which means resource circularity in industrial processes (McDonough & Braungart 2002; Eriksen et al. 2019). Advanced technologies and systems for the collection, sorting, and reprocessing of recyclable plastics have created new opportunities for diverting most of the plastic waste from disposal (Hopewell 2009). Although recycling of the disposed plastic polymers is a commonly adopted practice, it still needs significant effort and resources. Several factors can make the plastic recycling practice more complicated, such as collection of plastic, separation of plastics by type, washing of the wastes/darts from the plastics, and subsequent waste disposal generated from the plastics, potential pollution of the plastics to bring it up as finished goods.

India holds one of the world's best recycling statuses which is considerably better than that of other nations. According to a 2013 estimate by the CPCB, 60% of plastic waste is recycled in the country, with the majority of the recycling taking place in the informal sector (MoEFCC, 2018; MoHUA, Govt. of India 2019). 65% of the plastic is recycled at authorized facilities, 15% by the unorganized sector, and 10% is reused at home (Chattopadhyay and Kumar, 2020). India recycles 38% more plastic than the global average of 20% (Gopinath et al. 2020). Presently, only HDPE and PVC plastics are recycled, whereas PS, PP, and LDPE are not recyclable (NPC 2019). The country has more than 3500 numbers of organized recycling units for plastics along with an additional 4000 unorganized recycling facility (Satapathy 2017). 3.6 million tons per year of plastic recycling is carried out currently providing an employment to almost 1.6 million people (Satapathy 2017). Plastic is recycled in compliance with the Bureau of Indian Standards: IS 1453: 1988 entitled 'The Guidelines for Recycling of Plastics' (CPP 2010). Recycled plastics are to be labelled as 'recycled' and the percentage of recycling content is to be shown. Even though the recycling rate of plastic in India is more satisfactory than many other countries, there are many setbacks in the impeding the success rate, which are discussed below.

(i) Indiscriminate disposal and collection The main cause of inefficient recycling process can be identified as lack of segregation at source of generation. Incompatible combination of waste flux leads to often non-separated collection or non-trustworthy feedstock following which the recycling process becomes inefficient (Kumar et al. 2017). Mixed waste, inappropriate sorting, and consequent incompatibility between various types of plastic polymers during the recycling are the major issues to be addressed (Banerjee et al. 2014). Plastic identification codes and container codes are rarely found on any packaging materials which lead to difficulty in identification of the kind of plastics among general waste (Satapathy, 2017). (ii) *Incompetent recyclers*: As mentioned previously, the recycling process in the country is mostly run by the informal sector (Kumar and Agarwal 2020). The quality of the plastic feedstock for secondary plastic production often remains inadequate or inferior due to the un-trained unauthorized recyclers associated with the process. Without proper training on categorically separating the plastic waste and without a proper sorting facility operated by trained workers, the recycled plastic products would usually be below the desired standard. Recycled plastics potential suppliers do not adequately invest in sorting and processing due

to limited profitability of these operations. (iii) *Use of additives*: Addition of dyes, fillers, and many different additives in plastics are often difficult to be removed economically from plastic polymer. Due to the presence of such additives, the plastic becomes viscous that causes problems during recycling (Hahladakis et al. 2018, ICPE, India 2020). (iv) *Incompetent technology*: Plastic recycling requires a certain level of technology and equipment to depolymerize various types of plastic in a product, and separate and extract target plastic for recycling. India lacks developed and tested technologies or appropriate technical competency for plastic waste recycling (Satapathy 2017). (v) *Product compatibility*: The recycled plastic products face fierce competition with the virgin plastic industry as well as the higher quality recycled plastic products from other advanced economies. In terms of the production cost, South Asia notably India produces some of the cheapest recycled plastic products with low performance in the world. (vi) *Cost compatibility*: Recycling plastic involves the collection and transportation cost of plastic feedstock that adds to the manufacturing cost of recycled plastics. As of 2019, the cost of purchasing virgin plastic is already cheaper (decreasing further due to technological advancement) than recycled plastics. (vii) *User acceptance*: Most recycled products are often of cheaper quality of short life span. Hence, Indian consumer still holds a psychological barrier towards recycled plastic and lacks the confidence and acceptance for recycled plastic often due to its inconsistency in quality (Bhattacharya et al. 2018a, b). (viii) *Inadequate implementation of regulations, minimum funding, and transparency*: A significant extent of initial capital is required to implement an effective recycling unit. Lack of adequate government policies, regulations and funding, and tax incentives to support entrepreneurs in setting up recycling facilities are observed as barrier in Indian context (v) *Limited resources*: Resources and electricity consumption by these units are demanding, often not economical.

Plastic recycling as a part of circular economy, the emphasis should be on minimizing impurities and losses via product design, technological advancement, and more focused waste management of plastic (Erkinsen et al. 2019). Higher resource recovery efficiency, source-separation, considering high number of target fractions in sorting, and efficient material recovery facility are found to be critical for improved circularity of plastic waste (Eriksen et al. 2019). Furthermore, diverse and fragmented initiatives to raise recycling levels at various stages of the supply chain must be coordinated (Hahladakis and Iacovidou, 2018). Mixed plastic waste is a critical problem to be tackled before its conversion to wealth as during the traditional recycling process, most plastic materials are not compatible with each other (Hopewell et al. 2009; Eriksen et al. 2019). This demonstrates that the treatment of plastic waste by traditional recycling technologies is very vulnerable to contaminants present and also needs certain additional configurations depending on the resin, colour, and transparency for the separation of plastic waste (Hopewell et al. 2009; Chintala et al. 2018; Wang and Wang, 2011). As a need of the moment, the plastic recycling industry needs a prompt technology upgradation to increase production quality and efficiency and to deal with complicated mixed plastic feedstocks (OECD 2018). In order for mixed plastic recycling to thrive, high-performance sorting needs to be undertaken to ensure that plastic types are segregated to high purity; and it is essential to further establish end markets for each polymer recycling stream (Hopewell et al. 2009). Land should be allotted for development of dedicated recycling plant having appropriate recyclability design with adequate storage. Within government regulations, provisions may be made to encourage investment by incentivizing recyclers and by promoting standardization of recycled products with tax advantages. Usage of recyclable materials in the manufacturing of products, developing the domestic market, selling of recycled items, and building interest for them are crucial. Demonstration of established recycling technology for mixed plastic waste, efficient quality validation of the recycling input feed, product testing, certification and accreditation, tax advantages for recycling-based production and sales, and building confidence of user through demonstration would further help for penetration and acceptance of the recycled plastic. The best recycling route for polymers should be multiple reuses, followed by repeated mechanical recycling as many times as possible till its quality is retained at acceptable range, which can be ideally termed as cascading of recycled plastics (Hahladakis and Iacovidou, 2018). The cascading of recycled plastic packaging into lower-grade goods is often advocated as the most efficient method of recovering its value, particularly when there is contamination and/or deterioration (Hahladakis and Iacovidou, 2018; Eriksen et al. 2019). Once the polymer quality deteriorates, their monomers can be recovered through a chemical path and can be re-polymerized resulting in a circular economy (Lamberti et al. 2020).

Plastic waste valorisation: emerging technology options being deployed and level of their utilization in India

A number of key drivers such as mandates and regulations within government policies, increasing eco-awareness among consumers, increased focus on sustainable waste management options, and search for cost and environmentally effective waste management technology are accelerating the research on valorization or alternative utilization aspects of plastic waste in India. Growing research has been centred on alternate methods such as plastic liquefaction,

pyrolysis, and co-processing (Brems et al. 2013; Gopinath et al. 2020). Co-processing entails utilizing waste products as substitute fuel or raw material in manufacturing processes for the recovery of energy and resources (CPCB guidelines in cement processing 2017). Co-processing as a method of managing plastic waste has been increasingly sought, as it is more environmentally safe and is an efficient waste management option as opposed to landfilling and incineration (CPCB guidelines in cement processing 2017). Although many more advances are required to embrace plastic as a business model for fuel, the country has made breakthroughs to make plastic into fuel (Swacchindia 2017). The emerging technology options which are relatively at nascent stage or are being explored at laboratory scale to deal with plastic waste in India are discussed below. Out of 36 states, 15 states/UTs are using plastic waste in bitumen road construction, 2 states/UTs are sending the plastic waste to waste to energy plants, and 3 states are using plastic waste as refuse derived fuel (RDF) (CPCB annual reports on PWM rules 2019).

Plastic waste conversion to oil in India: an expanding technology backed by R&D focus

Pyrolysis is a thermal waste treatment method carried out in an oxygen-free environment under increased pressure and temperature for a short duration. Pyrolysis represents a relatively well-organized and safe operational way for energy recovery from plastic waste and is one of the finest techniques for the conversion of mass to energy along with liquid (pyrolytic oil), solid (carbon black), and gaseous fractions with high energy values (Almeida and Marques 2016; Miandad et al. 2016; Phanisankar et al. 2020; Maafa et al. 2021). Plastic pyrolysis gives an average yield of 45–50% oil, 35–40% gases, and 10–20% tar depending upon the method employed (Wong et al. 2015). The plastic hydrocarbon oil contains paraffins, isoparaffins, olephines, naphthalenes, aromatics, and non-condensable gas with calorific value: the exact composition and relative proportion which depend on the quality of the plastic waste and the process conditions (Ademiluyi and Akpan 2007). There are two types of pyrolysis process: thermal and catalytic. Thermal cracking oils typically have a poor octane value, greater residual content at moderate temperatures, and need further processing for fuel usage (Lopez et al. 2011). Catalytic pyrolysis oil production is preferable than thermal pyrolysis as it significantly reduces pyrolysis temperatures and conversion time, improves conversion efficiency, allows to degrade plastics selectively with only cost and recovery of catalyst as areas of concern, and there is no effluent generated in the process (Walendziewski and Steininger 2001; Al-Salem et al. 2017). The heating value of the pyrolytic oil is equivalent to diesel or gasoline (44–46 MJ kg^{-1}) (Pandey et al. 2020).

In India, approximately 100 kilotons of plastic were processed in 2016–2017 through pyrolysis, which Phanisankar et al. (2020) projected to go up to 1500 kilotons in 2019–2020. There are about 25–30 operating RDF processing plants, having a capacity of around 3395 tonnes per day (Dube et al. 2014). In Pune, Maharashtra, India's first thermo catalytic depolymerization process–based plastic waste converting system was established in 2009 that converts plastic to high-grade fuel (The Hindu 2018). One ton of plastic generates approximately 600–650 l of petrol, 20–25% synthetic gas, and 5–10% residual char, which can be effectively used to fill bitumen roads (www.rudra-envsolution.com). In one of the successful development, Indian Institute of Petroleum, Dehradun, in 2014 developed a novel process for the transformation of plastic waste (polyethylene or polypropylene) into fuel or diesel. The technology has been shown to efficiently convert 1-kg plastic into 750 ml automobile fuel petrol (Bhattacharya et al. 2018a). Among all the states, Goa (two plants Bicholim and Sonsoddo setup in 2016) is actively working with in fuel transformation of plastic waste (Swachhindia 2017).

A number of research works have reported on plastic waste conversion to oils by thermochemical depolymerization (Chintala et al. 2018). These studies reported comparable performance of engine and emission behaviour while operating on plastic-derived oils as compared to neat diesel operation. Single-cylinder diesel engine using plastic oil has been reported to exhibit a stable performance with comparable brake thermal efficiency to that of diesel, though CO emission under the former was higher (Mani et al. 2009; 2011; Devaraj et al. 2015; Kumar et al. 2013; Chintala et al. 2018). However, several engine modifications are required for applying pure plastic pyrolytic oil that refuses the use of existing engines, whereas a mixture of typical fuel and plastic oil can be utilized with minor engine modification and without much affecting engine output, so oil blends are preferable to use as fuel over raw pyrolytic oil for diesel engines (Singh et al. 2020a). Chintala et al. (2018) operated a 3.7-kW CI engine by plastic oil and observed comparable brake thermal efficiency of the test engine at all engine loads. Ukarde and Pawar (2021) investigated the catalytic thermo liquefaction of poly-olefinic plastic waste using a selective, stable, non-toxic, and cost-effective copper-doped TiO_2 catalyst. The process yielded more than 80% of oil with 85% conversion efficiency under less stringent conditions along with least gas and char formation. Similarly, Singh et al. (2020b) reported yield of better-quality oil from plastic waste using fly ash and zeolite catalyst as compared to that without catalyst, though the yield was lower. The authors recorded higher brake thermal performance and lower energy consumption in engine tests using the 80:20 diesel fuel:plastic pyrolytic oil blend. Das et al. (2020) reported suitability of 20% plastic oil-diesel blend produced using

zeolite catalyst as an alternative fuel for diesel engine. Based on yield and thermo-physical properties of different catalyst such as silica, alumina, zeolite, barium carbonate, Cleetus et al. (2013) reported combination of silica, alumina, and zeolite (pore size ~4 A) as the optimum catalyst for plastic oil production. While assessing the properties of plastic oil from HDPE, LDPE, PP, and styrene for their suitability as diesel engine fuel, Mangesh et al. (2020) reported polypropylene pyrolysis oil having better physicochemical properties. There has been report that emission characteristics of a diesel engine fuelled with waste plastic oil can be considerably lowered by using oil blends as compared to neat plastic oil (Cleetus et al. 2013; Kaimal and Vijayabalan 2016). In India, co-pyrolysis of biomass and plastic has been increasingly explored as an effective waste management strategy because of higher energy efficiency, better stability of oil, and higher heating value; however, such technological intervention is mostly at lab scale till now (Kumar and Srinivas 2020). The enhanced efficiency is due to synergetic effect by reduced activation energy of co-pyrolysis in comparison to separate pyrolysis of biomass and plastic (Pandey et al. 2020). Operating temperature, composition of feedstocks, catalysts, reaction time, and residence time of gas are the main factors that influence the quality and proportion of final products (Pandey et al. 2020). Table 1 represents the developments at R&D level in India related to value addition to plastic waste through plastic waste to oil conversion.

However, problems regarding delayed ignition, higher heat discharge rate, and increased NOx remain to be addressed to increase plastic oil fuel feasibility for diesel engines (Kalargaris et al. 2017). Future research work should focus on reducing the unsaturated hydrocarbons in oil by hydrotreating and hydrocracking to improve their combustion performance in diesel engines (Mangesh et al. 2017). Cleetus et al. (2013) estimated that 64% of the production cost is accounted for the cost of catalysts. Hence, prospects for cheaper catalysts should be further investigated to reduce the production cost.

Plastic waste as concrete component: proven technology mostly deployed at pilot scale

The latest developments in concrete studies have been centred on integrating new and lightweight materials as alternatives to natural concrete aggregates to minimize energy consumption and transportation costs significantly (Bale 2011; Rai et al. 2012). The incorporation of plastics into concrete imparts different properties to the plastic and concrete blend (Saikia and de Brito, 2012). The addition of PET (polyethylene terephthalate) has been shown to enhance tolerance to sulphuric acid (Araghi et al. 2015), polyolefins improve tensile strength, PVC enhances resistance to chlorine (Senhadji et al. 2015), and PE (polyethylene) develops thermal resistance (Ali et al. 2018). Mainly, there are two types of plastic wastes, namely plastic aggregates (low bulk density supplement coarser aggregates and fine aggregates for lightweight concrete) and plastic fibres (can substitute conventional steel fibres for stabilization) used in building material (Gu and Ozbakkaloglu 2016). Considerable research was carried out in some countries like USA and UK on feasibility of plastic incorporation in concrete; however, there have been limited studies in India (Mahesh et al. 2016). Recovered waste plastics can be used successfully as recycled aggregates in cement concrete at a percentage 7.5–10%, which might prove to be a valuable downcycling option for waste plastics (Makri et al., 2019). High-strength concrete is possible with incorporation of small amount of plastic aggregate, particularly for concrete with a low water/cement ratio (Islam et al., 2016). Because plastic aggregate concrete has a good workability and a low water/cement ratio in the concrete mix, it may be used confidently in structural concrete (Islam et al., 2016). Incorporation of plastic waste as aggregate in concrete can improve various properties such as lowering of density, increase in toughness, improving abrasion resistance, and better thermal insulation properties, making these suitable for various practical applications (Saikia and de Brito, 2012). In India, the concept of utilizing plastic in concrete penetrated after 'plastone', a block made from a mixture of waste plastic and stone developed by R. Vasudevan from Tamilnadu in 2012 (MoEFCC 2018). A plastone block that consumes 300 plastic bags and six PET bottles can sustain heat and resist water percolation (MoEFCC 2018). Today, more than 12 states are successfully implementing his idea (Brindha et al. 2019).

The reports from India have shown that concrete manufactured from recycled plastic through partial elimination of the sand portion has the same compressive properties as structural concrete (Gopinath et al. 2020). In most of the studies related to concrete, pulverized plastics (in a range of 0.2–30%) is partially replacing fine aggregate and properties such as compressive strength, tensile strength, flexural strength, thermal resistance, and CO emission are tested to identify the optimum percentage. Raghatate (2012) found that increasing the percentage of fibre plastic in concrete decreased compressive strength, but improved tensile strength, with an optimal strength at 0.8% addition. Mathew et al. (2013) determined that a 22% replacement of recycled plastic for coarse particles is best for structural concrete. Subramani and Pugal (2015), Gawatre et al. (2015), and Khajuria and Sharma (2019) all found that plastic waste performed optimally as coarse aggregates in concrete at 10%, 7.5%, and 2.5% substitution, respectively, with additional increases reducing concrete strength. Vanitha et al. (2015) recommended replacing 4% of aggregates with plastic aggregates in paver blocks and 2% in solid blocks when using M20 grade concrete. At a percentage of 5–10%,

Table 1 Value addition to plastic waste through plastic waste to oil conversion: developments at R & D level in India

Objective	Type of plastic/method	Input/output (vol/kg)	Reference
Pyrolysis of HDPE	HDPE/pyrolysis	Highest liquid yield (79.08 wt%) at a temperature of 550^0C along with 24.75 wt% gaseous product	Kumar and Singh 2011
Yield and thermo-physical properties of oil with different catalyst such as silica, alumina, zeolite, barium carbonate	Polythene/catalytic pyrolysis	Combination of silica, alumina and zeolite (pore size ~4 A) optimum catalyst for plastic oil production	Cleetus et al. 2013
Performance of single cylinder diesel engine fuelled with plastic oil	Waste plastic/pyrolysis	Lower emissions using blends compared to neat plastic oil The smoke and NOX were reduced by 22% and 17.8% Plastic blend at 25% gave better performance in engine	Kaimal and Vijayabalan 2016
Use of plastic oil in a 3.7 kW rated power CI engine to assess performance, combustion and emission behaviour	Municipal mixed plastic waste/thermochemical depolymerization	Thermal performance is comparable with diesel Slightly Higher carbon based emissions (HC, CO, & smoke) Decreased NOx emission Viable fuel for CI engine	Chintala et al. 2018
Property analysis and engine performance	Mixed plastic waste/non-catalytic pyrolysis	Brake thermal efficiency increases with plastic proportion and specific fuel consumption decreases	Singh et al. 2020a
Pyrolysis using catalyst-fly ash and zeolite	LDPE plastic bottle, polybag/catalytic pyrolysis	Carry bags with zeolite catalyst yield liquid fraction of 22% at a conversion rate of 47%	Singh et al. 2020b
Properties of plastic oil from HDPE, LDPE, PP and styrene	HDPE, LDPE, PP and styrene	Superior physicochemical properties of polypropylene pyrolysis oil	Mangesh et al. 2020
Investigation of conversion of plastics and biomass in a bubbling fluidized bed reactor	Mixed plastic waste/thermal co-pyrolysis	Thermal co-pyrolysis or co-gasification of biomass and plastics at temperatures is recommended above 600 °C Proper biomass to plastics ratio should be used to avoid melting of plastics It is crucial to operate the reactor well above the minimum fluidization velocity to avoid defluidization	Pandey et al. 2020
Energy, exergy and emission analysis of diesel engine fuelled with	Shredded medical plastic wastes/thermo catalytic pyrolysis of	The 20% plastic oil diesel blend is suitable for diesel engine Significant improvement in brake thermal efficiency by plastic Lower NOx and HC emission under low load, became higher under high load	Das et al. 2020
Copper doped TiO_2 catalyst as a selective, robust, non-toxic, inexpensive	Catalytic thermo liquefaction	With minimum char and gas formation CTL can convert any combination of plastic waste into HC-Oil with minimum carbon loss and > 80% yield	Ukarde and Pawar 2021

pulverized/unpulverized polyethylene may be utilized successfully as a partial replacement for fine or coarse aggregate in concrete without significantly compromising the mechanical characteristics (Mahesh et al. 2016). HDPE may be used in lieu of cement in conjunction with fine aggregate to produce paver blocks with excellent thermal characteristics, flexural strength, water resistance, and compressive strength (Brindha et al. 2019). Manjunath (2016) showed the feasibility of using e-waste plastic particles as a partial substitute for coarse aggregates. Rai et al. (2012) showed that by adding a super plasticizer, the workability of waste plastic mix concrete may be increased further by 10–15%. With the increasing requirement of concrete in various construction industries, concrete industry in India seems to be an appropriate industry which can consume substantial amount of plastic waste. Incorporating plastic to concrete appears to be advantageous considering its lighter weight than competing products; decreased fuel usage during shipping; reliability; sustainability; chemical, water, and impact resistance; excellent thermal and electrical insulation properties; and relatively lower manufacturing cost (Siddique et al. 2008). Though studied extensively at lab scale, this technology with proven benefits needs to be promoted at field scale for a better penetration and acceptance.

Plastic waste use in road construction in India: proven technology with faster adoption at commercial scale

Another option that has been increasingly explored is through plastic use in roads and pavement making (Duggal et al. 2020). Asphalt or bitumen is a vital component of road building. However, due to its low tolerance to water, anti-stripping agents are required. The usage of waste plastic to substitute for a specific percentage of bitumen is an encouraging solution in the quest for alternatives to change the conventional structure of bituminous roads (Biswas et al. 2020). Low affinity between aggregates and bitumen is one of the causes for structural faults in standard road design. However, in contrast to standard building techniques, the binding between plastic-covered aggregates and bitumen is better, rendering better stability to plastic roads (Vasudevan et al. 2012; Behl et al. 2014; Sasidharan et al. 2019). Apart from being environmentally friendly, plastic inclusion in road even in limited amounts (approximately 5–10% by bitumen) helps significantly to boost the durability, strength, fatigue, and resistance to deformation under extreme climate and water-induced damage (Vasudevan et al. 2012; IRC 2013; Gopinath et al. 2020).

Globally, India is the pioneering country which appears to be the most experienced in using plastic waste for road construction (Sasidharan et al. 2019). Number of research works has been reported in India (Table 2). A Government of India order in November 2015 has made it compulsory to use waste plastic, alongside bituminous mix during construction of roads (Plastindia 2018; MoEFCC 2018). This idea came into practice after the method of plastic waste recycling for road making was introduced by R. Vasudevan, from Tamilnadu, India. The following guidelines from Indian Road Congress that came into force in 2013 (IRC:SP:98–2013) discuss the requirements and usage of waste-plastic in dry process wear, its benefits, use, production, transport, and storage (IRC 2013; Sasidharan et al. 2019). Jambulingam Street in Chennai was one of India's first plastic roads built in 2002 (Bhattacharya et al. 2018a). Since 2002, plastics roads are performing satisfactorily without potholes (Vasudevan et al. 2010, 2012). Out of 36 states, 15 states/UTs are using plastic waste for polymer bitumen road construction (CPCB annual report on PWM rules, 2019). India has over 100,000 km of roads made of waste-plastic and the technology is being adopted at a faster rate with time (Indiatimes.com 2020; Biswas et al. 2020). By using plastic waste in road, it has been estimated to save 1 tonne of bitumen (worth 414$) per kilometre (Hindustantimes 2020).

The process of plastic road making generally involves shredding of plastic, heating of shredded plastic along with aggregate mix and bitumen in the mixing chamber at a temperature range of 160 °C, which can then be employed in road construction (Biswas et al. 2020). For per kilometre of road construction, 1 tonne of plastic is used, which further decrease the emissions of carbon dioxide by 3 tonnes km^{-1} in contrast to the normal building techniques (Vasudevan et al. 2012). Fifty grams of bitumen is required for each kilo of stone, a one-tenth of which can be substituted with plastic waste, resulting in decreased bitumen consumption.

Recent analysis also confirmed greater stability and steady degradation of plastic road than traditional oil roads. Several laboratory and field studies from India already confirmed superior performance of plastic-laden bituminous roads (Heriawan 2020; Bale 2011; Behl et al. 2014; Bhoot et al. 2012; Biswas et al. 2020). Investigation by Behl et al. (2014) established that most environmentally hazardous PVC can be safely used as a modifier along with bitumen (at 3–5% of bitumen). Ponnada and Krishna (2020) reported 7.5% of PET plastic as the optimum blending percentage to be added to bitumen for improving the rheological parameters, while Hake et al. (2020) found 10% plastic blend as the optimum. As per the reports of Bansal et al. (2017), limited replacement of bitumen by waste plastic (4–10%) demonstrates an improvement in strength of up to 16% relative to standard mixture. Marshall stability is also found to be increased by 51–64% in plastic modified concrete mix than conventional mix (Bindu and Beena, 2010; Bhageerathy et al. 2014). Inclusion of plastic in asphalt mix is also beneficial in preventing excessive drain down from the mixture (Bindu and Beena 2010). Furthermore, incorporating

Table 2 Value addition to plastic waste through utilization in bituminous road construction in India

Objective	Key finding	Reference
Behaviour of bituminous concrete mix containing shredded waste plastic bags as binder	• Penetration and ductility decreased, softening point increased with increase in plastic • Higher fatigue life than plain bituminous concrete	Punith and Veeraraghavan, 2003
Performance evaluation of plastic modified bituminous concrete mixes	• Significant improvement in properties like marshall stability, retained stability, indirect tensile strength and rutting in modified bituminous concrete mixes • Improved Aggregate impact value, Los Angeles abrasion value and reduced water absorption capacity	Sabina et al. 2009
Benefits of stabilizing stone mastic asphalt (SMA) mixture with shredded plastic	• Polymer bitumen blend is a better binder compared to plain bitumen • Blend increased Softening point and decreased Penetration value with a suitable ductility • Reduction in porosity, moisture absorption of and improves soundness • Higher Marshall Stability value and suitable Marshall Coefficient	Bindu and Beena 2010
Basic properties and field assessment on plastic coated aggregates	• Improved flexible pavement properties • 10% of bitumen replacement is recommended • Reduce coast of pavement, increase strength, can undertake heavy traffic	Vasudevan et al. 2010
LDPE modified bitumen property	• Marshal Stability can be increased by up to 25% • Significant improvement in the aggregate properties	Rokade 2012
Polymer (PE, PP, PS) coated stone aggregate as a raw material for pavement	• Improved binding property and poor wetting property • Marshall Stability value in the range of 18–20 kN and the • Increased load bearing capacity by 100%, no pothole formation	Vasudevan et al. 2012
Comparative assessment of bitumen, aggregate and bitumen-aggregate plastic mix	• Optimum recommended plastic use 10% based on Marshal Stability test • Plastic improved impact value, softening value, penetration and ductility value	Sasane et al. 2015
Physical properties of Plastic mix bitumen pavement	• Aggregate Impact reduced by 22% at 10% plastic addition than control • Crushing Value reduced by 30–50% (indicates strong aggregates) • Los Angeles Abrasion increased by 20–30% indicating hardness of aggregate	Unde and Potnis 2015
Modification of road aggregate by plastic addition	• Marshall stability of plastic modified mix was 51% more than that of control • Suitable for increased load carrying capacity, reduces the absorption of moisture	Rajput and Yadav 2016
0,1,3,5 and 7% by weight of plastic addition in bitumen	• Bitumen quality improved through plastic addition • Optimum plastic content was 5%	Naskar et al., 2010
Properties of bitumen replacement by plastic bottle	• Improved rheological properties of the bitumen by plastic addition • Higher Marshall stability by 7.5% replacement of than 5% replacement • 7.5% of the PET plastic is the optimum percentage	Ponnada and Krishna 2020

plastic to asphalt mixtures requires no alteration to current plant facilities or technology (Chesner et al. 2002) and process does not evolve any gas at 130–180 °C temperature (Gawande et al. 2012; Gawande 2013).

The main obstacle to widespread usage of polymer modified bitumen in paving practice has been their tendency towards gross phase separation under quiescent conditions (Rajasekaran et al., 2013; Barad 2015). In a developing country like India with hot and extremely humid climate that adversely affects road network, implementing eco-friendly plastic roads will provide substantial economic savings that go for the repairing and maintenance of bituminous roads.

Our assessment clearly shows that considerable numbers of lab scale research work has been reported addressing properties of plastic-laden bituminous road; however, field tests reports are very few (Biswas et al. 2020). With government being keen on encouraging the setting up of small plants for mixing plastic in bituminous road construction, road infrastructure with optimal use of bitumen and plastic will in one hand reduce the menace of plastic pollution and will produce improved durable road network (Biswas et al. 2020).

Co-processing of plastic in cement kilns in India: a mandatory step gaining increasing acceptance

After China, India is the world's second-largest cement manufacturer. Cement plants are investigating the co-processing of scrap materials to displace conventional fuel coal and petroleum coke with a view to reducing carbon and fuel costs. In India, pyrolysis of plastic waste and energy recovery through co-processing in cement kiln are gaining increasing acceptance (Plastindia 2018).

In 2012, in Madhya Pradesh, test trial for performance evaluation of co-processing of plastic waste in cement kiln was carried out in ACC Limited cement manufacturer (Envis 2009). After successful lab and pilot scale tests, the government issued guidelines for *Co-processing of Plastic Waste in Cement Kilns under Plastic Waste Management Rules*, 2016 in 2017 (CPCB guidelines in cement processing 2017). The detailed guidelines elaborate the collection, segregation, transportation, and co-processing of plastic-waste in cement kiln. Presently, out of the 28 states and 8 union territories, 14 states/UTs are sending plastic waste to cement plants for co-processing (CPCB annual report on PWM rules 2019). Currently, only 46 cement plants out of 238 across the country use waste plastic as fuel (Times of India 2019). In India's cement plants, fuel from waste accounts for approximately 10% of fuel consumption (www.cemnet.com 2019).

A significant benefit of such strategy is that an existing facility can deploy the use of plastic waste without much structural modification. As the process attains a very high temperature (~2,000 °C), it ensures complete destruction (≤99.9%) of waste and harmful pollutants. The process also ensures complete exhaust gas scrubbing due to counter-current flow of raw material resulting in trapping of heavy metals, sulphur, and other contaminants within clinker (MoEFCC 2018). Further plastic waste utilization through this pathway would require no processing (segregation or washing) and low-end plastic waste that possesses challenges during waste management would be resources for the cement industry (ENVIS 2009; Poddar and Paranjpe, 2015).

Due to the use of energy saving updated technology, regulated emission, Indian cement industry is one of the most energy-efficient industries globally (Kukreja et al. 2020). It has been estimated that cement industry's thermal substitution rate could be increased by 5.5% by using 90% of single-use plastic as fuel and removing traditional fuel (Kukreja et al. 2020). This will further help to achieve a 25% substitution by the year 2025 and will promote the circular economy (Confederation of Indian Industry 2016). However, for better implementation of the technology for cement manufacturers that have been better aligned with coal plants and fly ash industry, there are some issues that required to be made clear by regulatory body such as waste collection and transportation strategy to the cement kiln, and paying and buying options for plastic waste user and producer considering all the stake holders (www.cement.com 2019).

Gasification of waste plastic: advanced technology yet to be well established

Gasification of plastic waste has recently been attracting increased attention as thermo-chemical environment-friendly recycling method alternative to conventional methods (Saebea et al. 2020). The process involves partial oxidation of organic matter at high temperatures (1200–1500 °C) under mild oxidizing conditions producing fuel gas consisting primarily of carbon monoxide and hydrogen and a condensed liquid (Scheirs 1998; Brems et al. 2013). The condensed liquid following subsequent separation and processing results in the manufacture of tar which has several potential applications (Kato et al. 2003). The main advantage of gasification is feasibility to treat heterogeneous polymers with limited scope of pre-treatment.

In India, generation of power through gasification of plastic waste is gaining attention due to its flexibility, durability, and economy. Positronicsa private farm from West Bengalin collaboration with Council of Scientific & Industrial Research (CSIR), Central Mechanical Engineering Research Institute (CMERI), Durgapur, has started producing power from plastic waste using plasma gasification technology (Positronics India 2021). Furthermore, coal and biomass co-gasification of plastics seems as a promising recovery route due to the positive effect on process efficiency and greater process flexibility. However, at present, limited work has been done to explore the suitability of plastic gasification in India. The future breakthrough of the method would need more research attempt to develop the appropriate equipment design and product optimization.

Development of bioplastic in India: evolving plastic alternative technology in need of go to market strategy for better market penetration

In an effort to reduce the unfavourable effects of conventional plastic use, worldwide researchers concentrated on the creation of safe, and organic and biodegradable alternative sources that lead to the development of bioplastics, where

100% carbon comes from renewable agricultural sources. Bio-plastics are plastic made of renewable materials that are either biodegradable, bio-based, or both. These plastics are derived from renewable biomass sources, such as vegetable fats and oils, corn, pea starch, microbes, and microalgae. (Lackner 2015). Although currently the cost of bioplastic production is considerably higher than conventional plastics, they provide advantages over petrochemical-based plastics such as faster degradation in environment, lower GHG emission, and reduced consumption of fossil fuels, saving of up to 40% energy in production (van den Oever et al. 2017). Biodegradable plastic is segmented into polyester, polylactic acid (PLA), polyhydroxy-alkanoate (PHA), starch blends, and others (Knowledge Sourcing Intelligence, India 2020). The non-biodegradable bioplastics are bio-PE (polyethylene), bio-PET (polythene terephthalate), bio-PA (polyamide), and bio polycarbonate.

Though bioplastic availability is marginal in Indian market, it is proliferating gradually with many industry players taking pioneering steps. As per recent estimation, the bioplastic market is expected to expand at a compound annual growth rate of 23.91% to reach US$754.648 million by 2025 from US$208.475 million in 2019 (Knowledge Sourcing Intelligence, India 2020). Presently, Bio-PET shares majority of bioplastic market and predicted to grow at a rapid rate. Given that biomass availability as raw material for plastic production is not an issue in India, a number of bioplastic manufacturing firms are in play such Envigreen, Ecolife, Plastobags, Earthsoul India, and Truegreen (Gupta 2017). Jammu and Kashmir (J&K Agro Industries Ltd in joint venture with Earthsoul India) is the leader to have built a dedicated integrated biopolymer manufacturing facility (capacity 960 metric tons per year) (Bioplastic news 2014).

Considerable concern in the expansion of biobased polymers is being garnered in India due to their diverse applications and biodegradability. Environmental awareness campaigns, feedstock availability, and government funding are providing bioplastic producers with significant support. In addition, major industry players are encouraged to join this market to comply with the increasing customer awareness. The other driving factor that is poised to facilitate the market growth is the continuous R&D efforts, increasing focus of sustainability that various companies have turned their attention towards, with the objective of reducing their carbon footprint and change in consumer preference for environment-friendly option. Furthermore, the wide utilization of bioplastics by various industry verticals such as construction, textiles, and automotive among others is also providing to the growing need for bioplastics.

This segment represents significant developments at lab scale encompassing range of raw materials, production, and technology as shown in Table 3. The greatest metric of weakness is the bioplastic usage, which is projected to be just 2% of the overall. India is dealing with a problem of biodegradable labelling and certification. Compostable and degradable plastics should be clearly labelled and applied in complementary ways (Song et al. 2009). Code mark as per Bureau of Indian Standards' Specification ISIISO 17088: 2008 entitled 'Specifications for Compostable Plastics' shall be affixed in all biodegradable plastic along with the mark 'compostable' or 'bio-degradable during composting' (CPPR 2010). However, it should be remembered that biodegradable or compostable plastics need appropriate composting conditions (ENVIS 2009). High cost of bioplastics, misconceptions or lack of clear understanding, infrastructure, and limited funding are acting as constraints to achieve its large-scale market penetration. There are a lot of uncertainties related to this evolving technology at present such as its post-use processing and end-of-life options, exact degree of biodegradability in comparison to traditional plastic in terms of decomposability, landfill conditions and their effect on food welfare, and competitive use of feedstock.

The technology options for plastic waste valorization discussed in the 'Plastic waste conversion to oil in India: an expanding technology backed by R&D focus' section to the 'Development of bioplastic in India: evolving plastic alternative technology in need of go to market strategy for better market penetration' section are being investigated and implemented in Indian conditions at varying level of success. Research by Gopinath et al. (2020) established an energy, economy, and environmental index (EEE index) to evaluate the suitability of various methods of plastic processing integrating different parameters. Their report included the inclusion of plastic waste in concrete buildings and road making as the most suitable plastic-waste management solutions with an EEE index value of 42.85%, following pyrolysis and liquefaction of plastic waste with a total of EEE index 23.08% each. Landfilling with a 7.14% EEE index have been shown as the least all-inclusive disposal process and must be decreased to minimal. However, while implementing and promoting the new and emerging technologies to deal with urban plastic waste problem, their environmental and social implications should also be taken into account (Agarwal et al. 2015). The emerging utilization practices that are being explored for plastic waste management in India are shown in Fig. 4 along with their benefits and issues.

Current regulations in India to curb plastic pollution

India has taken clear measures in the right direction to control the production of plastics and plastic waste. In recent years, the central and state governments have introduced and enforced different types of regulation, including waste management laws, bans and buybacks, and Expanded Producers

Table 3 Some recent work on bioplastic reported from India

Objectives	Bioplastic type/substrate used	Key findings	Reference
PHB by waste activated sludge	Molasses spent-wash	• Highest PHB yield (Yp/s) (0.184 g g^{-1} COD consumed) for deproteinized spentwash supplemented with nitrogen	Khardenavis et al. 2009
PHB from fermentative effluents	Fermentative effluents	• Acid-rich effluents from HBR yielded higher PHB productivity (25%) especially at lower substrate loading conditions	Mohan et al. 2010
Optimization of PHA synthesis by mixed culture		• Influencing factors: iron, glucose, VFA, nitrogen, phosphorous concentration, VFA composition, pH, and microenvironment • Microenvironment (80%) and pH (11%) showed significant contribution in PHA production	Mohan and Reddy 2013
PHA using newly isolated strain of *Bacillus tequilensis*	Wastewater with synthetic acids (SA) and acidogenic fermented food waste(AFW) from H_2producing anaerobic bioreactors	• Isolate showed ability to grow and accumulate PHA in both substrates • Higher PHA synthesis with SA (59% dw) compared to AFW (36% dwt) with good substrate removal (SA, 87%; AFW, 59%) • PHA has co-polymer [P(3HB-co-3HV)], hydroxy butyrate (80–90%) and hydroxy valerate (10–15%) • High dehydrogenase activity leads to high PHA production	Reddy et al. 2014
PHB production utilizing by *Botryococcus braunii Kütz*	Waste water	• Maximum yield (247 mg l^{-1}) of PHB dry weight at 60% concentration of sewage waste water, pH 7.5 at 40 °C	Kavitha et al. 2016
PHA under varying pH and reactor microenvironment (aerobic and microaerophilic)	Wastewater	• Maximum PHA production in microaerophilic operation (56%), compared to aerobic (34%) • Neutral pH showed better PHA synthesis (56%) than basic (44%) and acidic (28%) redox environments • Proteobacteria, Acidobacteria and Firmicutes took part in PHA production	Amulya et al., 2015
PHB From *Chlorella Pyrenoidosa*	*Chlorella Pyrenoidosa*	• 27% of PHB content in the 14 days old algae culture	Das et al. 2018
Characterization of PHA by *Natrinema ajinwuensis*		• *Natrinema ajinwuensis* accumulated 61.02 ± 0.68% PHA dw at 72 h (0.210 g l^{-1} h^{-1} volumetric productivity) • Composition: poly(3-hydroxybutyrate-*co*-3-hydroxyvalerate) with 13.93 mol% 3-hydroxyvalerate content and 284 °C degradation temperature	Mahansaria et al. 2018
PHA from municipal secondary wastewater sludge	Municipal secondary wastewater sludge	• PHA recovery of 0.605 g under optimized condition • PHA contained mcl PHA (58%) along with other diverse PHA	Kumar et al. 2018

Table 3 (continued)

Objectives	Bioplastic type/substrate used	Key findings	Reference
PHA production and characterization from marine algal species		• All the PHAs were thermally stable below 260 °C • Most thermally stable PHA: from *Leptolyngbya valderiana* • Primary strains involved *Chlorella* sp., (ii) *Oscillatoria salina*, (iii) *Leptolyngbya valderiana* and (iv) *Synechococcus elongatus*	Roja et al. 2019
PHA biopolymers in aerobic granules having mixed sludge in aerobic granular reactors (AGRs)	Oily wastewater	• Smaller but compact Micrococcus granules accumulated 0.47 mg PHA mg^{-1} • *81%* hydrocarbon removal efficiency *by Micrococcus aloeverae* strain SG002 (0.47 mg PHA mg^{-1} dw) • Granule extracted PHA was characterized as co-polymer P (3HB-*co*-3HV) having 3.5–4.5 of butyrates to valerates (PHB:PHV) ratios	Ghosh and Chakraborty 2020
PHB production utilizing various wastes	Rice husk and straw hydrolysate, dairy wastewater, and rice mill wastewater	• Maximum bio H_2 (1.53 mol H_2 mol^{-1} glucose) and PHB (9.8 g l^{-1}) production using 100% rice straw hydrolysate at pH 7 (under dark fermentation by *Bacillus cereus*)	Dinesh et al. 2020
PHA synthesis by *E. coli* Isolated from sludge soil	Sago molasses	• *E. coli* produced 27.1 g l^{-1} of PHAs from 33.26 g l^{-1} of biomass (81%) at 35 °C, pH 7.5 48 h of incubation	Narayanan et al. 2020
Bioplastic with different plasticizers and *Boswellia serrata* as filler	Jack fruit peel cellulose	• Pure nanocellulose was extracted from jackfruit peel through acid hydrolysis • Prepared biodegradable and compostable films by solvent casting nanocellulose and plasticizers along with novel filler, *Boswellia serrate*	Reshmy et al. 2021

PHA polyhydroxyalkanoate, *PHB* poly-β-hydroxybutyrate, *dw* dry weight

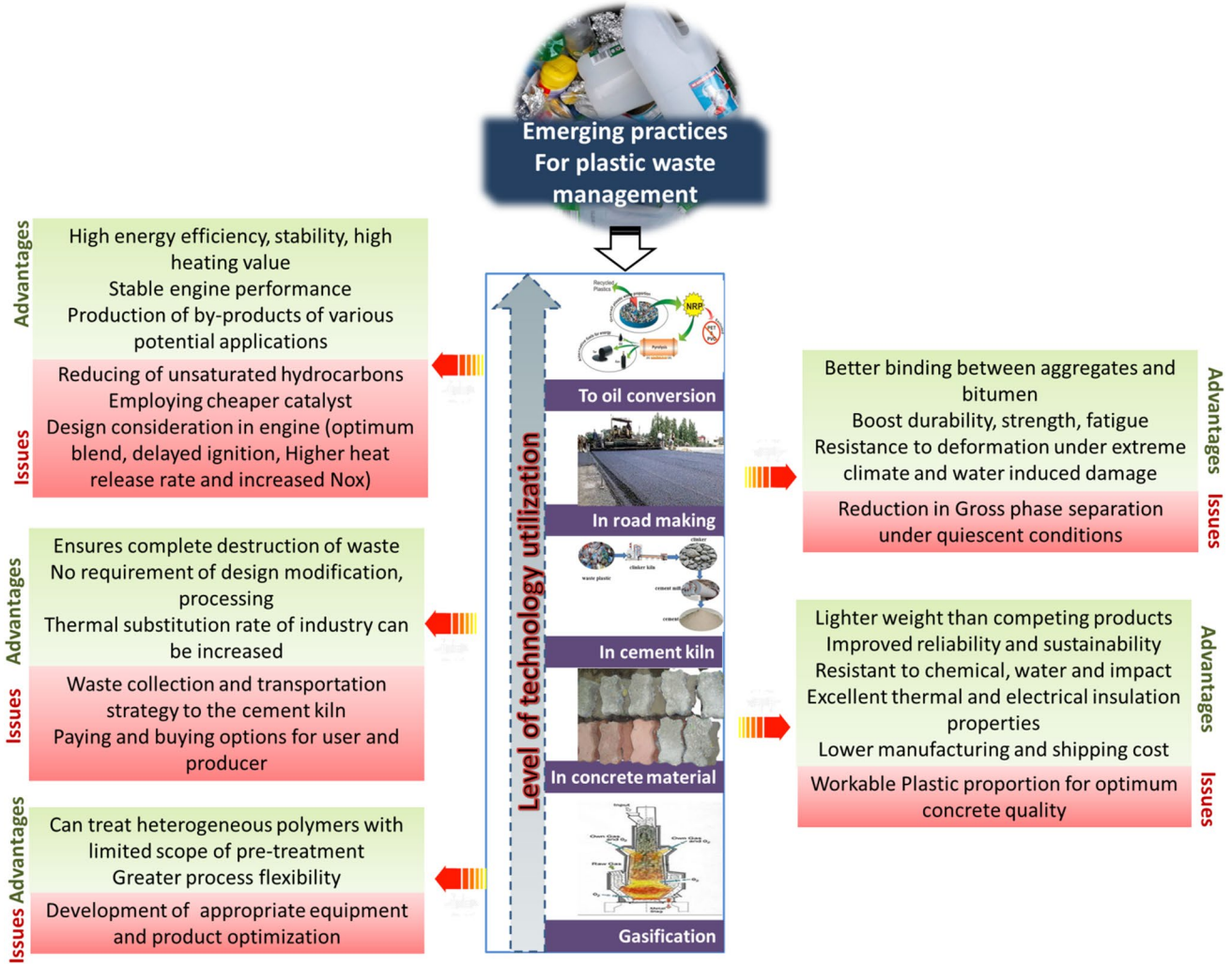


Fig.4 Emerging practices of plastic waste management in India: the advantages and issues involved

Responsibility (EPR) strategies, in order to counter the detrimental environmental and economic consequences of plastic waste. There are currently numerous laws in India to restrict and manage plastic waste and to govern the plastics industry, with new regulations constantly being implemented. Altogether, 18 states and union territories have executed some kind of ban on plastic manufacture, stock, sale, or use of plastic carry bag (MoHUA, Govt. of India 2019). However, closing the plastic loop appears to be a big obstacle where relevant data shows a lack of enforcement to implement the plastic regulations (Kumar and Agrawal 2020). Since the informal sector, a major role player in plastic waste dynamics is largely beyond the complete control of the official legislation, there is a growing challenge for the authority to track its operations.

Recycled Plastics Manufacture and Usage Rules, 1999 and 2003 (Amendment) was the first government legislation on plastic waste in India that aims to regulate the food item packaging in recycled plastics and manage the plastic littering issue. The rule permitted recycled and virgin polybags to be used for non-food applications but discouraged the same for food related usage. The rule amended in 2003 as *Plastics Manufacture, Sale and Usage Rules, 1999* prohibited the manufacture, storage, shipping, or selling of recycled or virgin plastic bags of less than 20 × 30 cm and 20 microns thickness (ICPE, India 2020). The rule needs recycling of plastics to be undertaken in accordance with the Bureau of Indian Standards specification ('The Guidelines for Recycling of Plastics': IS 14534: 1998) and the recycled plastics to be marked as 'recycled' (indicating the percentage of recycled material used). *Plastics (Manufacture, Usage and Waste Management) Rules, 2009* replaced the Recycled Plastics Use and Manufacturing Rules, 1999. Packets and containers manufactured with virgin plastic should be in natural shade without the use of colouring agents. Such plastics shall be manufactured for purposes except storage

and packaging food stuffs and only using pigments as per the standardized specification provided by the Bureau of Indian Standards' specifications: IS 9833: 1981 entitled 'List of pigments and colorants for use in plastics in contact with foodstuffs, pharmaceuticals and drinking water'. *Plastic Waste Management & Handling Regulations* that came into force in 2011 banned the multicoloured plastic bags because of the coloured agents used in these bags involve metals such as Cd and Pb that could be released to environment during use. It also put a pricing system for polybags to discourage its use. Government of India further superseded earlier *Plastic Waste Management & Handling Regulations of 2011* by *Plastic Waste Management Rules* on 18 March 2016. It stressed phasing out of multi-layered packages and the implementation of an expanded producers' (EPR) obligation for producers, importers, and brand owners to ensure that plastic goods are handled in environmentally sound way. This law suggests that any local authority would be accountable for developing a plastic waste storage, processing, and disposal facility. The ministry has notified the *Plastic Waste Management (Amendment) Rules, 2018*, which stipulates phasing out of non-recyclable multi-layered plastics (PWM rules 2018). However, 'non-recyclable multi layered plastic' was substituted by 'multi layered plastic, which is non-recyclable, non-energy recoverable or with no alternate use' through an amendment in March 2018 that effectively allows manufacturers of multi-layered plastics and plastic bag users to continue with business as usual (The Gazette of India 2018). *The single-use plastic (regulation) bill 2018* aims to create a mechanism to allow India to achieve its objective of removing plastic for one-use by 2022. On 5 June 2018, India declared its pledge to eradicate all single-use plastics by 2022 (Excell 2019).

The country enacted laws requiring EPR under Plastic Waste Management Rules 2016, a policy approach where producers must be responsible for the clean-up or recycling of their products until the end of its life (Excell 2019; MoEFCC 2020). The draft regulations suggest that producers can meet waste management goals over 5 years, beginning with 30% and moving to up to 90% in the fifth year after notification of the law. EPR includes management of the probable impacts of a product in every stage of manufacturing, distribution, usage, collection, reuse, recycling, reprocessing, and disposal. The brand, importers, and manufacturers along with state urban development departments have been delegated with the task of designing modalities for waste management schemes focused on EPR. Over the years, though the implementation has matured with realistic hopes, there is a need for a national framework for EPR, which covers the entire value chain, from polymer manufacturer to brand owners, importers and distributors, retailers, households, collectors, recyclers, and re-processors. Additional to the regulatory framework, the e-commerce companies that account for a major share of the single-use packaging plastic have been also instructed to reduce plastic packaging as much as possible. Such companies have introduced a range of initiatives aimed at minimizing the usage of plastics, including no box delivery idea, using software to maximize packaging form and volume. To come up with more stringent solution, industry should be pushed to focus on R&D and work on new packaging options.

Issues needing attention for improved plastic waste management in India

With the 'Make in India' initiative, initiated by the government with a specific emphasis on the chemical industry and an effort to transform the country into a global production hub, it is anticipated that tremendous growth in the plastic processing market, particularly in downstream industries, will take place (Mohanty 2017). With a step already being taken in that direction, plastics are bound to find tremendous use in the infrastructure space. However, the status of plastic waste management in the country is yet to attain a satisfactory level even after the existing ban on plastics below 50 microns, phasing out of multi-layered packaging and enacting EPR for producers, importers, and brand owners to ensure environment-friendly management of plastic products until the end of their lives. In the present assessment, following issues requiring attention for improving plastic waste management scenario in India could be summed up.

Reliable technology needs: As an urgent need of the moment, commercially feasible, technically viable, and proven plastic waste management and processing techniques need to be implemented in the country to reduce costs and improve process performance and environmental performance. In waste processing units, operators cannot separate out waste, due to dearth of equipment/technology/design designed to carry out waste separation (Enviro annotations 2019). Integrating advanced technology such as artificial intelligence or machine learning to the waste sorting or processing phase would help in effective waste segregation at reduced cost and time. The maturation of key technologies like big data and machine learning is opening new avenues for better coordination between different role players and stake holders of the supply and management chain. Assessment through Technology Readiness Level study will be helpful for feasibility of implementation of technology (Solis and Silveira, 2020).

Accurate assessment of plastic pollution threat at different sphere: India needs rigorous sampling and investigation for recognizing extent of plastic pollution (par-

ticularly microplastics) at various spheres to curb the plastic pollution menace (Raha et al. 2021). At what level various plastic originated contaminants (polychlorinated biphenyls, persistent organic pollutants, heavy metal, additives, etc.) are interacting with the biosphere and their combined eco-toxicological threats on the food network is required to be understood to steer any targeted pollution action plan.

Improved segregation/classification: The regulations on plastic waste management do exist but proper implementation and monitoring of these rules are still not being done in many parts of the country (Pandey and Manuja 2020). Absence of grossly ineffective source segregation, efficient solid waste management, especially in large cities, calls for a decentralized approach with segregated collection of dry and wet waste operated by skilled workers. Urban local body must formulate strategies to develop a detailed plan for waste assortment and outreach activities for sensitizing the citizens on segregating wastes and its benefits. In order to tackle the plastic pollution issue, a pan-India strategy for successful waste disposal and segregation with active involvement of the urban local authorities must be formulated.

Application of life cycle assessment tool: At global scale, a large number of life-cycle assessments (LCAs) have been carried out estimating environmental impacts of various options for end-of-life plastic management, though such studies are limited in Indian context. In a recent study by Aryan et al. (2019), the authors established incineration without energy recovery as the most detrimental plastic waste management option, while the recycling pathway as the most environmental friendly option. Generation of country-specific LCA data set is crucial considering various plastic types, input material, material flow, transport, distribution, and lifetime of plastic waste and most appropriate indicators to establish sound waste management practises while keeping into consideration diverse ecological impacts. From production, processing of plastics to application till end of life, the life cycle impact must be assessed for biodegradable plastic also.

***More emphasized R&D effort*:** More R&D effort at industry academia level should be emphasized addressing holistic issues of plastic life cycle. Redesigning products and distribution system to improve scope of to re-absorption into production chain through current manufacturing facilities, with few or no toxic by-products, appears to be beneficial. Future studies are necessary on development of packaging alternatives, technology development related to emerging technology for plastic pollution. Some of the specific research areas needing focussed approach could use alternative catalysts and their conversion efficiency during plastic conversion to fuel, optimum plastic proportion in plastic-based concrete, plastic biomass co-processing, biodegradability of bioplastic in real-life scenario, exploring plastic degrading microbes, and plastic composting, and understanding functions of the plastisphere in physico-chemical alteration of plastic waste is critical.

Improved regulations emphasizing radical resource efficiency: Despite initiatives to implement holistic waste management solutions, management of residual waste, particularly plastics, has posed an ever-increasing challenge. To alleviate the plastic disposal crisis, strengthening the existing unregulated collection and sorting recycling sector with appropriate policy and financial support from the government and private sector are of paramount importance. It is necessary to implement a holistic framework for EPR that explicitly identifies transparency and obligation of different stakeholders during a product's lifetime. Interventions at various manufacturing stages and waste management processes will ensure plastics' value and consistency and subsequently its reusability and recyclability. Policies related to plastic co-processing would be helpful to promote circular economy as well as plastic waste disposal crisis. Technology such as plastic in concrete, plastic in road making, and co-processing plastic in cement industry are already proven at pilot scale and seem to be ideal alternative for upscaling with the help of regulatory support. For plastic product with improved design, a centralized facility for testing and certification is essential to meet the global quality standards and commercialization of the product with govt. initiative. These regulatory policies will be better if paired with a growing public understanding of the risks and adverse effects of plastic pollutants and the significant effects on network use.

Waste management is a science that considers logistics, environmental impact, and social responsibility. Materializing a sustainable and effective plastic waste management framework a reality is a challenging, tough, and time-consuming task that can hardly be achieved by the government alone. The responsibility to ensure a successful and effective plastic waste management framework is shared among the different stakeholders, viz. waste generators, administration, local bodies, collectors and processors/recyclers, and manufacturers, which have a role to play. We present a framework (Fig. 5) that highlights on the role of each stakeholder to manage and reduce plastic waste. Accordingly, the role of top management (administration) lies mainly in regulating the volume of plastic that is entering the chain through various policy, promotion, and implementation, whereas the waste generators and processors have major role in limiting plastic waste entering dumping site/ocean. According to any waste management hierarchy, reducing and reusing waste

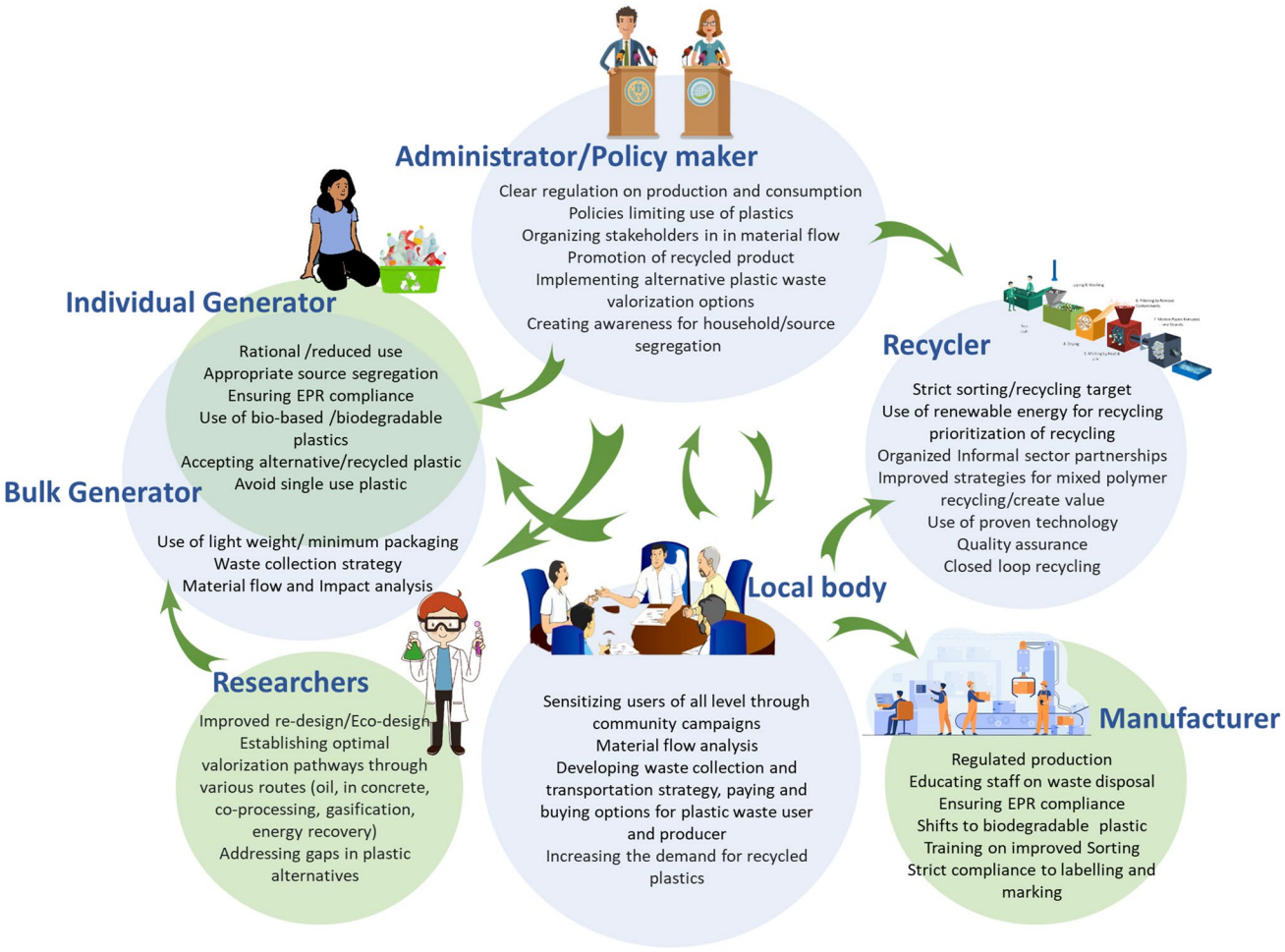


Fig. 5 Plastic waste management plan showing role of all the stakeholders in waste management programme

are preferable to its treatment and disposal, which could be primarily taken care by the generators/users. Rational usage of plastic, ensuring EPR compliance at individual level, appropriate source segregation, and higher acceptance of recycled plastic or plastic alternative at user's end would play the pivotal role in a successful waste management framework. Hence, in addition to a proper legislative framework, public knowledge and participation are must or vital element for a successful and sustainable plastic waste prevention and management strategy. Often, it is seen that people are aware of the negative consequences of poor waste disposal, yet the lack of positive attitudes in actual practice leads to chaos, which calls for an urgent need to rectify this gap in knowledge and practice. The local bodies can do their parts in reducing plastic waste by information dissemination aimed at raising public awareness, sensitization, and reducing stigma. It is critical to provide users the information, attitudes, awareness, and skills they need for execution of appropriate waste management procedures and thus able to trigger a whole societal transformation. Formulating a highly efficient collection and transportation strategy, material flow analysis within their jurisdiction, developing feasible paying and buying options for plastic waste user and producer (for co-processing), and increasing the demand for recycled plastics by the respective local authorities would enable to generate positive behaviours towards waste management. The waste processors can directly influence the plastic user by ensuring quality assurance and compliance to labelling and marking. The bulk manufacturer of plastic products should regulate production, educate and train staff on waste disposal and improved sorting, ensure EPR compliance, and shift to biodegradable plastic. Adopting a strict sorting/recycling target along with use of renewable energy for recycling, organized partnerships with informal sector, deploying improved strategies for mixed polymer recycling, and use of proven technology by the recyclers would aid to the development of a sustainable plastic waste management strategy. However, the scientific community needs to bridge the gap in actual practice of waste disposal and the targeted waste management framework by improved re-design/

eco-design of plastic products, by establishing the optimal valorization pathways through various routes such as plastic waste to oil, in concrete, co-processing, gasification, and energy recovery and by addressing gaps in plastic alternatives. The overall success of such framework will depend upon how these elements interact with each other and how these stakeholders could be brought together under a shared canopy of policies that are geared toward the development of a sustainable plastic waste management strategy.

Conclusion

The present assessment demonstrates explicitly the inadequacy of an effective plastic waste management solution in India. The lack of user awareness, expertise of people involved in collection and processing, inadequate technical competency, financial resources, and regulations contribute to the unsuccessful management of plastic wastes. A large proportion of waste plastics generated in the country are still processed by traditional practices. However, with stringent environmental legislations and a shift in people's attitude towards a greener solution, alternative technologies with less environmental impact are becoming the hope for the future. A range of state-of-the-art technologies are being explored that can change the present scenario faster when they are applicable at fields. The focus on plastic waste reuse is increasingly placed on emerging choices in road manufacture and the cement industry, which are backed by government mandates. Although prohibitions and taxation will limit the production and utilization of single-use plastics, regulatory policies will be more effective if coupled with spontaneous user involvement and increasing public insight into the risks and harmful consequences of plastics pollution. To resolve the plastic contamination problem, the role of the informal waste industry should be re-addressed. The responsibility must be equally shared among the different stakeholders to ensure a successful and effective plastic waste management framework. There is need of accurate evaluation of the danger of plastic waste in various spheres to curb the menace, to employ proven robust and reliable technologies, enhanced segregation/classification using smart technology, further research and development initiative with particular focus on co-processing and plastic alternative, application of life cycle approach tool for risk assessment of different technology options considering Indian condition, and improved legislation that underline the radical resource efficiency.

Acknowledgements The authors sincerely acknowledge the Director, Defence Research Laboratory, Tezpur. SK also acknowledges DRDO for post-doctoral fellowship.

Author contribution SK and NK prepared the primary manuscript. SC conceptualized and edited the manuscript. SKD and DVK supervised the work and revised the manuscript. All the authors have read and approved the final manuscript.

Funding The funding source for this project work was supported by DRDO fellowship.

Data Availability All data generated or analyzed during this study are included in this published article.

Declarations

Ethics approval and consent to participate Not applicable.

Consent for publication Not applicable.

Competing interests The authors declare no competing interests.

References

Ademiluyi T, Akpan C (2007) Preliminary evaluation of fuel oil produced from pyrolysis of low density polyethylene water-sachet wastes. J Appl Sci Environ Manage 11(3). https://doi.org/10.4314/jasem.v11i3.55069

Agarwal R, Chaudhary M, Singh J (2015) Waste management initiatives in India for human well being. European Scientific J. 105–127 SPECIAL/ edition ISSN: 1857 – 7881 (Print) e - ISSN 1857- 7431

Ali MR, Maslehuddin M, Shameem M, Barry MS (2018) Thermal-resistant lightweight concrete with polyethylene beads as coarse aggregates. Constr Build Mater 164:739–749. https://doi.org/10.1016/j.conbuildmat.2018.01.012

Almeida D, Marques MDF (2016) Thermal and catalytic pyrolysis of plastic waste. Polímeros 26(1):44–51. https://doi.org/10.1590/0104-1428.2100

Al-Salem SM, Antelava A, Constantinou A, Manos G, Dutta A (2017) A review on thermal and catalytic pyrolysis of plastic solid waste (PSW). J Environ Manage 197:177–198. https://doi.org/10.1016/j.wasman.2009.06.004

Amulya K, Jukuri S, Venkata Mohan S (2015) Sustainable multistage process for enhanced productivity of bioplastics from waste remediation through aerobic dynamic feeding strategy: process integration for up-scaling. Bioresour Technol 188:231–239. https://doi.org/10.1016/j.biortech.2015.01.070

Araghi HJ, Nikbin IM, Reskati SR, Rahmani E, Allahyari H (2015) An experimental investigation on the erosion resistance of concrete containing various PET particles percentages against sulfuric acid attack. Constr Build Mater 77:461–471. https://doi.org/10.1016/j.conbuildmat.2014.12.037

Aryan Y, Yadav P, Samadder SR (2019) Life Cycle Assessment of the existing and proposed plastic waste management options in India: A case study. J Clean Prod 211:1268–1283. https://doi.org/10.1016/j.jclepro.2018.11.236

Bale AS (2011) Potential reuse of plastic waste in road construction: a review. Inter J Adv Engineering Technol 2:233–236

Banerjee A (2019) Why does India report only half the plastic waste it generates every year? Available at: https://www.business-standard.com/article/current-affairs/why-does-india-report-only-half-the-plastic-waste-it-generates-every-year-119040200120_1.html. Accessed 26 May 2021

Banerjee T, Srivastava RK, Hung YT (2014) Plastics waste management in India: an integrated solid waste management approach. Handbook of Environment and Waste Management: Volume 2: Land and Groundwater Pollution Control, November, 1029–1060 https://doi.org/10.1142/9789814449175_0017

Bansal S, Misra AK, Bajpai P (2017) Evaluation of modified bituminous concrete mix developed using rubber and plastic waste materials. Inter J Sustain Built Environ 6(2):442–448. https://doi.org/10.1016/j.ijsbe.2017.07.009

Barad MM (2015) Use of plastic in bituminous road construction. J Inform Knowl Res Civil Eng 3(2):208–212

Behl A, Sharma G, Kumar G (2014) A sustainable approach: utilization of waste PVC in asphalting of roads. Constr Building Mater 54:113–117. https://doi.org/10.1016/j.conbuildmat.2013.12.050

Bhageerathy KP, Alex AP, Manju VS, Raji AK (2014) Use of biomedical plastic waste in bituminous road construction. Inter J Eng Adv Technol 3(6):89–92

Bhattacharya RRNS, Chandrasekhar K, Deepthi MV, Roy P, Khan A (2018b) Fact sheet on plastic waste in India. World Environmental Day. http://www.teriin.org/sites/default/files/files/factsheet.pdf. Accessed 21 May 2021

Bhattacharya RRN, Chandrasekhar K, Roy P, Khan A (2018a) Challenges and opportunities: plastic waste management in India. The Energy and Resources Institute https://www.teriin.org/sites/default/files/2018-06/plastic-waste-management_0.pdf. Accessed 28 May 2021

Bhoot K, Malviya K, Prajapat TK (2012) Potential reuse of plastic waste in road construction: a review. Inter J Scientific Res Eng 1(9):1–5

Bindu C, Beena KS (2010) Waste plastic as a stabilizing additive in Stone Mastic Asphalt. Inter J Engg Technol 2(6):379–387

Bioplastic news (2014) The future of bioplastics in India, https://bioplasticsnews.com/2014/07/26/the-future-of-bioplastics-in-india/. Accessed 11 May 2021

Biswas A, Goel A, Potnis S (2020) Performance comparison of waste plastic modified versus conventional bituminous roads in Pune city: a case study. Case Studies Constr Mater 13:00411. https://doi.org/10.1016/j.cscm.2020.e00411

Brems A, Dewil R, Baeyens J, Zhang R (2013) Gasification of plastic waste as waste-to-energy or waste-to-syngas recovery route. Nat Sci 5(6):695–704. https://doi.org/10.4236/ns.2013.56086

Brindha M, Raj CY, Nivas IS, Suliman SS (2019) Manufacturing of plastone using plastic and quarry dust. Inter J Eng Sci 20892.

Chattopadhyay S, Kumar N (2020) A circular plastics economy strategy for India, Confederation of Indian industry.Available at: https://sustainabledevelopment.in/wp-content/uploads/2020/06/1589884216Manage-Plastics-Report_web.pdf. Accessed 1 May 2021

Chintala V, Godkhe P, Phadtare S, Tadpatrikar M, Pandey JK, Kumar S (2018) A comparative assessment of single cylinder diesel engine characteristics with plasto-oils derived from municipal mixed plastic waste. Energy Converse Manage 166:579–589. https://doi.org/10.1016/j.enconman.2018.04.068

Cleetus C, Thomas S, Varghese S (2013) Synthesis of petroleum-based fuel from waste plastics and performance analysis in a CI engine. J. Energy 1–10. https://doi.org/10.1155/2013/608797

Confederation of Indian industry (2016) Promoting alternate fuels and raw material usage in indian cement industry. Approach paper for achieving 25% Thermal Substitution Rate in Indian cement Industry. http://www.ciiwasteexchange.org/doc/annexure_6.pdf. Accessed 11 May 2021

CPCB annual report on PWM rules (2019) Annual report for the year 2018–19 on implementation of plastic waste management rules. Available at: https://cpcb.nic.in/uploads/plasticwaste/Annual_Report_2018-19_PWM.pdf. Accessed 30 June 2021

CPCB guidelines in cement processing (2017) Guidelines for co-processing of plastic waste in cement kilns. Available at: https://cpcb.nic.in/uploads/plasticwaste/Co-processing_Guidelines_Final_23.05.17.pdf. Accessed 15 May 2021

CPCB (2018) Life cycle assessment (LCA) study of plastic processing products, Central Pollution Control Board (CPCB), https://cpcb.nic.in/uploads/plasticwaste/LCA_Report_15.05.2018.pdf. Accessed 29 Apr 2021

CPPR (2010) Waste regulation in India: an overview, Centre for Public Policy Research, https://www.cppr.in/wp-content/uploads/2012/10/Waste-Regulation-in-India-An-Overview-Bernard.pdf. Accessed 22 Apr 2021

Das AK, Hansdah D, Mohapatra AK, Panda AK (2020) Energy, exergy and emission analysis on a DI single cylinder diesel engine using pyrolytic waste plastic oil diesel blend. J Energy Institute 93(4):1624–1633. https://doi.org/10.1016/j.joei.2020.01.024

Das SK, Sathish A, Stanley J (2018) Production of biofuel and bioplastic from Chlorella pyrenoidosa. Materials Today: Proceedings 5(8):16774–16781. https://doi.org/10.1016/j.matpr.2018.06.020

Devaraj J, Robinson Y, Ganapathi P (2015) Experimental investigation of performance, emission and combustion characteristics of waste plastic pyrolysis oil blended with diethyl ether used as fuel for diesel engine. Energy 85:304–309. https://doi.org/10.1016/j.energy.2015.03.075

Dhanshyam M, Srivastava SK (2021) Effective policy mix for plastic waste mitigation in India using system dynamics. Resour Conser Recycl 168:105455. https://doi.org/10.1016/j.resconrec.2021.105455

Dinesh GH, Nguyen DD, Ravindran B et al (2020) Simultaneous biohydrogen (H2) and bioplastic (poly-β-hydroxybutyrate-PHB) productions under dark, photo, and subsequent dark and photo fermentation utilizing various wastes. Inter J Hydrogen Energy 45(10):5840–5853. https://doi.org/10.1016/j.ijhydene.2019.09.036

Dube R, Nandan V, Dua S (2014) Waste incineration for urban India: valuable contribution to sustainable MSWM or inappropriate high-tech solution affecting livelihoods and public health? Inter J Environ Technol Manage 17(2–4):199–214

Duggal P, Shisodia AS, Havelia S, Jolly K (2020) Use of waste plastic in wearing course of flexible pavement. In: Lecture Notes in Civil Engineering. Springer, 177–187. https://doi.org/10.1007/978-981-13-7615-3_16.

Economictimes India (2020) India retains position as third largest economy in PPP, global actual Individual consumption, capital formation. Accessed 21 Apr 2021

Economictimes, India (2019a) Industry seeks clarity on single-use plastic ban; officials say each state following its own policies is fragmenting industry. Available at: https://economictimes.indiatimes.com/news/company/corporate-trends/industry-seeks-clarity-on-single-use-plastic-ban/articleshow/71606595.cms?from=mdr. Accessed 15 Apr 2021

Economictimes, India (2019b) Government sets up over 56 MW waste-to-energy capacity in last 3 years. Available at: https://economictimes.indiatimes.com/small-biz/productline/power-generation/government-sets-up-over-56-mw-waste-to-energy-capacity-in-last-3-years/articleshow/72182964.cms. Accessed 12 May 2021

Enviro annotations (2019) India's solid waste data draws urgent attention. https://enviroannotations.medium.com/indias-solid-waste-data-draws-urgent-attention-a61b5c7f4ae4. Accessed 15 May 2021

Envis (2009) Envis Eco-Echoes, Indian centre for plastics in the environbment. http://icpe.in/envis_newsletter/s-Eco-Oct-Dec-09.pdf. Accessed 1 May 2021

Eriksen MK, Damgaard A, Boldrin A, Astrup TF (2019) Quality assessment and circularity potential of recovery systems for household plastic waste. J Industrial Ecol 23(1):156–168

Excell C (2019) World Resources Institute, 127 countries now regulate plastic bags. Why aren't we seeing less pollution? Available at:

https://www.wri.org/insights/127-countries-now-regulate-plastic-bags-why-arent-we-seeing-less-pollution. Accessed 7 Apr 2021
Chesner WH, Collins RJ, MacKay MH, Emery J (2002) User guidelines for waste and by-product materials in pavement construction (No. FHWA-RD-97-148, Guideline Manual, Report No. 480017). Recycled Materials Resource Center
FICCI (2019) Plastic Ban: Single-use plastic has no place on this planet. Available at: http://www.ficci.in/ficci-in-news-page.asp?nid=18656#:~:text=Of%20this%2C%2043%20per%20cent,and%20flow%20into%20the%20sea. Accessed 11 Apr 2021
Gawande A, Zamare G, Renge VC, Tayde S, Bharsakale G (2012) An overview on waste plastic utilization in asphalting of roads. J Eng Res Studies 3(2):1–5 (E-ISSN0976–7916)
Gawande AP (2013) Economics and viability of plastic road: A review. J Curr Chem Pharm Sc 3(4):231–242
Gawatre DW, Damal VS, Londhe SS, Mane A, Ghawate H (2015) Environmental issues of plastic waste use in concrete. Inter J Innovative Res Adv Engin 2(5):114–118
Ghosh S, Chakraborty S (2020) Production of polyhydroxyalkanoates (PHA) from aerobic granules of refinery sludge and Micrococcus aloeverae strain SG002 cultivated in oily wastewater. Intl Biodeterior Biodegrad 155:105091. https://doi.org/10.1016/j.ibiod.2020.105091
Gopinath KP, Nagarajan VM, Krishnan A, Malolan R (2020) A critical review on the influence of energy, environmental and economic factors on various processes used to handle and recycle plastic wastes: Development of a comprehensive index. J Clean Prod 274:123031. https://doi.org/10.1016/j.jclepro.2020.123031
Gu L, Ozbakkaloglu T (2016) Use of recycled plastics in concrete: a critical review. Waste Manage 51:19–42. https://doi.org/10.1016/j.wasman.2016.03.005
Gupta (2017) https://www.ecoideaz.com/expert-corner/pioneering-bioplastic-companies-india-bring-positive-change. Accessed 18 May 2021
Hahladakis JN, Velis CA, Weber R, Iacovidou E, Purnell P (2018) An overview of chemical additives present in plastics: migration, release, fate and environmental impact during their use, disposal and recycling. J Hazard Mater 344:179–199. https://doi.org/10.1016/j.jhazmat.2017.10.014
Hahladakis JN, Iacovidou E (2018) Closing the loop on plastic packaging materials: what is quality and how does it affect their circularity? Sci Total Environ 630:1394–1400. https://doi.org/10.1016/j.scitotenv.2018.02.330
Hake SL, Damgir RM, Awsarmal PR (2020) Utilization of plastic waste in bitumen mixes for flexible pavement. Trans Res Procedia 48:3779–3785. https://doi.org/10.1016/j.trpro.2020.08.041
Heriawan A (2020) Upcycling plastic waste for rural road construction in India: an alternative solution to technical challenges. Available at: https://think-asia.org/bitstream/handle/11540/11643/sawp-69-upcycling-plastic-waste-india.pdf?sequence=1. Accessed 19 May 2021
Hindustantimes (2020) 1 lakh km of road built using plastic waste, govt aims to double it. https://www.hindustantimes.com/india-news. Accessed 9 May 2021
Hopewell J, Dvorak R, Kosior E (2009) Plastics recycling: challenges and opportunities. Philos Trans R Soc Lond B Biol Sci 364(1526):2115–2126. https://doi.org/10.1098/rstb.2008.0311
IBEF (2021) Indian plastics industry and exports, India Brand Equity Foundation. Available at: https://www.ibef.org/exports/plastic-industry-india.aspx. Accessed on 1 May 2021
ICPE, India (2020) Additives for recycling and degradation of polymers. http://www.icpe.in/ncl/pdf/NCL%20Report_Chapter%204.pdf. Accessed 3 June 2021
IINS (2016) Indian plastics industry set to buck the global trend; set to grow 12% this year. India Infoline News Service. https://www.indiainfoline.com/article/news-top-story/indian-plastics-industry-set-to-buck-the-global-trend-set-to-grow-12-this-year-116091200141_1.html. Accessed 1 June 2021
Indiatimes.com (2020) This Indian genius created roads with plastic waste, now entire world uses his technique. https://www.indiatimes.com/technology/science-and-future/meet-padma-shree-r-vasudevan-who-made-roads-stronger-with-waste-plastic-that-saves-our-planet-505994.html. Accessed 2 June 2021
IRC (2013) Guidelines For The Use Of Waste Plastic in Hot Bituminous Mixes (Dry Process) in Wearing Course. Indian Road Congress. (IRC: SP: 98–2013). Available at: https://www.tce.edu/sites/default/files/PDF/IRC-Spec=Road-with-plastic-waste.pdf. Accessed 23 June 2021
Islam MJ, Meherier MS, Islam AR (2016) Effects of waste PET as coarse aggregate on the fresh and harden properties of concrete. Constr Build Mater 125:946–951
Jaganmohan M (2020) Plastic industry in India - statistics and facts. Available at: https://www.statista.com/topics/6902/plastic-industry-in-india/. Accessed 2 June 2021
Joshi R, Ahmed S (2016) Status and challenges of municipal solid waste management in India: a review. Cogent Environ Sci 2(1):1139434. https://doi.org/10.1080/23311843.2016.1139434
Kaimal VK, Vijayabalan P (2016) A study on synthesis of energy fuel from waste plastic and assessment of its potential as an alternative fuel for diesel engines. Waste Manage 51:91–96. https://doi.org/10.1016/j.wasman.2016.03.003
Kalargaris I, Tian G, Gu S (2017) Combustion, performance and emission analysis of a DI diesel engine using plastic pyrolysis oil. Fuel Proc Technol 157:108–115. https://doi.org/10.1016/j.fuproc.2016.11.016
Kalyani KA, Pandey KK (2014) Waste to energy status in India: a short review. Rene Sustain Energy Rev 31:113–120. https://doi.org/10.1016/j.rser.2013.11.020
Kato K, Nomura S, Uematsu H (2003) Waste plastics recycling process using coke ovens. J Mater Cycles Waste Manage 5:98–101. https://doi.org/10.1007/s10163-003-0089-3
Kavitha G, Kurinjimalar C, Sivakumar K, Kaarthik M, Aravind R, Palani P, Rengasamy R (2016) Optimization of polyhydroxybutyrate production utilizing waste water as nutrient source by *Botryococcus braunii* Kütz using response surface methodology. Inter J Biol Macromol 93:534–542. https://doi.org/10.1016/j.ijbiomac.2016.09.019
Khajuria A, Sharma P (2019) Use of plastic aggregates in concrete. Inter J Innov Technol Exploring Eng ISSN 2278–3075:9
Khardenavis AA, Vaidya AN, Kumar MS, Chakrabarti T (2009) Utilization of molasses spentwash for production of bioplastics by waste activated sludge. Waste Manage 29(9):2558–2565. https://doi.org/10.1016/j.wasman.2009.04.008
Knowledge Sourcing Intelligence, India (2020) https://www.knowledge-sourcing.com/report/india-bioplastics.market#:~:text=India%20bioplastics%20market%20is%20projected,zenith%20during%20the%20forecast%20period. Accessed 5 June 2021
Kukreja K, Sharma P, Mohapatra B, Saxena A (2020) Indian cement industry: a key player in the circular economy of India. In Enhancing Future Skills and Entrepreneurship. Springer, Cham. 181–192. https://doi.org/10.1007/978-3-030-44248-4_18
Kumar PK, Srinivas S (2020) Catalytic co-pyrolysis of biomass and plastics (polypropylene and polystyrene) using spent FCC catalyst. Energy Fuels 34(1):460–473. https://doi.org/10.1021/acs.energyfuels.9b03135
Kumar S, Singh RK (2011) Recovery of hydrocarbon liquid from waste high density polyethylene by thermal pyrolysis. Braz J Chem Eng 28:659–667. https://doi.org/10.1590/S0104-66322011000400011
Kumar S, Prakash R, Murugan S, Singh RK (2013) Performance and emission analysis of blends of waste plastic oil obtained by

catalytic pyrolysis of waste HDPE with diesel in a CI engine. Energy Convers Manage 74:323–331. https://doi.org/10.1016/j.enconman.2013.05.028

Kumar A, Agrawal A (2020) Recent trends in solid waste management status, challenges, and potential for the future Indian cities–a review. Current Research in Environmental Sustainability 2:100011. https://doi.org/10.1016/j.crsust.2020.100011

Kumar A, Samadder SR, Kumar N, Singh C (2018) Estimation of the generation rate of different types of plastic wastes and possible revenue recovery from informal recycling. Waste Manage 79:781–790. https://doi.org/10.1016/j.wasman.2018.08.045

Kumar S, Smith SR, Fowler G, Velis C, Kumar SJ, Arya S, Rena Kumar R, Cheeseman C (2017) Challenges and opportunities associated with waste management in India. Royal Soc Open Sci 4(3):160764. https://doi.org/10.1098/rsos.160764

Lackner M (2015) Bioplastics. Kirk-Othmer Encyclopedia of Chemical Technology. 1–41. https://doi.org/10.1002/0471238961.koe00006

Lamberti FM, Román-Ramírez LA, Wood J (2020) Recycling of bioplastics: routes and benefits. J Poly Environ 1-21. https://doi.org/10.1007/s10924-020-01795-8

Lebreton LCM, Van Der Zwet J, Damsteeg JW, Slat B, Andrady A, Reisser J (2017) River plastic emissions to the world's oceans. Nat Commun 8:1e10. https://doi.org/10.1038/ncomms15611

Lewis H, Retamal M, Atherton A (2018) Addressing plastic pollution in India: background paper prepared for Mumbai Business forum on Addressing Plastic Pollution in India, prepared for Stewart Investors. University of Technology Sydney, Australia, Institute for Sustainable Futures

Lopez A, Marco ID, Caballero BM, Laresgoiti MF, Adrados A (2011) Influence of time and temperature on pyrolysis of plastic wastes in a semi-batch reactor. Chem Eng J 173:62–71. https://doi.org/10.1016/j.cej.2011.07.037

Maafa IM (2021) Pyrolysis of polystyrene waste: a review. Polymers 13(2):225. https://doi.org/10.3390/polym13020225

Mahansaria R, Dhara A, Saha A, Haldar S, Mukherjee J (2018) Production enhancement and characterization of the polyhydroxyalkanoate produced by Natrinema ajinwuensis (as synonym) *Natrinema altunense* strain RM-G10. Inter J Biol Macromol 107:1480–1490. https://doi.org/10.1016/j.ijbiomac.2017.10.009

Mahesh M, Rao BV, Sri CS (2016) Re-use of polyethylene plastic waste in concrete. Inter J Eng Develop Res 4(4):693–702

Makri C, Hahladakis JN, Gidarakos E (2019) Use and assessment of "e-plastics" as recycled aggregates in cement mortar. J Hazard Mater 379:120776

Mangesh VL, Padmanabhan S, Ganesan S, Prabhudevrahul D, Reddy TDK (2017) May. Prospects of pyrolysis oil from plastic waste as fuel for diesel engines: a review. In IOP Conference Series: Materials Science and Engineering 197:012027

Mangesh VL, Padmanabhan S, Tamizhdurai P, Ramesh A (2020) Experimental investigation to identify the type of waste plastic pyrolysis oil suitable for conversion to diesel engine fuel. J Clean Prod 246:119066. https://doi.org/10.1016/j.jclepro.2019.119066

Mani M, Nagarajan G, Sampath S (2011) Characterisation and effect of using waste plastic oil and diesel fuel blends in compression ignition engine. Energy 36:212–219. https://doi.org/10.1016/j.energy.2010.10.049

Mani M, Subash C, Nagarajan G (2009) Performance, emission, and combustion characteristics of a DI diesel engine using waste plastic oil. Appl Therm Eng 29:2738–2744. https://doi.org/10.1016/j.applthermaleng.2009.01.007

Manjunath BTA (2016) Partial replacement of e-plastic waste as coarse-aggregate in concrete. Procedia Environ Sci 35:731–739. https://doi.org/10.1016/j.proenv.2016.07.079

Mathew P, Varghase S, Paul T, Varghase E (2013) Recycled Plastic as coarse aggregates for structural concrete. Inter J Innov Res Sci Eng Technol 2(3):687–690

McDonough W, Braungart M (2002) Cradle to cradle: remaking the way we make things New York. North Point Press, NY

Miandad R, Barakat MA, Aburiazaiza AS, Rehan M, Nizami AS (2016) Catalytic pyrolysis of plastic waste: a review. Process Saf Environ Protect 102:822–838. https://doi.org/10.1016/j.psep.2016.06.022

MoHUA (Ministry of Housing & Urban Affairs, Government of India) (2019) Plastic waste management issues, solutions & case studies, 2019. Source: http://164.100.228.143:8080/sbm/content/writereaddata/SBM%20Plastic%20Waste%20Book.pdf. Accessed 10 June 2021

MoEFCC (Ministry of Environment, Forest and Climate Change, Government of India) (2018) Beat plastic pollution: good news from India. MoEFCC, New Delhi

MoEFCC (2020) Guideline document uniform framework for extended producers responsibility, Available at: http://moef.gov.in/wp-content/uploads/2020/06/Final-Uniform-Framework-on-EPR-June2020-for-comments.pdf. Accessed 20 Apr 2021

Mohan SV, Reddy MV (2013) Optimization of critical factors to enhance polyhydroxyalkanoates (PHA) synthesis by mixed culture using Taguchi design of experimental methodology. Bioresour Technol 128:409–416. https://doi.org/10.1016/j.biortech.2012.10.037

Mohan SV, Reddy MV, Subhash GV, Sarma PN (2010) Fermentative effluents from hydrogen producing bioreactor as substrate for poly (β-OH) butyrate production with simultaneous treatment: an integrated approach. Bioresour Technol 101(23):9382–9386. https://doi.org/10.1016/j.biortech.2010.06.109

Mohanty S (2017) Recycling of plastics in Indian perspective. UNIDO Office. https://www.unido.org/sites/default/files/files/2018-11/Plenary 2 - Plastics - Mohanty.pdf. Accessed 1 May 2021

Mukherjee S, Mukhopadhyay S, Hashim MA, Sengupta B (2015) Contemporary environmental issues of landfill leachate: assessment and remedies. Crit Rev Environ Sci Technol 45(5):472–590. https://doi.org/10.1080/10643389.2013.876524

Narayanan M, Kumarasamy S, Ranganathan M, Kandasamy S, Kandasamy G, Gnanavel K, Mamtha K (2020) Production and characterization of polyhydroxyalkanoates synthesized by *E. coli* Isolated from sludge soil. Materials Today: Proceedings. https://doi.org/10.1016/j.matpr.2020.05.725

Naskar M, Chaki TK, Reddy KS (2010) Effect of waste plastic as modifier on thermal stability and degradation kinetics of bitumen/waste plastic blend. Thermochim Acta 509:128–134. https://doi.org/10.1016/j.tca.2010.06.013

NPC (2019) Tool Kit on Plastic Waste Management Rules 2016, National Productivity Council, Available at: https://www.npcindia.gov.in/NPC/Files/delhiOFC/EM/Tool-kit-on-plastic-management.pdf. Accessed 14 June 2021

OECD (2018) Improving plastics management: trends, policy responses, and the role of international co-operation and trade. OECD Environment Policy Paper no. 12. Prepared by the OECD for the G7 Environment, Energy and Oceans Ministers. Available at: http://www.indiaenvironmentportal.org.in/files/file/Improving%20Plastics%20Management.pdf. Accessed 13 May 2021

Panda AK, Singh RK, Mishra DK (2010) Thermolysis of waste plastics to liquid fuel: a suitable method for plastic waste management and manufacture of value added products—a world prospective. Rene Sustain Energy Rev 14(1):233–248

Pandey S., Manuja, S (2020) Plastic waste management: turning challenges into opportunities. https://www.teriin.org/

policy-brief/plastic-waste-management-turning-challenges-opportunities. Accessed 7 Apr 2021
Pandey U, Stormyr JA, Hassani A, Jaiswal R, Haugen HH, Moldestad BM (2020) Pyrolysis of plastic waste to environmentally friendly PRODUCTS. Energy Produc Manage in the 21st century IV: The Quest for Sustain Energy 246:61–74
Phanisankar BSS, Vasudeva RN, Manikanta JE (2020) Conversion of waste plastic to fuel products. Materials Today: Proceedings https://doi.org/10.1016/j.matpr.2020.02.880
Pillai J (2021) Managing plastic waste–what emerging economies like India can learn from developed nations. Reinf Plast. https://doi.org/10.1016/j.repl.2021.04.011
Plastindia (2018) Report on The Indian Plastics Industry. Available at: https://plastindia.org/pdf/Indian-Plastics-Industry-Report-2018-2.pdf. Accessed 30 June 2021
Poddar SK, Paranjpe PA (2015) Energy recovery through plastic waste in cement industry: a review. Int J Recent Dev Eng Technol 4(5):15–18
Ponnada S, Krishna KV (2020) Experimental investigation on modification of rheological parameters of bitumen by using waste plastic bottles. Materials Today: Proceedings 32:692–697
Positronics, India (2021) www.positronicsindia.co.in. Accessed 24 May 2021
Punith VS, Veeraraghavan A (2003) Laboratory fatigue studies on bituminous concrete mixed utilizing waster sherdded plastic modifier, Proceedings of 21st ARRB Transport Research (ARRB) and 11th Road Engineering Association and Australia (REAAA) Conference, Caims, Australia, May 19–23, 2003
PWM rules (2018). Plastic Waste Management (Amendment) Rules, 2018, The Gazette of India, 2018, Ministry of Environment, Forest and Climate Change
Raghatate AM (2012) Use of plastic in a concrete to improve its properties. Inter J Adv Eng Res Studies 1(3):109–111
Raha UK, Kumar BR, Sarkar SK (2021) Policy framework for mitigating land-based marine plastic pollution in the Gangetic Delta Region of Bay of Bengal- a review. J Clean Prod 278:123409. https://doi.org/10.1016/j.jclepro.2020.123409
Rai B, Rushad ST, Kr B, Duggal SK (2012) Study of waste plastic mix concrete with plasticizer. International Scholarly Research Notices. ISRN Civil Engineering. Article ID 469272
Rajasekaran S, Vasudevan R, Paulraj S (2013) Reuse of waste plastics coated aggregates-bitumen mix composite for road application—green method. American Journal of Engineering and Research 2(11):1–13
Rajmohan KVS, Ramya C, Raja Viswanathan M, Varjani S (2019) Plastic pollutants: effective waste management for pollution control and abatement. Curr Opin Environ Sci Health 12:72–84. https://doi.org/10.1016/j.coesh.2019.08.006
Rajput PS, Yadav RK (2016) Effect of plastic waste on properties of road aggregate. Inter J Innov Res Sci Technol 2(11):2012–2015
Reddy MV, Amulya K, Rohit MV, Sarma PN, Mohan SV (2014) Valorization of fatty acid waste for bioplastics production using Bacillus tequilensis: integration with dark-fermentative hydrogen production process. Inter J Hydrogen Energy 39(14):7616–7626. https://doi.org/10.1016/j.ijhydene.2013.09.157
Reshmy R, Philip E, Vaisakh PH, Raj S et al (2021) Development of an eco-friendly biodegradable plastic from jack fruit peel cellulose with different plasticizers and Boswellia serrata as filler. Sci Total Environ 144285. https://doi.org/10.1016/j.scitotenv.2020.144285
Roja K, Sudhakar DR, Anto S, Mathimani T (2019) Extraction and characterization of polyhydroxyalkanoates from marine green alga and cyanobacteria. Biocatal Agri Biotechnol 22:101358. https://doi.org/10.1016/j.bcab.2019.101358
Rokade S (2012) Use of waste plastic and waste rubber tyres in flexible highway pavements. International Conference on Future Environment and Energy IPCBEE, Vol.28 IACSIT Press, Singapore, 105–108
Sabina KTA, Sangita SD, Sharma BM (2009) Performance evaluation of waste plastic/polymer modified bituminous concrete mixes. J Sci Ind Res 68:975–979
Saebea D, Ruengrit P, Arpornwichanop A, Patcharavorachot Y (2020) Gasification of plastic waste for synthesis gas production. Energy Rep 6:202–207. https://doi.org/10.1016/j.egyr.2019.08.043
Saikia N, De Brito J (2012) Use of plastic waste as aggregate in cement mortar and concrete preparation: a review. Construc Build Mater 34:385–401
Sasane N, Gaikwad H, Patil JR, Khandekar SD (2015) Application of waste plastic as an effective construction material in flexible pavement. Inter Res J Eng Technol 2(3):1943–1948
Sasidharan M, Eskandari Torbaghan M, Burrow MPN (2019) Using waste plastics in road construction. K4D Helpdesk Report. Brighton, UK: Institute of Development Studies. Available at: https://assets.publishing.service.gov.uk/media/5d41b34040f0b60a86a5e5dc/595_Use_of_Waste_Plastics_in_Road_Construction.pdf. Accessed 27 May 2021
Satapathy S (2017) An analysis of barriers for plastic recycling in the Indian plastic industry. Benchmarking: An Inter J https://doi.org/10.1108/BIJ-11-2014-0103
Scheirs J (1998) Polymer Recycling. Wiley, New York
Senhadji Y, Escadeillas G, Benosman AS, Mouli M, Khelafi H, Ould Kaci S (2015) Effect of incorporating PVC waste as aggregate on the physical, mechanical, and chloride ion penetration behavior of concrete. J Adhes Sci Technol 29:625–640. https://doi.org/10.1080/01694243.2014.1000773
Siddique R, Khatib J, Kaur I (2008) Use of recycled plastic in concrete: a review. Waste Manage 28(10):1835–1852. https://doi.org/10.1016/j.wasman.2007.09.011
Singh K, Mathur A (2019) Plastic pollution in India: an evaluation of public awareness and consumption behaviour. OIDA Inter J Sustain Develop 12(7): 25–40. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3548729. Accessed 6 Apr 2021
Singh RK, Ruj B, Sadhukhan AK, Gupta P, Tigga VP (2020a) Waste plastic to pyrolytic oil and its utilization in CI engine: performance analysis and combustion characteristics. Fuel 262:116539. https://doi.org/10.1016/j.fuel.2019.116539
Singh RK, Ruj B (2016) Plasticwaste management and disposal techniques-Indian scenario. International J Plastics Technol 19(2):211–226. https://doi.org/10.1007/s12588-015-9120-5
Singh TS, Verma TN, Singh HN (2020b) A lab scale waste to energy conversion study for pyrolysis of plastic with and without catalyst: Engine emissions testing study. Fuel 277:118176. https://doi.org/10.1016/j.fuel.2020.118176
Solis M, Silveira S (2020) Technologies for chemical recycling of household plastics–a technical review and TRL assessment. Waste Manage 105:128–138. https://doi.org/10.1016/j.wasman.2020.01.038
Song JH, Murphy RJ, Narayan R, Davies GBH (2009) Biodegradable and compostable alternatives to conventional plastics. Phil Trans R Soc B 364:2127–2139. https://doi.org/10.1098/rstb.2008.0289
Srivastav R (2019) India's plastic waste situation wasn't created today. Available at: https://www.downtoearth.org.in/blog/waste/india-s-plastic-waste-situation-wasn-t-created-today-67061. Accessed 6 Apr 2021
Subramani T, Pugal VK (2015) Experimental study on plastic waste as a coarse aggregate for structural concrete. Inter J App Innov Eng Manage 4(5):144–152
Swachhindia (2017) https://swachhindia.ndtv.com/recycling-plastic-india-conversion-plastic-to-fuel-the-unrealized-p-otential-9436/. Accessed 18 June 2021

Swati TI, Vijay VK, Ghosh P (2018) Scenario of landfilling in India: problems, challenges, and recommendations. Handbook of environmental materials management. Springer, Cham 1–16

Taylor D (1999) Talking trash: the economic and environmental issues of landfills. Environ Health Pers 107(8):A404–A409. https://doi.org/10.1289/ehp.99107a404

The Gazette of India (2018) http://egazette.nic.in/WriteReadData/2018/184349.pdf. Accessed 9 June 2021

The Hindu (2018) How India converts plastic to crude oil. https://www.thehindu.com/sci-tech/energy-and-environment/plastic-surgery/article25604517.ece. Accessed 3 June 2021

The Hindu (2019) Wasted effort: half of India's waste-to-energy plants defunct. https://www.thehindu.com/news/national/wasted-effort-half-of-indias-waste-to-energy-plants-defunct/article26273068.ece. Accessed 29 May 2021

Time of India (2019) India has a 26,000-tonne plastic waste problem. http://timesofindia.indiatimes.com/articleshow/67646745.cms?utm_source=contentofinterest&utm_medium=text&utm_campaign=cppst. Accessed 3 Apr 2021

Ukarde TM, Pawar HS (2021) A Cu doped TiO2 catalyst mediated Catalytic Thermo Liquefaction (CTL) of polyolefinic plastic waste into hydrocarbon oil. Fuel 285:119155. https://doi.org/10.1016/j.fuel.2020.119155

Unde SR, Potnis SC (2015) Effective utilization of plastic waste in flexible pavemmet and analysis by experiments. Inter J Eng Sci Res Technol 4:882–891

Van den Oever M, Molenveld K, van der Zee M, Bos H (2017) Bio-based and biodegradable plastics: facts and figures: focus on food packaging in the Netherlands (No. 1722). Wageningen Food & Biobased Research. Available at: https://edepot.wur.nl/408350. Accessed 5 July 2021

Vanitha S, Natrajan V, Praba M (2015) Utilization of waste plastic as a partial replacement of coarse aggregates in concrete Blocks. Ind J Sci Technol 8:0974–5645

Vasudevan R, Sekar ARC, Sundarakannan B, Velkennedy R (2012) A technique to dispose waste plastics in an ecofriendly way–application in construction of flexible pavements. Constr Build Mater 28(1):311–320. https://doi.org/10.1016/j.conbuildmat.2011.08.031

Vasudevan RNSK, Velkennedy R, Sekar ARC, Sundarakannan B (2010) Utilization of waste polymers for flexible pavement and easy disposal of waste polymers. Inter J Pavement Res Technol 3(1):34–42

Verma R, Vinoda KS, Papireddy M, Gowda ANS (2016) Toxic pollutants from plastic waste-a review. Procedia Environ Sci 35:701–708. https://doi.org/10.1016/j.proenv.2016.07.069

Walendziewski J, Steininger M (2001) Thermal and catalytic conversion of waste polyolefines. Catal Today 65(2–4):323–330. https://doi.org/10.1016/S0920-5861(00)00568-X

Wang JL, Wang LL (2011) Catalytic pyrolysis of municipal plastic waste to fuel with nickel-loaded silica-alumina catalysts. Energy Sources Part A 33(21):1940–1948. https://doi.org/10.1080/15567030903436814

Wong SL, Ngadi N, Abdullah TAT, Inuw IM (2015) Current state and future prospects of plastic waste as source of fuel: a review. Ren Sustain Energy Rev 50:1167–1180. https://doi.org/10.1016/j.rser.2015.04.063

www.cemnet.com (2019) India's fuel shift to plastic waste is coming. Available at: https://www.cemnet.com/News/story/167506/india-s-fuel-shift-to-plastic-waste-is-coming.html. Accessed 6 May 2021

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Scientia Iranica B (2021) 28(6), 3181–3200

Sharif University of Technology
Scientia Iranica
Transactions B: Mechanical Engineering
http://scientiairanica.sharif.edu

Prediction of an optimum engine response based on different input parameters on common rail direct injection diesel engine: A response surface methodology approach

M. Kumar[a], N.A. Ansari[a], A. Sharma[b], V.K. Singh[a], R. Gautam[a], and Y. Singh[c,*]

a. *Department of Mechanical Engineering, Delhi Technological University, Delhi, India.*
b. *Department of Mechanical Engineering, G L Bajaj Institute of Technology and Management, Greater Noida, UP, India.*
c. *Department of Mechanical Engineering, Graphic Era Deemed to be University, Dehradun, Uttarakhand, India.*

Received 20 October 2020; received in revised form 12 January 2021; accepted 19 April 2021

KEYWORDS
Linseed;
Biodiesel;
Blend;
Performance;
Emission;
Common rail direct injection;
Response surface methodology.

Abstract. With increasing growth in industrialization all over the world, the need for reducing harmful pollutants gains significance, especially in the transportation sector. In this respect, the present study evaluates the effects of different engine operating parameters on the performance and emissions of a CRDI (Common Rail Direct Injection) diesel engine. The main objective of this study was to optimize the emissions as well as efficiency parameters to achieve the optimal configuration parameters for the engine using the desirability approach of Response Surface Methodology (RSM) technique, a suitable optimization approach to saving a lot of repetitive testing. It was observed that after RSM modeling, the optimized engine settings of input factors were diesel/linseed blend concentration 8.10%, FIP (Fuel Injection Pressure) 600 bar, EGR (Exhaust Gas Recirculation) level 4.667%, and load on engine 9.33 kg. On these constant hold values, the optimized output torque, BTE (Brake Thermal Efficiency), BMEP (Brake Mean Effective Pressure), mechanical efficiency, HC (hydrocarbon), and CO_2 carbon dioxide were calculated as 20.04 Nm, 26.035%, 3.474 bar, 52.503%, 28.14 ppmv, and 7.319 %vol respectively. The aforementioned predicted values were experimentally validated, and the errors in the predicted values were in a limited range.

© 2021 Sharif University of Technology. All rights reserved.

1. Introduction

With a considerable shortage of the world petroleum supplies and rising consumption of petroleum products, especially transportation fuels, the need for liquid fuel replacement has gained much significance for all countries. Singh et al. [1] found that in the automobile industry, non-renewable diesel fuel is one of the primary fossil fuels whose burning had a detrimental effect on the atmosphere. Due to the gradual, yet rapid, depletion of fossil fuels, the search for renewable fuel sources is of importance. Gautam and Kumar [2] observed that emissions of carbon monoxide (CO), Unburnt Hydro-Carbon (UHC), carbon dioxide (CO_2), smoke particles, and oxides of nitrogen (NOx) from the Compression Ignition (CI) engines had a severe impact on human beings as well as the environment.

*. *Corresponding author.*
E-mail addresses: yashvirkankara21@gmail.com (Y. Singh)

doi: 10.24200/sci.2021.56745.4885

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.

ReactiveFnJ: A Choreographed Model for Fork-Join Workflow in Serverless Computing

Urmil Bharti (urmil.bharti@rajguru.du.ac.in)
Shaheed Rajguru College of Applied Sciences for Women https://orcid.org/0000-0003-2655-3001
Anita Goel
University of Delhi
S. C. Gupta
Indian Institute of Technology

Research Article

Keywords: Serverless computing, FaaS, Event-driven function composition, Choreography, Parallel computing, Distributed computing, Fork and Join, Orchestration

Posted Date: January 10th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-942960/v1

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

Abstract

Function-as-a-Service (FaaS) is an event-based reactive programming model where functions run in ephemeral stateless containers for short duration. For building complex serverless applications, function composition is crucial to coordinate and synchronize the workflow of an application. Some serverless orchestration systems exist, but they are in their primitive state and do not provide inherent support for non-trivial workflows like, Fork-Join. To address this gap, we propose a fully serverless and scalable design model *ReactiveFnJ* for Fork-Join workflow. The intent of this work is to illustrate a design which is completely choreographed, reactive, asynchronous, and represents a dynamic composition model for serverless applications. Our design uses two innovative patterns, namely, Relay Composition and Master-Worker Composition to solve execution time-out challenges. As a Proof-of-Concept (PoC), the prototypical implementation of Split-Sort-Merge use case, based on Fork-Join workflow is discussed and evaluated. The *ReactiveFnJ* handles embarrassingly parallel computations, and its design does not depend on any external orchestration services, messaging services, and queue services. *ReactiveFnJ* facilitates in designing fully automated pipelines for distributed data processing systems, satisfying the Serverless Trilemma in true essence. A file of any size can be processed using our effective and extensible design without facing execution time-out challenges. The proposed model is generic and can be applied to a wide range of serverless applications that are based on the Fork-Join workflow pattern. It fosters the choreographed serverless composition for complex workflows. The proposed design model is useful for software engineers and developers in industry and commercial organizations, total solution vendors and academic researchers.

1. Introduction

Serverless computing is an emerging paradigm and is gaining popularity in the cloud owing to its simplicity, billing model, and inherent elasticity. This cloud computing execution model greatly simplifies the usage of cloud resources and suits well to highly scalable, event-driven applications in the cloud. Serverless architecture is especially effective at supporting modern applications with unpredictable scale and user demand [2].

The FaaS programming model of serverless allows programmers to develop cloud applications as individual functions that can run and scale independently. This model is event-driven since functions are activated in reaction to specific cloud events like, a state change in an object store, receipt of a message, a file upload, or insertion of a record in database. Though FaaS looks like a promising option for deploying cloud applications, it has few limitations also. Most notably, FaaS functions are stateless, short-lived, and cannot communicate directly with each other [15] [24]. Thus, executing complex, burst-parallel jobs pose a significant challenge for serverless execution frameworks [13]. The composition of workflows in such jobs require extensive fine-grained communication and synchronization between independent functions that is challenging to implement in a serverless framework [8]. These FaaS challenges force developers to resort to alternate ways to establish function communication like publish-

subscribe, passing data over some slow and expensive storage medium, or serverless orchestration services, but all these alternatives yield too high latency and cost [27].

At a high level, there are two approaches for function composition in a serverless application: orchestration and choreography [28], as shown in Fig. 1. In orchestration, a controller module orchestrates and controls the interaction between serverless functions. The controller module governs the flow according to the needs of the business logic. The choreography model is an event-driven paradigm in which every function works autonomously as a loosely coupled service. The functions work in a pipeline based on the triggered events. Each function performs its task, and its completion triggers the next function/s down the pipeline. In an event-driven architecture where each component plays a more architecturally aware role, the choreography model is used in the design of workflows instead of an orchestration model [4]. [23].

The Fork-Join model is a programming method that exploits parallelism in applications based on inherent *divide and conquer* algorithms [22]. This execution model has already been successfully used for building parallel systems where an incoming task splits into subtasks that are processed by a set of parallel servers. The implementation of the Fork-Join model becomes more practical in serverless computing as these platforms are inherently scalable and do not need resource provisioning in advance. Since horizontal scaling in serverless is entirely automatic, elastic, and managed by its provider, dynamic parallel processing in Fork-Join can best exploit these characteristics.

Currently, Fork-Join workflow cannot be composed using any of the available serverless orchestration services [4] like Amazon Step Functions (ASF) (December 2016), Azure Durable Functions (June 2017), and IBM Composer (October 2017). These services lack the ability to dynamically launch functions in parallel. Though ASF, the most mature and performant project [13], supports the *Parallel* state type to execute tasks in parallel. However, the application developers must list all the tasks to execute in parallel in an array in the state machine, thereby restricting the flexibility of the concurrency level [34].

In our research, we present *ReactiveFnJ*, a fully choreographed, vendor-neutral, and platform-independent serverless design model for Fork-Join workflow. The *ReactiveFnJ* uses an innovative algorithmic design that is purely event-driven, reactive, ST-Safe, and is free from hard execution time limits imposed by a serverless provider. The *ReactiveFnJ* handles embarrassingly parallel computations, but it does not depend on any external orchestration services, messaging services, and queue services. This model is generic, and several other use cases can be implemented by substituting specialized FaaS functions in our PoC implementation. *ReactiveFnJ* model is the first work that exhibits all the characteristics mentioned above to the very best of our knowledge. Our design and its implementation will also lead us to find answers of the following research questions:

RQ1: Does any serverless design exist for long running computation?

RQ2: Does the Serverless challenge of executing a complex, burst-parallel workflow can be addressed programmatically without using any external orchestration service?

RQ3: Is it possible to design a serverless system using pure event-driven architecture for complex workflows?

This manuscript is organized into several sections. Section 2 presents the technical background, motivation behind the proposed *ReactiveFnJ* serverless design model and summarizes our contributions. The detailed design and algorithms related to our serverless Fork-Process-Join pipeline are described in Sect. 3. Section 4 gives a detailed description of the implementation of our design concepts in the AWS Lambda. Section 5 shows the evaluation insights of the proposed models, followed by a discussion and lessons in Sect. 6. The related works about the serverless applications implemented for parallel processing are described in Sect. 7. The concluding remarks and scope for future works are delineated in Sect. 8.

2. Background, Motivation, And Contribution

This section discusses current serverless challenges, the motivation behind this work, and the significant contributions of this research.

A. Serverless and its Challenges

The serverless model was originally designed to execute event-driven, stateless functions in response to user actions or changes in the storage tier, e.g., uploading a photo to Amazon S3, inserting/updating a record in DynamoDB, etc. [21]. Some recent works have shown that the large-scale parallelism and auto-scaling features provided by serverless platforms make them well-suited for burst-parallel fine-grained tasks and parallel computation workflows [12]. In essence, the FaaS model is apt for embarrassingly parallel computing use cases such as linear algebra [31], optimization algorithms [10], data analytics [30], and real-time machine learning classifications [9].

Building a complex serverless application with numerous short-lived, concurrent functions requires new design guidelines. Beyond simple examples, serverless applications need to be designed as a composition of functions. In most of the cases, a serverless workflow composition needs an orchestration service that provides a coordination mechanism between FaaS functions [16]. These coordination services automatically trigger the execution of each function in the workflow and synchronize their behaviors and states. FaaS orchestration services such as AWS Step Functions or IBM Composer offer limited capabilities to coordinate serverless functions [5]. For instance, even in the AWS Step function, there is no provision for multiple functions to synchronize in parallel when the number of parallel instances is dynamic. In place of orchestration services, developers use some indirect ways to synchronize the dynamic parallel execution of functions via notification services, queue services, and in-memory data store/cache services [12] but they have their own limitations and downsides.

Few researchers have come up with solutions for handling embarrassingly parallel computations in serverless. PyWren [19] uses its own ad-hoc external orchestration service, and ExCamera [11] relies on an external server to synchronize the parallel executions. These solutions add significant latency and cost.

Nevertheless, all the systems implemented so far for parallel computing do not comply with the four requirements claimed by Amazon for a serverless application: (i) No server management, (ii) Flexible scaling, (iii) Pay for value, and (iv) Automated high availability [4].

Barcelona-Pons. *et al*. argued that serverless functions follow a trigger-based model, so a FaaS composition should also be trigger-based [4]. This means that the termination of one or many functions should trigger the next stage (function) using asynchronous events in a workflow. From this perspective, any serverless composition achieving dynamic parallelism should also be trigger-based. Therefore, we emphasize that it is of utmost desire to build a serverless application as a complete reactive system of FaaS function compositions.

Any function composition is referred as Serverless Trilemma (ST) safe if it satisfies three main principles of Serverless Trilemma stated by Baldini et al. [3]: (1) Substitution - Each composition should behave like a function and could be substituted in any other pipeline, (2) Black-box - Each component of the workflow should be a black-box and abstracts from rest of the system i.e. implementation details of functions remain hidden from others, and (3) No Double billing - FaaS is a pay-per-use model, i.e., fine-grained resource measurement based on usage and there should not be double-billing of cloud function.

B. Fork-Join Model in Serverless

The generic Fork-Join model of parallel processing splits a compute-intensive task into smaller sub-tasks to process them parallelly using the available CPU cores [29]. Therefore, the Fork-Join design model renders execution speed-up by running forked tasks in parallel and combining their results. This model can be best utilized in serverless implementation for use cases like, sorting [20], searching [17], matrix multiplication [6], string matching, MapReduce patterns like counting and summing, Collating, and Filtering [14]. In all the mentioned cases where the volume of data fluctuates, and resource requirements cannot be anticipated, the pay-per-use model of serverless is best used.

The above use cases can be best implemented using the generic Fork-Join workflow of parallel processing in serverless but designing this scalable workflow efficiently is a challenge in serverless frameworks. Currently, function composition for Fork/Join is not directly supported by the existing serverless orchestration services [4]. The available orchestration services are not designed for managing parallel Fork-Join workflows in a scalable and efficient way.

C. Motivation

Execution time limit is a major constraint of FaaS that hinders its implementation to those applications where running time might go beyond the set time limits. We experimented in the AWS Lambda, a serverless computing service provided by Amazon Web Services and found that a *Single FaaS* function can only sort a file of size 560 kb in a default environment setting. If this function gets an input file of size more than 560 kb, it ceases to complete due to execution time out.

It is just an example, but there are many compute-intensive scientific and business applications where constituent functions may time-out before their completion and hence applications are unable to harness the power of serverless computing. There is a class of inherently parallel applications in any domain, where the initial task can be split into a large number of independent sub-tasks (Fork), and then each sub-task can be implemented as autonomous functions in serverless. This design brings up two challenges in serverless: (i) to provide some coordination mechanism to run in parallel a multitude of functions derived from a single task and (ii) to devise a synchronization mechanism to join the results of all split tasks and to prepare the aggregated output (Join).

One can attempt to use an existing orchestration service, but it may not be a viable solution as they do not support the execution of tasks in parallel when number of tasks are dynamically determined at runtime. As an example, Amazon Step Functions support the *Parallel* state type to execute tasks in parallel, but application developers have to provide all the tasks to execute in parallel in an array construct of the state machine [8]. This restricts the flexibility of concurrency level and hence its usage in scenarios, where launching of functions in parallel is dynamic in nature.

The main downsides of currently available orchestration services are (i) Billing is based on the number of transitions happening during workflow execution (ii) Latency issues when working complex workflows (iii) Non adherence to serverless trilemma (violation of substitution and double billing principles) (iv) extra efforts are required to build a workflow in the orchestration services.

To handle these problems, we aim to design and implement a function composition mechanism for Fork-Join workflow which can be used into a broad spectrum of applications.

D. Contributions

The main contribution of this work is an algorithm-based design for serverless Fork-Join workflow. We have proposed a trigger-based serverless design model namely, *ReactiveFnJ*, for Fork-Join workflow. This design model can be utilized in compute intensive and burst-parallel applications. Our proposed model employs innovative design patterns that can process a file of any size without being time-out. In *ReactiveFnJ*, multiple component synchronization and coordination is crafted by *Relay Composition* and *Master-Worker Composition* design patterns thereby making it a choreographed and pure event-driven system. *ReactiveFnJ* is an asynchronous dynamic serverless composition design model that fully exploits the scalability, availability, and built-in fault tolerance of serverless infrastructure.

To prove the feasibility and viability of *ReactiveFnJ* design, we build a prototypical implementation to sort a large input file as a proof of concept. We call this sorting requirement as Split-Sort-Merge (SSM) use case throughout this research article. In SSM, the main aim is to sort a data file of any size (theoretically) where individual records are of variable length. There are several approaches to sort a large data file but the main challenge here is to develop an approach for a serverless architecture, where serverless functions are stateless and have a constrained execution environment. There exist multiple traditional serverful deployment frameworks for SSM, like, Map-Reduce and Apache Spark however, these

frameworks suffer from cluster management, load balancing and task fairness issues [1]. Thus, developing/migrating these applications to serverless platforms illustrates unique opportunities.

After the successful implementation of SSM, it can be claimed that applications designed using *ReactiveFnJ* will not have dependency on any external orchestration service, will not time out and will be free from vendor lock-in problems.

To summarize, the present work makes the following notable contributions:

- A pure event-driven choreographed system design for Fork-Join workloads suitable for serverless deployments.
- State of the art serverless function composition model satisfying all competing constraints of ST Trilemma (ST-Safe).
- Employs asynchronous push-based design exploiting recursive calling of stateless functions to handle long running compute intensive workloads.
- Self-driven function composition mechanism without relying on any external orchestration service to handle the complexities of state management.
- An Algorithm-level design solution that overcomes the execution time limitation imposed by serverless providers.

Our approach shall be a tipping point where a single machine/container is not big enough to perform a big computation and a serverless function fails for the same reason. Using our proposed novel algorithmic design approach, it shall be feasible to execute burst-parallel, compute intensive applications having large data-at-scale by leveraging the scalability benefit of FaaS.

3. Reactivefnj: Design

This section presents the design of *ReactiveFnJ* in detail. The main design components, their interaction, design challenges, and employed algorithms are discussed in depth.

ReactiveFnJ design has three main components i.e., 1) *Fork*, 2) *Process* and 3) *Join* as shown in Fig. 2. Using the well-established divide and conquer principle, The Fork component divides the main task into smaller subtasks for parallel processing. Once the main task is subdivided, each subtask is processed independently by *Process* component and intermediate results are produced. Responsibility of the Join module is to combine these intermediate results to produce the result.

3.1 The Fork component

The first component of *ReactiveFnJ* i.e., Fork is designed to read a data file of any size and split it into smaller files where size of each split file is always less than or equal to maximum file-size. Maximum file-size is a configurable parameter and can be set according to the serverless execution environment settings.

Challenge in design

In a conventional way, data records of the input file can be read to create small size file splits. Maximum number of records in a split file can be passed as an input parameter. This design for reading and splitting a file will eventually fail in case the input data file is too big. In other words, reading and carving smaller files will continue till the function itself does not surpass the memory usage and execution time-out limits imposed by service providers. This conventional design challenge for serverless is alleviated by a *Relay Composition* pattern *for Fork* as described below.

Relay Composition **for Fork** - In this design, a big file is read, and split files are created by initiating a recursive *Relay Composition* pattern. This pattern can handle an input file of any size, theoretically, without being execution time-out. To make this composition more efficient, data is read and written in byte chunks of almost fixed number of bytes. Let *MaxSplitSize* (*MSS*) indicate the maximum byte chunk that can be read/written by a single function without execution time-out. MSS is an estimated value and can be calculated empirically to know the maximum number of bytes that a *Single FaaS* function can read/write.

The number of split files N can be calculated as follows:

$$N = \left\lceil \left(\frac{Input\ FileSize}{MSS} \right) \right\rceil$$

$$\text{For } 1, 2 \ldots, N \text{ there are } F_1, F_2, \ldots, F_N \text{ split files where } |F_i| \leq MSS, \forall F_i ; i \in Z^+$$

The *Relay Composition* starts reading the first byte chunk and writes it to a new file and triggers the next instance asynchronously. New instance starts reading from the byte position in the file where the previous instance had stopped. The last record in a byte chunk may be incomplete due to variable length records and fixed *MSS* value. The composition design takes care of this case by discarding the partial read record and adjusting *StartByteLocation* parameter value as given in Algorithm 1. *StartByteLocation* works as a relay baton and is used to pass the next read position of the input file to the subsequent instance. Hence the *StartByteLocation* is being relayed in every successive recursive instance till the end of the file is reached. This recursive style in *Relay Composition* where *StartByteLocation* being passed in successive calls as shown in Fig. 3. Thus, a file of any size, having variable length records can be forked in serverless infrastructures without being time-out.

Algorithm 1: Read and Split file of any size

Input: InputFilePath, Byte Chunk Size, OutputFilePath

Output: Split Files

```
1. # Reads and splits input file into multiple files of specified size
2. function splitFile(FilePath inputFile, Number numBytesToRead, FilePath outputFile)
3. open inputFile in read mode,  open outputFile in write mode
4. inputFileSize = get number of bytes in inputFile
5. bytesTillNewLine = seek newLine/EOF after numBytesToRead bytes from startPosition
6. while (startPosition + numBytesToRead + bytesTillNewLine < inputFileSize )
7. bytesData = readChunk from startPosition till (startPosition +
8. numBytesToRead + bytesTillNewLine) position
9. write bytesData to outputFile, close outputFile,open new outputFile
10. startPosition = (startPosition + numBytesToRead + bytesTillNewLine)
11. if (startPosition + numBytesToRead + bytesTillNewLine < inputFileSize ) then
12. bytesTillNewLine = seek newLine/EOF after numBytesToRead bytes from startPosition
13. else
14. #remaining file size is less than the bytes to read
15. bytesData = readChunk from startPosition till file end
16. write bytesData to outputFile, close outputFile
17. end if
18. end while
19. end
```

Each recursive instance creates a new split file that can be processed by the Process component in the design pipeline.

3.2 The Process component

The second component of *ReactiveFnJ* is designed to perform computation on all the splits carved in the previous step. All the split files are processed in parallel by independent functions. In essence, the design of *Process* component harnesses the power of scalability offered by a serverless infrastructure. Hence, parallel processing of sub-parts helps in reducing the overall execution time in serverless. This component is the realization of a serverless idea of dynamically creating a compute cluster on demand without any overhead of cluster management. This component can implement use cases where

computation is data independent and therefore, can be executed in parallel. Some exemplary use cases include eCommerce, clickstream analytics, contact centre, legacy app modernization, and DevOps functions. In Algorithm 2, in-memory sort on a split file has been demonstrated but it could be any computation required to run in parallel.

Algorithm 2: In-memory sort of an Individual file
Input: InputFilePath, OutputFilePath
Output: Sorted OutputFile
1. #Sorts the input file in ascending order
2. function sortFile(FilePath inputFile, FilePath outputFile)
3. Open inputFile in read mode
4. For each Row in inputFile do
5. Add row to dataList
6. For End
7. Sort data in dataList in ascending order
8. Open outputFile in write mode
9. For each datarow in dataList do
10. Write datarow to outputFile
11. For End
12. End

3.3 The Join component

The results of the *Process* component are combined in the Join component to converge the result. The main job of this component is to join the processed split files in parallel. As per the design, in the first iteration, all available split files will be paired first and then joined to create a single file. The join component keeps on iterating this till a single file is left as shown in Fig. 2. Design of *Join* component is extremely efficient as number of iterations are growing in a binary logarithmic fashion as shown in Fig. 4 below. The number of iterations *i* required to join N number of files can be determined using our formula $2^{i-1} < N \leq 2^i$ where $i \in Z^+$.

The design of *Join* component was the biggest challenge we faced. Being a serverless component, the *Join* design should adhere to the following serverless principles i) complete decentralized scheduling, ii) reactive inter-module communication, iii) pure asynchronous push-based communication approach. To

claim a component to be truly serverless, it should not use any external orchestration service to synchronize its execution workflows along with the principles mentioned above.

Challenge in design:

The conventional way of joining N processed files in parallel is not a viable design in a serverless environment. In the traditional design, all processed files are joined in parallel in pairs. Each pair of files initiates the event-driven merging process, and a sorted file is created. Newly built joined files are again ready for join and this process will continue till a single file is left as shown in Fig. 5. This typical design works well when the size of a joined file can be handled under the limits of a serverless environment. But it shall eventually fail when file size starts growing as the joining process exceeds the serverless execution time limit. This design challenge is resolved by a reactive *Master-Worker Composition* (MWC) pattern as described below.

Master-Worker Composition **for Join -** MWC is an innovative design solution to handle the exponential increase in file size, the main cause of execution time-out, during the joining process. The main highlight of this design is that it can join any number of files without any constraint on file size.

In MWC, a parallel recursive join process is formulated which is devoid of time-out constraint. In this composition, we have two important modules, first is *MasterReactiveMerge* (MRM) and other one is *WorkerRecursiveMerge* (WRM). The MRM is responsible for initiating a join between a pair of split files. As MRM is reactive so whenever a pair of files is available for join, it is invoked automatically. Event-driven characteristic of serverless architecture helps MRM to run in parallel for each pair of files. Every instance of the MRM module calls a WRM module, responsible for joining two files. First invocation of this module keeps on joining the two files till the size of the joined file reaches a limit called *JoinSafeLimit* which is passed as a parameter to this module. Files joined by WRM could be of any size so its multiple sequential instances may be required to join files of big size. As the file size increases, the number of instances increases proportionally. One instance of WRM joins two files upto *JoinSafeLimit* and before getting time-out calls next WRM instance with appropriate parameters like read positions for both the files. Two files are read and joined to build a single joined file by one or more WRM instances, where the resultant file is appended by each instance. The key design features of MWC are as follows:

1. The Joining process is completely automatic where the number of files and their sizes are not known in advance.
2. There is no central controller responsible for joining files. The design of this component accomplishes the complete decentralized scheduling goal.
3. No additional messaging-service or shared-memory is used for inter-module communication. The design is based on event-driven architecture and makes use of trigger-based communication.
4. The design uses both sequential and parallel composition of serverless functions and synchronizes the burst parallelism.
5. The design ensures that executions are never double billed as design is based on pure asynchronous push-based communication approach. We understand that this design is an example of state-of-the-art

serverless function composition.
6. No external rendezvous server, ad-hoc orchestrator services and current serverless orchestration systems are used in design for task scheduling and synchronization. The design is completely based on the reactive programming model which is highly recommended for serverless.

4. Implementation

Amazon AWS Lambda is a mature FaaS platform, so we opted it for our experimental implementation. The *ReactiveFnJ* model is developed using FaaS and BaaS services provided by AWS.

To validate the design of *ReactiveFnJ*, we implement Split-Sort-Merge (SSM) case study. In this study, the main goal is to sort a data file of any size having variable length records. SSM is a perfect case study based on the generic Fork-Process-Join model of parallel processing. In SSM, Fork component split input file into smaller files, Process component does in-memory sorting of individual split files, and Join component merges the sorted files. We have typically chosen this use case to sort a very big text file that conventionally cannot be sorted using the resources limitations of a single serverless function container. There are many solutions available for handling these scenarios like Cloud IaaS, Hadoop Map Reduce, on-premises cluster etc. But we understand, serverless Function as a Service (FaaS) is the most attractive and methodical option because of its simplicity, billing flexibility and inherent elasticity. Study and implementation of this prototype aims to leverage existing event-based technology of serverless architectures to enable triggered compositions in complex workflows.

4.1 General overview of SSM

In the implementation of SSM, following assorted AWS services are used - i) AWS Simple Storage Service (Amazon S3) for storage of input file, split files, intermediate sorted files and final resultant sorted file, ii) AWS Lambda for fast, containerized, stateless execution of following sub-tasks of workflow: a) splitting b) sorting and c) merging, and iii) Amazon S3 Event Notifications to send event messages to coordinate stateless Lambda functions, iv) AWS Identity and Access Management (IAM) to manage and secure access to various AWS resources [35], v) AWS CloudWatch to monitor and observe data in the form of logs, metrics, and events for complete Fork- Join pipeline.

We implemented the AWS Lambda functions using Python 3.8, which readily offers great library support to manage critical operations like creation, deletion of S3 bucket folders at runtime, setting/retrieving input/output file path, read/write CSV format files etc. A brief description of important python modules used is shown in Table 1.

Table 1: Important Python modules used in implementation of SSM

Python Module Name	Brief Description
csv	CSV (Comma Separated Values) is a simple file format to store tabular data (numbers and text) in plain text. This module implements classes to read and write tabular data in CSV format. Programmers can also describe the CSV formats understood by other applications or define their own special-purpose CSV formats.
os	Provides functions for interacting with the operating system i.e., functions for creating and removing a directory (folder), fetching its contents, changing, and identifying the current directory etc. OS comes under Python's standard utility modules. This module provides a portable way of using operating system-dependent functionality.
botocore	Botocore is a low-level interface to a growing number of Amazon Web Services. The botocore package is the foundation for the AWS CLI as well as boto3.
boto3	It provides Amazon Web Services (AWS) SDK for Python. *boto3* is built atop of library *botocore*. It enables Python developers to create, configure, and manage AWS services, such as S3. The SDK provides an object-oriented API as well as low-level access to AWS services.
s3fs	Pythonic file interface to S3. It builds on top of *botocore*. The top-level class *S3FileSystem* holds connection information and allows typical file-system operations.
json	The *JSON* (JavaScript Object Notation) is a standardized format that allows exchanging of data across the internet. It is a lightweight data interchange format used to work with JSON data.

4.2 Lambda Function Outline

A list of AWS Lambda functions written by us for the SSM implementation are shown in Table 2.

Table 2: Lambda functions used in the Implementation of SSM

S.No	AWS Lambda	Description
1	BL_ReadAndSplit (λBR&S)	Recursive function that reads and split a data file of any size using *Relay Composition* pattern. It uses byte level read/write operations.
2	Sort (λS)	*Sort* function implements in-memory sorting of a split file. It runs in parallel to sort the data contained in the split files.
3	MasterReactiveMerge (λMRM)	Master function responsible for initiating first instance of worker i.e., WorkerRecursiveMerge. It runs in parallel to initiate pair-wise merging of all the sorted split files.
4	WorkerRecursiveMerge (λWRM)	Worker function performs recursive merging of a pair of sorted files of any size. This function also runs in parallel.

The interaction among all Lambda functions of the SSM pipeline is based on AWS S3 triggers and is shown in Table 3 below:

Table 3: AWS S3 Triggers configuration for each Lambda function

TriggerName	Events	Filter	Type	Lambda Function
SplitInput	PUT	*input/*, .csv	Lambda	BL_ReadAndSplit
ProcessInput	PUT	*to_process/*, .csv	Lambda	Sort
JoinInput	PUT	*to_join/*, .csv	Lambda	MasterReactiveMerge
MergeInput	PUT	*to_merge/*, .csv	Lambda	WorkerRecursiveMerge

4.3 Specifics of Lambda functions

In this section, Lambda functions of the SSM pipeline have been comprehensively discussed.

Our implementation uses five AWS S3 folders namely 1) *input* – for storing input file, 2) *to_process* – for storing file splits, 3) *to_join* - for storing sorted split files, intermediate merged files and final merged file, 4) *to_merge* – for temporarily storing pairs of files undergoing merge, and 5) *archive* – to archive input file after successful processing. The complete architecture design of SSM is shown in Figure 6 below:

4.3.1 BL_ReadAndSplit: A large file can be read recursively to overcome the restriction of execution time-out for an AWS Lambda function. The recursive Lambda function, λBR&S, reads a byte chunk and writes it to a new file. It is invoked when an input file is uploaded to the folder "/input" of S3 bucket. Data byte chunks are read from the input file and split files are created in a new folder "to_process" (created at runtime). A configurable parameter *ByteChunkSize* is used by this function that determines the size of split files. Just before getting time-out, this function asynchronously invokes itself by passing the new read byte position in the input file. First instance of λBR&S creates the first split file and invokes the next instance to create the next split file and so on. This process continues till the end of the file. This recursive implementation works perfectly for any size of input file. Invoking a function asynchronously using *lambda.invoke* and setting *Invocation-type* flag as "Event", place invoke requests in Lambda service queue for processing the requests as they arrive. Keeping *lambda.invoke* as the last statement in this function's code eliminates double billing as it would not force the Lambda to wait for each invocation to finish. Thus, recursive invocations of this function where each invocation communicates and coordinates with each other, attains "No Double Billing" condition for the serverless function composition mechanism of ST Trilemma [26].

After the successful completion of this function, the input file is deleted from its folder and is moved to a "\input_archive" folder in S3 bucket. Naming convention of a split file carries two important attributes of information i.e, i) Total number of Splits and ii) File Split Number as shown in Figure 7. Split file name and achieve folder helps in detecting and debugging failures, if any, via AWS CloudWatch log.

4.3.2 Sort: Sort is the AWS Lambda function implemented for sorting a split file. All split files invoke λS on their creation and resultant sorted files are stored in the "/to_join" folder of S3 bucket. Hence λS instances will run in parallel. The coordination between *BL_ReadAndSplit* and *Sort* is done through AWS S3 event notification. In the timeline, one *ReadAndSplit* Lambda function reads a large file and creates multiple small files. Small files will be created in sequence one after another and each of these files will also get sorted and stored in a new S3 bucket folder as shown in Figure 6. Choreography of all these tasks is fully event-driven.

4.3.3 Parallel Reactive Merge is a serverless solution to initiate merging of two sorted files of any size. This implementation solves the challenge of merging files when the size of sorted files starts growing and the execution time-out limit reached before the merging process completes. So Parallel Reactive Merge overcomes the serverless execution time-out constraint. For this, two Lambda functions are implemented: 1) MasterReactiveMerge(λMRM) 2) WorkerRecursiveMerge(λWRM).

In this implementation, a parallel reactive merging process is devised that will complete the merging without being time-out. A Lambda function MasterReactiveMerge is invoked whenever a sorted file is dropped in the *to_join* folder of S3 bucket. This Lambda function picks up another file available in *to_join* for merging, moves these two files to the *to_merge* folder and invokes the first instance of WRM. The WRM merges records from two sorted files and writes them to a new file. The WRM takes input parameters as: i) Working Directory, ii) File1, iii) File2 iv) Start Position in File1, v) Start Position in File2, vi) Output File, and vii) Byte Threshold Value. First invocation of WRM carries Start Position for both the input files as zero. Before getting time-out, WRM calls itself and the next instance starts appending records to the same file created by the first instance. Hence λWRM, recursively calls itself till the single merged file is created as depicted in Figure 8. Last instance of λWRM moves the merged file to the "/to_join" folder and deletes both the input files. One instance of λMRM followed by one or more instances of λWRM is initiated for every pair of files in "/to_join" and this parallel merging process continues till a single file is left in the "/to_join" folder.

5. Evaluation

In this section, we will describe the quantitative assessment for our proposed design model. This will facilitate us to demonstrate the feasibility of our algorithmic approach in designing fully choregraphed Fork-Join workflow in a serverless architecture. This research work also led us to find answers to aforementioned research questions

RQ1: Serverless design using *Relay Composition* pattern can be successfully used for long running computation. Our Fork component allows us to conclude that use cases where processing time might exceed the execution time limit can also make use of serverless technology. Results of our experiments are presented in Table 4 where *MaxSplitSize* is taken as 450 KB.

Table 4
Experimental Results - Single Lambda v/s Relay Composition pattern

Input File Size (KB)	Design Used	Time Out	No. of Files	(No. of Files x *MaxSplitSize* (KB)) + Size of last split file (KB)
200	*SF	No	1	N/A
300	SF	No	1	N/A
600	**SF**	**Yes**	**1**	**N/A**
600	**RC	No	2	(1x450) + 150
700	RC	No	2	(1x450) + 250
800	RC	No	2	(1x450) + 350
1000	RC	No	3	(2x450) + 100
1200	RC	No	3	(2x450) + 300
1400	RC	No	4	(3x450) + 50
1600	RC	No	4	(3x450) + 250

*SF- Single FaaS

**RC- *Relay Composition*

RQ2: Execution a complex, burst-parallel workflow can be addressed programmatically without using any external orchestration service. The algorithmic approach is a priori more powerful considering the availability of basic control flow instructions in any serverless runtime [30]. Our design model can handle burst-parallel workload by utilizing auto scalability of serverless alongwith algorithmic design. Successful implementation of Fork-Join workflow without using any external orchestration service proves this argument.Results of SSM pipeline based on *ReactiveFnJ* are given in Figure 9.

RQ3: It is possible to design a serverless system using pure event-driven architecture for complex workflows. The *ReactiveFnJ*, a pure event-driven system, using two novel serverless design patterns i.e., *Relay Composition* and *Master-Worker Composition*. These patterns are used in prototypical implementation of SSM pipeline that initiates when an input file gets uploaded in the AWS S3 bucket. Subsequent steps like splitting, sorting, and merging are triggered automatically without any human arbitration. This shows that complex parallel systems can be designed in a pure event-driven architecture and hence answers RQ3.

To summarize our results, the *ReactiveFnJ* is a choreographed design model for the development of distributed applications based on Fork-Join workflow with serverless architectures. This design can handle long running tasks by overcoming execution time-out constraint and is not dependent on any external orchestration service for its function composition.

6. Discussion And Lessons

Our model for Fork-Join workflow can be best utilized in serverless technology as this technology provides auto-scalability, built-in fault tolerance, availability, and abstraction of underlying infrastructure. To make Fork-Join workflow available, we build *ReactiveFnJ*, a serverless composition model. The *ReactiveFnJ* uses an innovative design that is purely event-driven, reactive, ST-Safe, choreographed composition model that conquers the hard time-limit forced by serverless environment. This model is completely generic, and several other workflows/use cases can be implemented by substituting specialized FaaS functions in SSM pipeline.

With the intent to develop other complex compute intensive and data parallel applications, patterns analogous to our innovative composition patterns i.e., *Relay Composition* and *Master-Worker Composition* should be discovered. We have articulated following general design guidelines that will help serverless system designers and architects to engineer their applications in an efficient and effective manner.

- Identification of iterative patterns in a long-running task. An iterative pattern can be implemented as a recursive serverless function where each function asynchronously invokes the next instance by relaying the appropriate parameters till the task is completed.

- Recognition of computing patterns that are data independent and hence, can be executed in parallel. These patterns can be implemented as serverless functions runnable in parallel exploiting on-demand scale features of serverless infrastructure.

- For coordination between serverless functions *Master-Worker Composition* is applied. This is a *Master-Worker Composition* where a master function is invoked automatically on an event and is responsible for invoking and dispatching parameters to Worker. The Worker recursively calls itself till the task is completed. In this composition Master is reactive in nature and Worker is recursive in its functioning

7. Related Work

Composition of serverless functions is sparsely covered in the current scientific literature but immensely important in practice [24]. Serverless frameworks must pave the way for complex function composition mechanisms to build responsive, compute intensive, and burst-parallel distributed serverless applications [8].

In some previous studies, researchers have explored and harnessed the power of FaaS in parallel processing applications. For massively parallel computations systems like PyWren [19] and ExCamera [11], have used their own external ad-hoc orchestration services to synchronize the parallel execution of cloud functions. PyWren utilized a polling mechanism for the Amazon S3 bucket to consolidate its

results. ExCamera configured an additional external server to synchronize the interactions of parallel running executions. Dai et al., [10] proposed a trigger-based computing framework called 'Domino'. They exploited both synchronous and asynchronous mechanisms to coordinate triggers. Workflow management systems like Apache Airflow, Oozie [18] and Camunda [32] exist but they depend on a dedicated long-running stateful execution engine to handle the state machine of the orchestration. Different sequential workflows in varied language runtime environments are implemented in [7] on AWS serverless platform and IBM Cloud Function. In our previous work, we proposed a design approach that exploits data parallelism in serverless infrastructure for massively parallel computations. In this design approach, we implemented a running prototype in AWS Lambda for distributed matrix multiplication use case [6].

Serverless popular function orchestration systems are- AWS Step Functions, Azure Durable Functions, IBM Cloud Functions, and Google Cloud Functions. They render help to create services involving compositions of serverless functions. But currently they do not support Parallel Fork-Join workflows natively [4]. As per L´opez et al., [2] serverless applications must follow a trigger-based interaction mechanism. They expressed that the FaaS orchestration system should also be event-triggered. It means, in a complex business workflow, the termination of one FaaS should trigger the next function in the pipeline using asynchronous events. Witte et al., [33] illustrated a cloud specific serverless implementation for seismic imaging application. They use AWS Step Function visual workflow for their implementation and exploit the mathematical properties of imaging optimization problem. Shankar et al., [31] developed a serverless system Numpywren for highly parallel linear algebra algorithms. For their implementation, authors introduced a new domain-specific language "LAmbdaPACK". Zhang et al., [36] proposed a new serverless framework "Kappa", that used a check-pointing mechanism to control function execution time-out. This framework simplifies application development and provides concurrency mechanisms to coordinate parallel computations.

Concurrency and parallel computation of serverless are well suited for requirements like sorting a huge data file. Efficient sorting of large data sets is an extensively discussed area as it is central to large businesses, banks, and institutions. Sorting counts for roughly one-fourth of the total computer cycles [25]. Usually, if a data file is small and can fit into the memory limits of a FaaS container, it can be easily programmed for sorting. But in case a file is huge then sorting by a single FaaS will time-out before its completion. Due to resource limitations of a serverless environment, processing of very large files is not supported directly. Recently, Amazon EFS volume mounted to handle such large files [26]. However, latency and additional cost are some challenges around this way-out.

After investigating previous and recent research works, a clear gap surfaced regarding the non-existence of reactive Fork-Join workflow runnable for serverless infrastructures. To the best of our knowledge, ours is the first work that presents the design and implementation of *ReactiveFnJ* - a pure reactive, choreographed, ST-Safe, algorithmic solution and answer the intrinsic constraint of serverless environments. This makes the best use of the scalability and fault tolerance feature of serverless.

ReactiveFnJ is judicious to be used for large-sized datasets where divide and conquer strategies can be applied.

8. Conclusions And Future Work

Current serverless technology provides its users with fine billing granularity and affordability to run arbitrary functions on-demand. We can reap great benefits of serverless when it comes to the runtime scalability of functions. Subtle barriers of this architecture are resource limit of functions, designing complex applications that require state and composability of long running functions.

This paper is an academic endeavour that designs a choreography model *ReactiveFnJ* for Fork-Join workflow in Serverless architectures. This model is pure event-driven, ST-Safe, not limited to serverless time constraints and makes the best use of scalability and fault tolerance features of serverless. This research proves the viability of a serverless design for a complex burst-parallel workflow where every serverless challenge is resolved at algorithmic level rather than using any external orchestration service, shared memory, or messaging services for its task scheduling and synchronization. This solution is platform independent as a result it is free from vendor lock-in problem also. Our model will render help to serverless architects and developers in fabricating compositions based on Fork-Join workflows. We contemplate that each component of the Fork-Join pipeline can be substituted by a variety of serverless functions to generate diverse workflows. The proof-of-concept and evaluation results authenticate that *ReactiveFnJ* can deliver sufficient performance and suggests its adoption in serverless frameworks for large-scale distributed computing.

Abbreviations

FaaS: Function as a Service; ST: Serverless Trilemma; SSM: Split-Sort-Merge; MSS: MaxSplitSize; MWC: Master-Worker Composition; MRM: MasterReactiveMerge; WRM: WorkerRecursiveMerge.

Declarations

Acknowledgements

Not applicable.

Authors’ contributions

All authors take part in the discussion of the work described in this paper. Urmil Bharti contributed to the model design and validation experiment of this work. She drafted the manuscript and coordinated the review task among authors. The author(s) read and approved the final manuscript.

Authors’ information

Ms. Urmil Bharti has done B.Sc. Computer Science (University of Delhi, India), M.Sc. Computer Science (DAVV Indore, India) and M.Phil. Computer Science (MKU, India). She has over 14 years of teaching experience as Assistant Professor in constituent colleges of University of Delhi. Earlier she worked in IT industry for more than 10 years. Her last designation in industry was Senior Quality Analyst. She is currently doing her research in Cloud and Distributed Computing. Her key research area is cloud computing, serverless technology and software engineering. She has authored several national and international research publications.

Dr. Anita Goel is an Associate Professor in Department of Computer Science, Dyal Singh College, University of Delhi, India. She has received her Ph.D. in Computer Science and Masters in Computer Applications from Jamia Millia Islamia and Department of Computer Science (University of Delhi), respectively. She has a work experience of more than 30 years. She is a visiting faculty to Delhi Technological University and NIIT University. From 2009-10, she was Fellow in Computer Science, at Institute of Life Long Learning (ILLL) in University of Delhi. She has served as member of program committee of International conferences like IEEE BigData Congress 2015 and ICWI 2015. She has guided several students for their doctoral studies and has travelled internationally to present research papers. She has authored books in Computer Science and has several national and international research publications.

Dr SC Gupta is B.Tech (EE) from IIT Delhi and has worked at Computer Group at Tata Institute of Fundamental Research and NCSDCT (now C-DAC Mumbai), Till recently, he worked as Deputy Director General, Scientist-G and Head of Training at National Informatics Centre, New Delhi and was responsible for keeping its 3000 scientists/ engineers uptodate in various technologies. He has extensive experience in design and development of large Complex Software Systems. Currently he is a Visiting Faculty at Dept of Computer Science and Engineering, IIT Delhi. His research interests include Software Engineering, Data Bases and Cloud Computing. He has been teaching Cloud Computing at IIT Delhi, which includes emerging disruptive technologies like SDN and SDS. He has guided many M.Tech & PhD Research students in these technologies and has many publications in Software Engineering and Cloud Technology in National and International Conferences and Journals.

Funding

Not applicable.

Availability of data and materials

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Arfat Y, Usman S, Mehmood R, Katib I (2020) Big data for smart infrastructure design: Opportunities and challenges. Smart Infrastruct Appl 491–518

2. Arjona A, López PG, Sampé J, Slominski A, Villard L (2021) Triggerflow: Trigger-based orchestration of serverless workflows. Futur Gener Comput Syst 124:215–229. doi: 10.1016/j.future.2021.06.004

3. Baldini I, Cheng P, Fink SJ, Mitchell N, Muthusamy V, Rabbah R, Suter P, Tardieu O (2017) The serverless trilemma: Function composition for serverless computing. Onward! 2017 - Proc 2017 ACM SIGPLAN Int Symp New Ideas, New Paradig Reflections Program Software, co-located with SPLASH 2017 89–103. doi: 10.1145/3133850.3133855

4. Barcelona-Pons D, García-López P, Ruiz Á, Gómez-Gómez A, París G, Sánchez-Artigas M (2019) FaAS orchestration of parallel workloads. In: WOSC 2019 - Proceedings of the 2019 5th International Workshop on Serverless Computing, Part of Middleware 2019. pp 25–30

5. Barcelona-Pons D, Sánchez-Artigas M, París G, Sutra P, García-López P (2019) On the FaaS track: Building stateful distributed applications with serverless architectures. Middlew 2019 - Proc 2019 20th Int Middlew Conf 41–54. doi: 10.1145/3361525.3361535

6. Bharti U, Bajaj D, Goel A, Gupta SC (2021) A Novel Design Approach Exploiting Data Parallelism in Serverless Infrastructure. In: Advances in Computing and Network Communications. Springer, pp 247–260

7. Bharti U, Bajaj D, Goel A, Gupta SC (2021) Sequential Workflow in Production Serverless FaaS Orchestration Platform. In: Proceedings of International Conference on Intelligent Computing, Information and Control Systems. pp 681–693

8. Carver B, Zhang J, Wang A, Anwar A, Wu P, Cheng Y (2020) Wukong: A scalable and locality-enhanced framework for serverless parallel computing. In: SoCC 2020 - Proceedings of the 2020 ACM Symposium on Cloud Computing. Association for Computing Machinery, Inc, pp 1–15

9. Christidis A, Davies R, Moschoyiannis S (2019) Serving machine learning workloads in resource constrained environments: A serverless deployment example. Proc - 2019 IEEE 12th Conf Serv Comput Appl SOCA 2019 55–63. doi: 10.1109/SOCA.2019.00016

10. Dai D, Chen Y, Kimpe D, Ross RB (2021) Trigger-Based Incremental Data Processing with Unified Sync and Async Model. IEEE Trans Cloud Comput 9:372–385. doi: 10.1109/TCC.2018.2830348

11. Fouladi S, Wahby RS, Shacklett B, Balasubramaniam KV, Zeng W, Bhalerao R, Sivaraman A, Porter G, Winstein K (2017) Encoding, fast and slow: Low-latency video processing using thousands of tiny threads. Proc 14th USENIX Symp Networked Syst Des Implementation, NSDI 2017 363–376

12. García-López P, Sánchez-Artigas M, Shillaker S, Pietzuch P, Breitgand D, Vernik G, Sutra P, Tarrant T, Ferrer AJ (2019) ServerMix: Tradeoffs and Challenges of Serverless Data Analytics. pp 1–15

13. Garcia Lopez P, Sanchez-Artigas M, Paris G, Barcelona Pons D, Ruiz Ollobarren A, Arroyo Pinto D (2019) Comparison of FaaS orchestration systems. In: Proceedings - 11th IEEE/ACM International Conference on Utility and Cloud Computing Companion, UCC Companion 2018. pp 109–114

14. Giménez-Alventosa V, Moltó G, Caballer M (2019) A framework and a performance assessment for serverless MapReduce on AWS Lambda. Futur Gener Comput Syst 97:259–274. doi: 10.1016/j.future.2019.02.057

15. Hassan HB, Barakat SA, Sarhan QI (2021) Survey on serverless computing. J Cloud Comput 10. doi: 10.1186/s13677-021-00253-7

16. Hellerstein JM, Faleiro J, Gonzalez JE, Schleier-Smith J, Sreekanti V, Tumanov A, Wu C (2019) Serverless computing: One step forward, two steps back. In: CIDR 2019 - 9th Biennial Conference on Innovative Data Systems Research

17. Holubiev V, Ihnatiuk B, Voytyuk I (2018) Next-generation serverless system for contextual search based on rich media content. CEUR Workshop Proc 2300:211–214

18. Islam M, Huang AK, Battisha M, Chiang M, Srinivasan S, Peters C, Neumann A, Abdelnur A (2012) Oozie: towards a scalable workflow management system for hadoop. In: Proceedings of the 1st ACM SIGMOD Workshop on Scalable Workflow Execution Engines and Technologies. pp 1–10

19. Jonas E, Pu Q, Venkataraman S, Stoica I, Recht B (2017) Occupy the cloud: Distributed computing for the 99%. SoCC 2017 - Proc 2017 Symp Cloud Comput 445–451. doi: 10.1145/3127479.3128601

20. Klimovic A, Wang Y, Kozyrakis C, Stuedi P, Pfefferle J, Trivedi A (2020) Understanding ephemeral storage for serverless analytics. Proc 2018 USENIX Annu Tech Conf USENIX ATC 2018 789–794

21. Kuhlenkamp J, Werner S, Tai S (2020) The Ifs and Buts of Less is More: A Serverless Computing Reality Check. In: Proceedings - 2020 IEEE International Conference on Cloud Engineering, IC2E 2020. pp 154–161

22. Landset S, Khoshgoftaar TM, Richter AN, Hasanin T (2015) A survey of open source tools for machine learning with big data in the Hadoop ecosystem. J Big Data 2:1–36. doi: 10.1186/s40537-015-0032-1

23. Leite LAF, Oliva GA, Nogueira GM, Gerosa MA, Kon F, Milojicic DS (2013) A systematic literature review of service choreography adaptation. Serv Oriented Comput Appl 7:199–216

24. Leitner P, Wittern E, Spillner J, Hummer W (2019) A mixed-method empirical study of Function-as-a-Service software development in industrial practice. J Syst Softw 149:340–359. doi: 10.1016/j.jss.2018.12.013

25. Leu F-C, Tsai Y-T, Tang CY (2000) An efficient external sorting algorithm. Inf Process Lett 75:159–163

26. Obrutsky S (2016) Cloud Storage : Advantages , Disadvantages and Enterprise Solutions for Business. EIT New Zeal

27. Pu Q, Berkeley UC, Venkataraman S, Stoica I, Berkeley UC, Nsdi I (2019) Shuffling , Fast and Slow : Scalable Analytics on Serverless Infrastructure This paper is included in the Proceedings of the. In: Nsdi

28. Ristov S, Pedratscher S, Fahringer T (2021) AFCL: An Abstract Function Choreography Language for serverless workflow specification. Futur Gener Comput Syst 114:368–382. doi: 10.1016/j.future.2020.08.012

29. Rizk A, Poloczek F, Ciucu F (2015) Computable bounds in fork-join queueing systems. Perform Eval Rev 43:335–346. doi: 10.1145/2796314.2745859

30. Sampé J, Vernik G, Sánchez-Artigas M, García-López P (2018) Serverless data analytics in the IBM cloud. Middlew Ind 2018 - Proc 2018 ACM/IFIP/USENIX Middlew Conf (Industrial Track) 1–8. doi: 10.1145/3284028.3284029

31. Shankar V, Krauth K, Pu Q, Jonas E, Venkataraman S, Stoica I, Recht B, Ragan-Kelley J (2018) Numpywren: Serverless Linear Algebra

32. Wiemuth M, Burgert O (2019) A workflow management system for the OR based on the OMG standards BPMN, CMMN, and DMN. In: Medical Imaging 2019: Image-Guided Procedures, Robotic Interventions, and Modeling. p 1095127

33. Witte PA, Louboutin M, Modzelewski H, Jones C, Selvage J, Herrmann FJ (2020) An event-driven approach to serverless seismic imaging in the cloud. IEEE Trans Parallel Distrib Syst 31:2032–2049

34. Yu T, Liu Q, Du D, Xia Y, Zang B, Lu Z, Yang P, Qin C, Chen H (2020) Characterizing serverless platforms with serverlessbench. In: SoCC 2020 - Proceedings of the 2020 ACM Symposium on Cloud Computing. pp 30–44

35. Zahoor E, Asma Z, Perrin O (2017) A formal approach for the verification of AWS IAM access control policies. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Springer, Cham, pp 59–74

36. Zhang W, Fang V, Panda A, Shenker S (2020) Kappa: A programming framework for serverless computing. In: Proceedings of the 11th ACM Symposium on Cloud Computing. pp 328–343

Figures

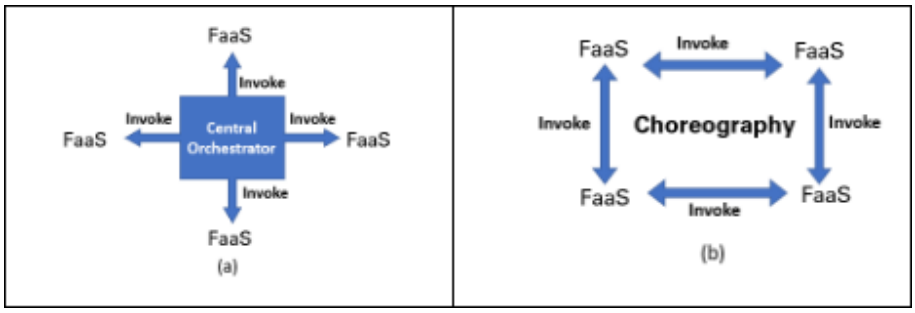


Figure 1

Function Composition Approaches (a) Orchestration (b) Choreography

Figure 2

Fork-Process-Join pipeline

Figure 3

Relay Composition used in spitting a big file

Figure 4

Join Component – Binary Logarithmic Growth Curve

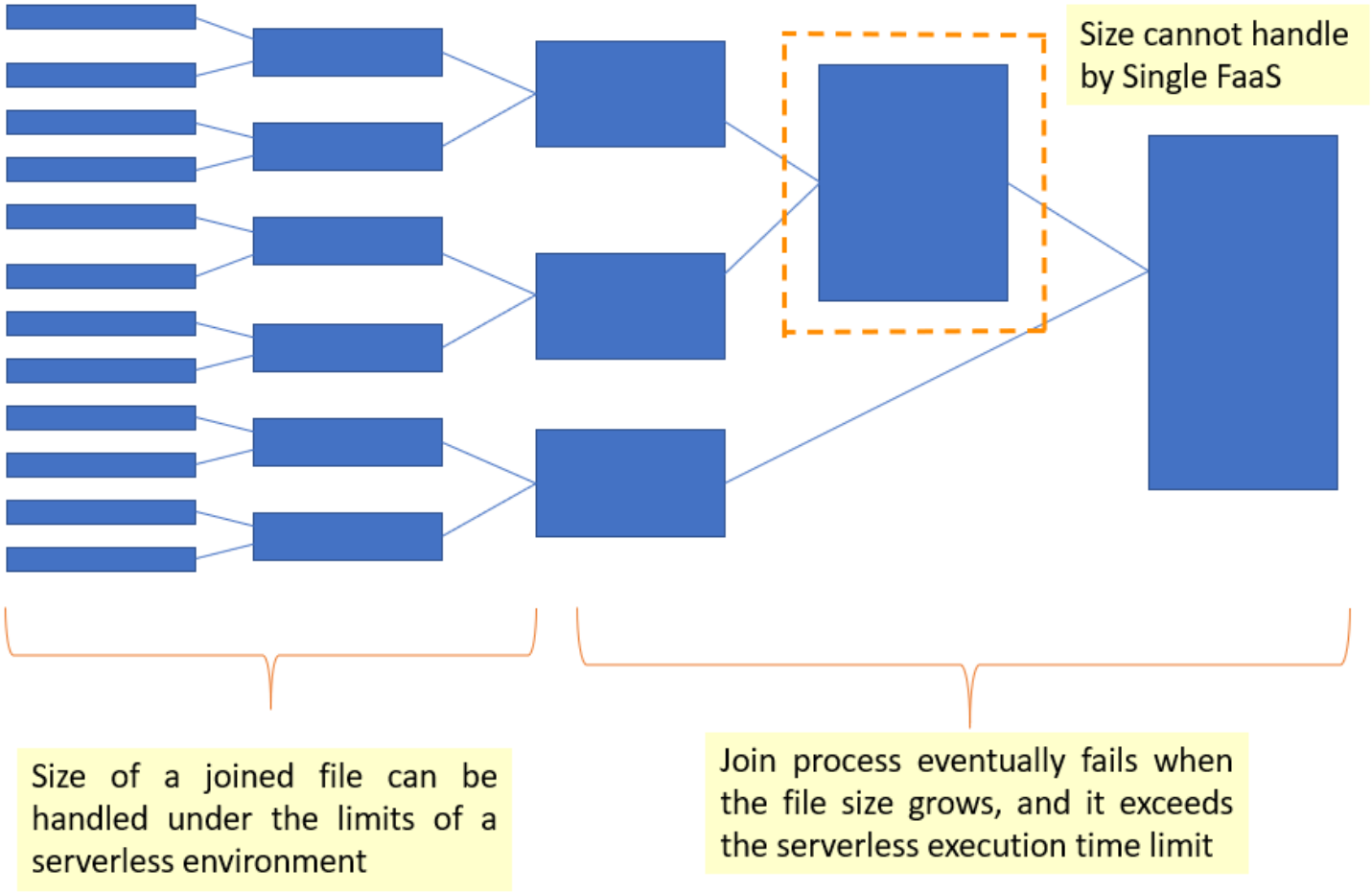


Figure 5

Conventional method to Join N files in Parallel

Figure 6

The SSM Architecture

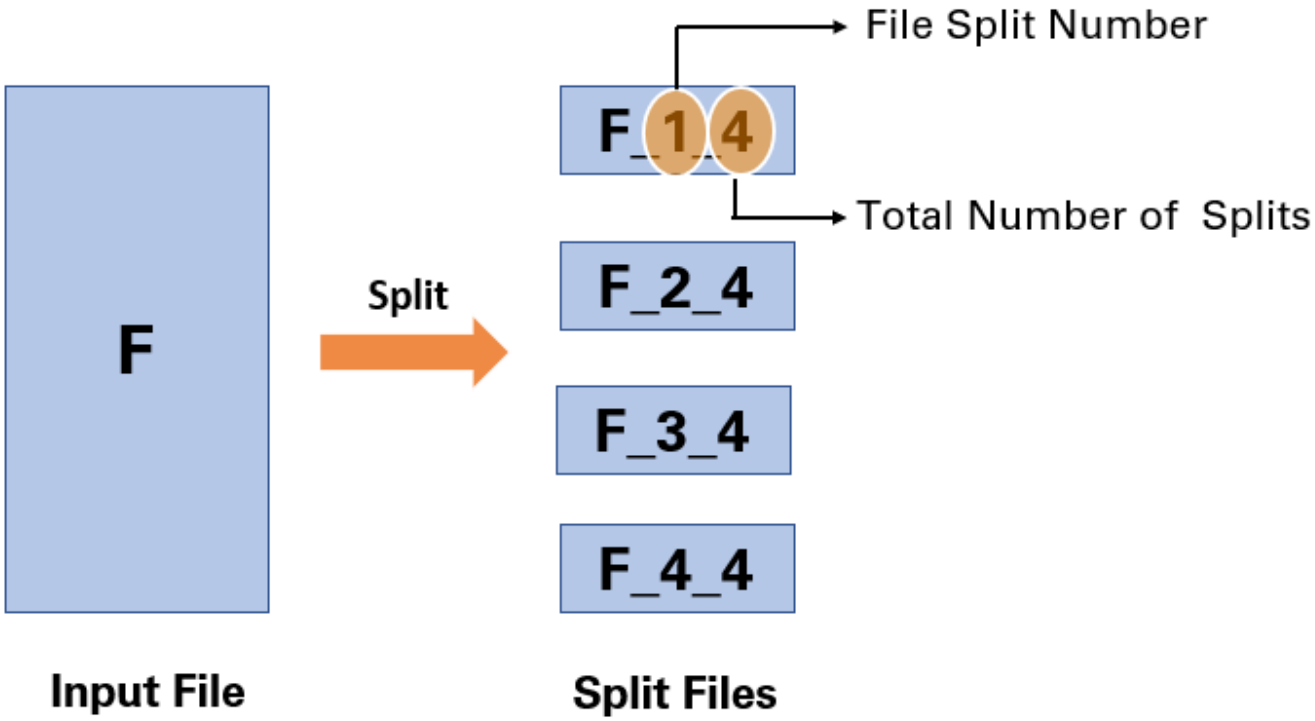


Figure 7

Split File Naming convention

Figure 8

(a) Parallel Reactive Merge (b) File Merge by Worker Recursive Merge

Figure 9

Executions Results of Single Lambda vs FnJ Pipeline

AIP Advances

Reflectocardiography (RCG): A leadless, optical reflection based facile measurement of human cardiac rhythm

Cite as: AIP Advances **12**, 015113 (2022); https://doi.org/10.1063/5.0076791
Submitted: 28 October 2021 • Accepted: 25 December 2021 • Published Online: 10 January 2022

Arindam Kushagra, Uddipan Dasgupta, Srishti Roychowdhury, et al.

View Online Export Citation CrossMark

ARTICLES YOU MAY BE INTERESTED IN

Experimental study of the dispersion of cough-generated droplets from a person going up- or downstairs
AIP Advances **12**, 015002 (2022); https://doi.org/10.1063/5.0073880

Giant parametric amplification and spectral narrowing in atomically thin MoS_2 nanomechanical resonators
Applied Physics Reviews **9**, 011404 (2022); https://doi.org/10.1063/5.0045106

Airborne virus transmission under different weather conditions
AIP Advances **12**, 015019 (2022); https://doi.org/10.1063/5.0082017



AIP Publishing

© 2022 Author(s).

AIP Advances ARTICLE scitation.org/journal/adv

Reflectocardiography (RCG): A leadless, optical reflection based facile measurement of human cardiac rhythm

Cite as: AIP Advances 12, 015113 (2022); doi: 10.1063/5.0076791
Submitted: 28 October 2021 • Accepted: 25 December 2021 •
Published Online: 10 January 2022

View Online Export Citation CrossMark

Arindam Kushagra,[1,a)] Uddipan Dasgupta,[1] Srishti Roychowdhury,[2] Srijani Samanta,[3] and Shivani Srivastava[4]

AFFILIATIONS

[1] Amity Institute of Nanotechnology, Amity University Kolkata, Kolkata 700135, India
[2] Amity Institute of Biotechnology, Amity University Kolkata, Kolkata 700135, India
[3] Department of Biotechnology, Delhi Technological University, Delhi 110042, India
[4] Department of Mathematics, Techno Main Salt Lake Campus, Kolkata 700091, India

[a)] **Author to whom correspondence should be addressed:** arindam.kushagra@gmail.com

ABSTRACT

Cardiac diseases have been one of the leading causes of death in developed and developing countries. This can be attributed to the prevailing eating habits as well as unhealthy work–life balance and lifestyle. There have been many instances where scientists, researchers, and innovators working in the domains of engineering, biotechnology, and healthcare have proposed novel ways to monitor the cardiac activity of a diseased person. Such monitoring systems started with the advent of electrocardiography followed by other methods, such as echocardiography, phonocardiography, and photoplethysmography, which led to better patient care. Recently, there has been growing interest to monitor the cardiac activity by mechanical movements of the body parts/surfaces, which has led to the beginning of new fields, such as gyrocardiography, ballistocardiography, seismocardiography, and forcecardiography. All these mechanical sensors have very sophisticated components attached to them, thus making them relatively less available in low-cost settings. In this work, we propose a proof-of-concept to optically measure the seismic thumps due to the heartbeats by keeping a reflecting surface over the epigastric region of the abdominopelvic cavity. Due to these thumps, the reflecting surface moved vertically and the position of the image of an otherwise stationary object in the reflecting surface was monitored using a camera. This mechanism is an inexpensive alternative to replace the costly MEMS components that constitute mechanical cardiograms and would reduce the cost of measurements substantially. It would be of great interest to scientists and healthcare researchers to explore the viable alternatives to low-cost settings.

© 2022 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/). https://doi.org/10.1063/5.0076791

INTRODUCTION

Presently, maximum clinical fatalities are caused by cardiac diseases.[1] Many researchers worldwide have been working toward the alleviation and early detection of cardiac malfunctioning.[2] However, when the heart starts malfunctioning, it becomes extremely important to detect the cardiac rhythms to monitor its activity for an individual's well-being. Conventionally, the cardiac activity is monitored by electrocardiography (ECG), a process which requires electrical leads to be pasted to certain regions of the subject's body.[3] The involvement of leads increases the complexity of measuring the cardiac activity owing to multiple sources of error (in the case of improper electrical connections to and from the leads).[4] To avoid such complications arising due to poor electrical connections in the ECG, the field was developed further to use sound waves to monitor the cardiac activities and this phenomenon is termed echocardiography or phonocardiography (PCG).[5] Other advanced methods have been developed and incorporated into wearable devices for the ease in monitoring of cardiac activities of the test subject without affecting the lifestyle of the subject.[6] These methods include the measurements of the dissolved oxygen concentration in the blood (SpO_2) measured by pulse oximetry[7] and

measurement of the pulse rate of the heartbeat by monitoring the reflections of IR light from the arteries/veins by monitoring the volume of blood flowing through them; this technique is termed photoplethysmography (PPG).[8] However, with the advancements in technological fields, researchers have been able to monitor the mechanical vibrations on the surface of our body caused by the heartbeats. Techniques such as gyrocardiography (which is done by placing a gyroscope, a MEMS device, on the chest wall),[9,10] ballistocardiography (measurement of recoil forces of the body in reaction to cardiac ejection of blood into the vasculature),[11,12] seismocardiography (measurement of local chest vibrations due to heartbeat),[13,14] and forcecardiography[15] fall under the purview of mechanical detection of cardiac activities. All the four listed techniques that utilize mechanical detection strategies employ sophisticated MEMS devices, such as accelerometers and gyroscopes. Monitoring cardiac activities and their performance would be severely hampered if there is a slight misalignment in these MEMS devices. The proposed invention overcomes this limitation of lack of functionality by having an optical readout mechanism that would use a light source, an optical detector, and a reflecting surface. It is a facile way to detect the cardiac activities by using a simple optical setup, thus bypassing the need of sophisticated MEMS devices.

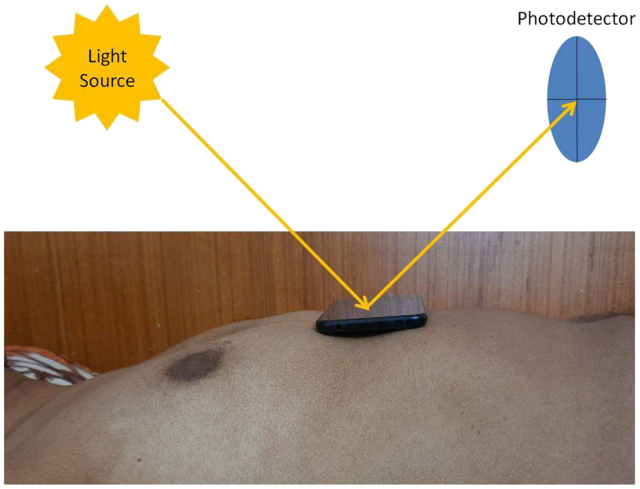


FIG. 1. Measurement schematic showing the reflecting surface put on the epigastric region of the abdominopelvic cavity. The object whose reflection was monitored for the recording is depicted by a light source, and the recording camera has been depicted as a photodetector.

MATERIALS AND METHODS

Seven male human subjects of east-central Indian ethnicity (age group 20–55 years) were laid supine with the reflecting surface kept on the epigastric region of the abdominopelvic cavity. A black reflecting surface was used to detect the cardiac rhythm, as shown in Fig. 1. The seismic thumps on the body surface were detected by means of the deflection of the image of an object, and the observations were recorded using an open-source screen recorder, OBS Studio[16] at 30 fps.

The employed light source can be a LED bulb or any sharply pointed object whose reflection can be tracked using a simple video recording device, such as a laptop/mobile camera, which acts as a photodetector in this work. The measurements were done in two conditions, with and without clothes on the abdominopelvic regions of the test subject. Video recordings were done using an open-source software program, OBS Studio at 30 fps. Each frame was extracted using another open-source software program, FFmpeg.[17] The movement of the object in each frame was traced using an image processing software program, ImageJ,[18] and the respective vertical coordinates were plotted against the passing time, thus generating a cardiograph, which can be further utilized for analysis of cardiac rhythms to ascertain the well-being of an individual.

The employed methodology and the reported protocol in this work (which is leadless and non-invasive and requires simple putting of the reflecting surface on the abdominopelvic cavity) are to measure the seismic vibrations experienced on the body surface due to the cardiac activities (therefore, can be considered as a variation of seismocardiography)[19] and thus adhere to the guidelines as approved by the ethical committee of Northern Jutland (N-20120069). The measurements were recorded after obtaining informed consent of the seven test subjects. However, only *Subject 01* consented for shirtless recording.

Summarily, the reported reflectocardiography (RCG) protocol can be conducted by any person able to handle a simple recording device (such as a mobile phone camera) in any setting. The specification table, mentioning the requirements to perform RCG measurements, is given in Table I.

TABLE I. Specification table, enlisting the parameters required to perform the RCG measurements.

Measurable attribute	Specifications
Subject's anatomical position	Supine, with head slightly tilted upward by ~15 cm using a pillow/headrest
Anatomical placement of the reflecting surface	Epigastric (sub-sternal) region of the abdominopelvic cavity
Size of the reflecting surface	Diagonally ~6 in.
Height of the stationary object whose reflection is monitored	~30–200 cm
Recording device	Any simple video recording device (e.g., a mobile phone camera)

© Author(s) 2022

RESULTS AND DISCUSSION

This work discusses a unique strategy to monitor the cardiac activities. The RCG protocol reported in this paper has an advantage over the prior discussed mechanical vibration monitoring techniques due to the use of an optical setup, thus bypassing the need to design and fabricate complicated MEMS devices using costly setups. The optical setup demonstrated in the present invention includes a fixed object whose reflection is continuously monitored using a camera, and the mechanical vibrations on the body surface are detected by the movement of the reflecting surface via the said camera. Due to the said movements of the reflecting surface, the image of the fixed object would also move, which would give us a video-graphic trend of the heartbeats of the subject. The movement of the reflection (image) of the fixed object is further analyzed by an image processing software program to get the image coordinates. These coordinates are then plotted against time to get a dynamic response of the cardiac activity for *Subject 01*, showing trends analogous to the ECG (as shown in Figs. 2 and 3).

However, the baseline correction was done with the help of Origin 8.5®, a graph plotting and data analysis software program, by going for the baseline correction mode after clicking the relevant tabs as per the following steps: *Analysis → Peaks and Baseline → Peak Analyzer → Open Dialogue → Recalculate (Auto) → Goal (Subtract Baseline) → Next → Baseline Mode (Asymmetric Least Squares Smoothing) → Next → Asymmetric Factor (0.001), Threshold (0.05), Smoothing Factor (4), Number of Iterations (10) → Next → Subtract → Finish*. After the baseline correction was done using the aforementioned steps, the modified RCG patterns obtained are shown in Fig. 3.

This subject showed normal cardiac functions, as the duration between subsequent peaks lie well within the normal range of 0.6 –1.2 s.[20] Video clips of the recordings of all the seven subjects of the movement of the reflecting surface along with the generated RCG

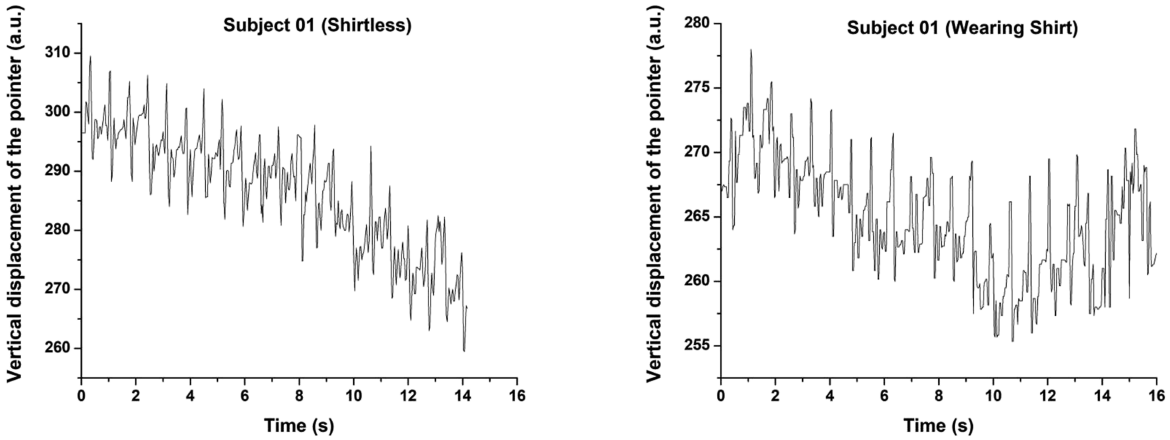


FIG. 2. The shirtless (left panel) and with shirt on (right panel) raw cardiographs (without baseline correction) depicting the rhythms of heartbeat employing the proposed RCG measurements on *Subject 01* have been shown. The peaks denote the ventricular depolarization events, with RR intervals lying between 0.6 and 1 s as marked in the diagrams.

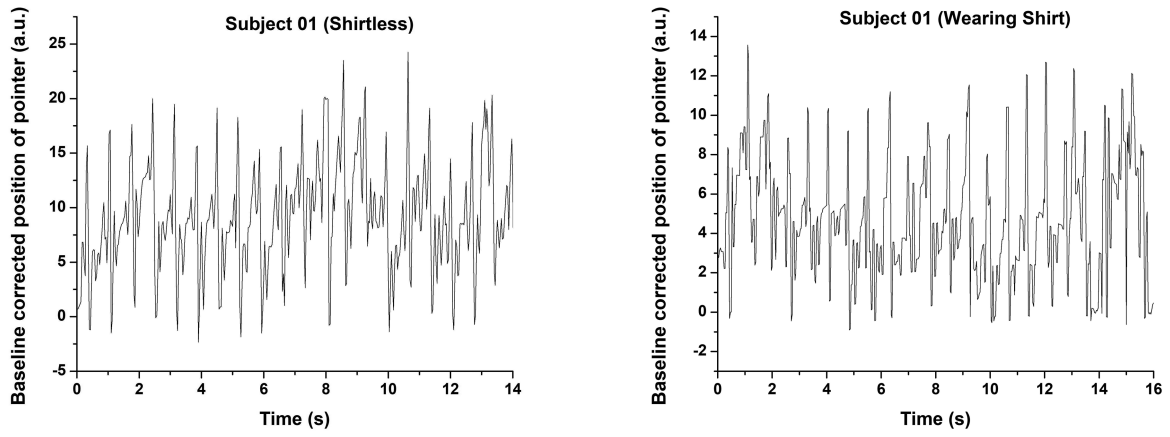


FIG. 3. The shirtless (left panel) and with shirt on (right panel) modified cardiographs (with baseline correction) depicting the rhythms of heartbeat employing the proposed RCG measurements on *Subject 01* have been shown.

© Author(s) 2022

TABLE II. Mean and standard deviation values of measured RR intervals for seven different subjects, following the reported RCG protocol.

Subject no.	Mean (s)	Standard deviation (s)	Age (years)	Gender
		Shirtless		
01	0.681 667	0.036 635	37	Male
		Wearing shirt		
01	0.780 702	0.129 727	37	Male
02	1.105 263	0.089 762	22	Male
03	1.101 667	0.088 836	26	Male
04	1.154 386	0.134 811	45	Male
05	1.147 917	0.179 699	56	Male
06	1.087 719	0.118 223	29	Male
07	1.162 963	0.154 196	32	Male

patterns are given as the supplementary material. The measurement principle of the proposed setup works on the identification of seismic thumps on the abdominal cavity by monitoring the movement of the reflection of a stationary object. When the said movement is traced with passing time, a cardiograph is generated. Since the proposed measurement protocol of cardiac rhythm uses the monitoring of the reflection of a stationary object, we term this type of measurement as *reflectocardiography (RCG)*. The vibrational noise that appears in the cardiograph can be reduced/nullified by applying suitable filters during the image processing part of data analysis. The presented data show around 15–25 heartbeats as recorded by the reported measurement protocol. The measurement data (indicated by RR intervals, signifying ventricular depolarization), as recorded by following the reported RCG protocol, are given in Table II.

This method is superior to the already existing methods employing the monitoring of seismic thumps over the body surface due to cardiac activities because it employs a very simple setup comprising only a reflecting surface and a video recording device *in lieu* of complicated MEMS devices (used in gyrocardiography, ballistocardiography, and seismocardiography). The reflection of any sharply pointed fixed object in the surrounding can be taken as a reference, thus increasing the scope of its application in extremely low-cost settings with a lack of sophisticated instruments. The only measurement parameter to be considered is the monitoring of the reflection of the fixed object by the video recording device. Training to operate such a setup is also not required, as it does not involve any complicated setup and can be operated by any person capable of handling the video recording devices. It can also discern the ventricular depolarization events and their intervals (marked by RR intervals[20]), which has a direct impact in assessing the well-being of an individual suffering from either of the following ailments: bradycardia, tachycardia, ventricular fibrillation, atrial fibrillation, bigeminy, trigeminy, and premature ventricular contraction. As discussed earlier, the RCG protocol employs a leadless method to detect the cardiac activities with a capability to measure the same while the subject is wearing the clothes. This is a significant advantage over the ECG and can be suitable to be employed in special cases where sticking the electrodes to the skin is not feasible, e.g., burn patients. It is expected to support and, in special cases, even replace the conventional methods (e.g., ECG) to monitor cardiac activities. The baseline of the RCG protocol wandered due to the breathing activities of the test subjects; however, the said breathing activity did not interfere with the recording of seismic thumps over the body surface due to the cardiac activities. This baseline wandering can be corrected by employing suitable algorithms and filters that are regularly being used in biomedical instrumentation. The reported RCG technique has an inherent limitation that pertains to the stiffness of the abdominopelvic cavity. The said stiffness, which could be majorly due to some muscular distress as well as excess body fat, would impede the transmission of seismic thumps due to cardiac activity to the body surface. This would lead to a substantial decrease in the amplitude of the detected thumps. To overcome this limitation, another suitable anatomical location (e.g., jugular vein) can be chosen by the researchers and healthcare professionals to increase the efficacy of the reported RCG protocol. In this paper, we have reported a qualitative study to monitor cardiac activities by means of optical reflection-based detection of seismic thumps experienced over the body cavity and it has inherent advantages over the other reported methods owing to its extremely low-cost and ease of monitoring as highlighted in the specification table (Table I). However, quantitative estimations pertaining to the efficacy of the RCG method over ECG, seismocardiography, gyrocardiography, and ballistocardiography are beyond the scope of the current work and may be taken up by fellow researchers and scientists working in the field of healthcare applications.

CONCLUSIONS

In this work, we report a novel way to detect the cardiac rhythms in the human body using a very simple and cost-effective method by placing a reflective surface on the abdominopelvic cavity and track the seismic movements of the reflective surface by monitoring the reflection of an otherwise stationary object. The RCG protocol is significantly improved over the already known methods, such as gyrocardiography, ballistocardiography, and seismocardiography, that employ the monitoring of seismic movements on the body surface from the pumping of blood. Such techniques employ the micrometer-scale mechanical devices, MEMS, that monitor the

© Author(s) 2022

body movements and transduce the physical movements to a digital readout to record the cardiac activities. Our proposed method does not require any such MEMS devices to monitor the cardiac activity, rather it simply relies on the tracking of the image of a stationary object/light source and does not require any sophisticated handling of the setup and is a big step toward the development of point-of-care cardiac activity monitoring devices. This work would be of immense interest to healthcare professionals, biomedical engineers, and scientists working toward the goal of a sustainable healthcare infrastructure development in low-income and middle-income countries worldwide.

SUPPLEMENTARY MATERIAL

Video recordings and extracted RCG patterns of seven subjects have been incorporated in the associated supplementary material of this manuscript.

ACKNOWLEDGMENTS

A.K. acknowledges Dr. R. R. Bhattacharjee, HoD AINTK, for his valuable input during this work.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts of interest to disclose.

Ethics Approval

The employed methodology and the reported protocol in this work (which is leadless and non-invasive and requires simple placement of the reflecting surface on the abdominopelvic cavity) are to measure the seismic vibrations experienced on the body surface due to the cardiac activities and thus adhere to the guidelines approved by the ethical committee of Northern Jutland (Approval number: N-20120069 reported in Ref. 19). The measurements were recorded after obtaining informed consent from the test subjects.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1]Cardiovascular diseases, https://www.who.int/health-topics/cardiovascular-diseases; accessed on 08 October 2021.
[2]J. Stewart, G. Manmathan, and P. Wilkinson, "Primary prevention of cardiovascular disease: A review of contemporary guidance and literature," JRSM Cardiovasc. Dis. **6**, 2048004016687211 (2017).
[3]Electrocardiogram, https://www.hopkinsmedicine.org/health/treatment-tests-and-therapies/electrocardiogram; accessed on 08 October 2021.
[4]J. García-Niebla, P. Llontop-García, J. I. Valle-Racero *et al.*, "Technical mistakes during the acquisition of the electrocardiogram," Ann. Noninvasive Electrocardiol. **14**(4), 389–403 (2009).
[5]Echocardiogram, https://www.mayoclinic.org/tests-procedures/echocardiogram/about/pac-20393856; accessed on 08 October 2021.
[6]K. Bayoumy, M. Gaber, A. Elshafeey *et al.*, "Smart wearable devices in cardiovascular care: Where we are and how to move forward," Nat. Rev. Cardiol. **18**(8), 581–599 (2021).
[7]Pulse Oximetry, https://www.hopkinsmedicine.org/health/treatment-tests-and-therapies/pulse-oximetry; accessed on 08 October 2021.
[8]J. Allen, "Photoplethysmography and its application in clinical physiological measurement," Physiol. Meas. **28**(3), R1–R39 (2007).
[9]S. Siecinski, P. S. Kostka, and E. J. Tkacz, "Gyrocardiography: A review of the definition, history, waveform description, and applications," Sensors **20**(22), 6675 (2020).
[10]M. J. Tadi *et al.*, "Gyrocardiography: A new noninvasive monitoring method for the assessment of cardiac mechanics and the estimation of hemodynamic variables," Sci. Rep. **7**(1), 6823 (2017).
[11]L. Giovangrandi, O. T. Inan, R. M. Wiard *et al.*, "Ballistocardiography—A method worth revisiting," in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (IEEE, 2011), pp. 4279–4282.
[12]C.-S. Kim, S. L. Ober, M. S. McMurtry *et al.*, "Ballistocardiogram: Mechanism and potential for unobtrusive cardiovascular health monitoring," Sci. Rep. **6**, 31297 (2016).
[13]P. Castiglioni, A. Faini, G. Parati, and M. Di Rienzo, "Wearable seismocardiography," in *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (IEEE, 2007), pp. 3954–3957.
[14]M. Di Rienzo, E. Vaini, P. Castiglioni *et al.*, "Wearable seismocardiography: Towards a beat-by-beat assessment of cardiac mechanics in ambulant subjects," Auton. Neurosci. **178**(1–2), 50–59 (2013).
[15]E. Andreozzi, A. Fratini, D. Esposito *et al.*, "Forcecardiography: A novel technique to measure heart mechanical vibrations onto the chest wall," Sensors **20**(14), 3885 (2020).
[16]OBS Studio, https://obsproject.com/; accessed on 08 October 2021.
[17]FFmpeg, https://www.ffmpeg.org/; accessed on 08 October 2021.
[18]ImageJ, https://imagej.nih.gov/ij/; accessed on 08 October 2021.
[19]K. Sørensen, S. E. Schmidt, A. S. Jensen, P. Søgaard, and J. J. Struijk, "Definition of fiducial points in the normal seismocardiogram," Sci. Rep. **8**(1), 15455 (2018).
[20]Normal Electrocardiography (ECG) Intervals, https://emedicine.medscape.com/article/2172196-overview; accessed on 08 October 2021.

© Author(s) 2022

The current issue and full text archive of this journal is available on Emerald Insight at:
https://www.emerald.com/insight/2398-628X.htm

Resolving stress of university teacher: exploring role of workplace spirituality and constructive workplace deviance

Naval Garg
University School of Management and Entrepreneurship, Delhi Technological University, Delhi, India
Sarika Kumari
Department of Telecommunication, Ministry of Telecommunications, Delhi, India, and
B.K. Punia
Haryana School of Business, Guru Jambheshwar University of Science and Technology, Hisar, India

Received 24 April 2020
Revised 24 February 2021
8 September 2021
4 December 2021
Accepted 11 December 2021

Abstract

Purpose – The study explores the relationship between workplace spirituality and work stress among university teachers. It also investigates the mediating effect of constructive deviance amid the association between workplace spirituality and stress among Indian university teachers.

Design/methodology/approach – The association between spirituality and stress is studied using correlation and multiple regression. The mediating effect of constructive deviance is examined using the Sobel test and bootstrapping estimates using Hayes' PROCESS macro. The hierarchical regression is used to report direct and indirect effects.

Findings – Findings reveal a significant negative association between the six dimensions of workplace spirituality and stress. The results also concluded the mediation effect of constructive deviance, which means workplace spirituality promotes constructive deviance that influences educators' stress levels.

Originality/value – The study is based on primary data collected by the author. It is one of the first explorations of the mediating effect of constructive deviance in the relationship of six dimensions of workplace spirituality and work stress among teachers.

Keywords Workplace spirituality, Constructive deviance, Stress, Mediation

Paper type Research paper

Introduction

Traditionally, teaching occupation was considered a low-stress profession (French *et al.*, 1982), but the situation changed drastically with the commercialization of the education sector in India (Ravichandran and Rajendran, 2007). Not only Indian teachers but lecturers from many countries reported a high level of stress (e.g. Chaplain, 2008; Skaalvik and Skaalvik, 2015). Several factors contributed to the increased stress level of teachers (Newberry and Allsop, 2017). These factors included students' and parents' rising demands and aspirations, various regulatory authorities and regulations, and repeated quality audits by the National Assessment and Accreditation Council (NAAC), National Board of Accreditation (NBA) and National Institute of Ranking Framework (NIRF), etc. Another major issue that contributed to strained university educators is the overlapping of teaching, research, administration, counseling, academic leadership and other similar job assignments. Other statutory duties such as deputation for election, census or nonacademic works also add misery to a teacher's life. Parry *et al.* (2016, pp. 88–89) argued that "teachers are required to acquire many new skills to cope effectively in a very fast-changing society. They have to

South Asian Journal of Business Studies
© Emerald Publishing Limited
2398-628X
DOI 10.1108/SAJBS-04-2020-0118

equip themselves in terms of creativity, innovations, and critical thinking." Ghania *et al.* (2014) blamed students' misbehavior, workload, time and resource difficulties for highly stressed teaching fraternity. Teachers generally have higher workloads than other professionals (Hooftman *et al.*, 2015). Aggarwal (2012) highlighted the issue of role insufficiency and role ambiguity among teaching professionals. Bhuvaneshwari (2013) reasoned long working hours and improper work–life balance, while Manabete *et al.* (2016) named role ambiguity, poor relations with the boss and work overload as major stressors for university teachers.

Further, previous researchers discussed several serious repercussions of stressful lecturers. It not only creates mental health-related issues for teachers, but it also negatively influences their families, students, colleagues and society at large. Previous researchers such as Harmsen *et al.* (2018) and Ramberg *et al.* (2018) reported adverse effects of stress in the form of low motivation and morale; absenteeism and turnover; poor teaching quality; fatigue and depressed mood; lack of interest, energy and passion; anxiety; tension; and frustration, etc. The most undesirable implication of distressed teachers is its direct impact on students' learning and development, academic achievement and life success (Ronfeldt *et al.*, 2013). Hanif *et al.* (2011) reported a negative relationship between stress and teaching quality. Anxiety influences one's ability to manage the classroom effectively (Jennings and Greenberg, 2009). Professors' well-being is an indispensable forecaster of pupil's learning (Klusmann *et al.*, 2016). Similarly, Shen *et al.* (2015) reported that students' motivation level pessimistically influences the burnout level of teachers. Specifically, educators' stress is negatively related to pupils' academic grades, test scores and perceived teacher–student support (Arens and Morin, 2016).

These grim implications of teacher's stress have encouraged researchers to explore interventions to minimize their anxiety levels. They are the soul of the institutions, and there is a need to empower them psychologically (Paul *et al.*, 2020). The exploration of spiritual motivation at the workplace is often neglected (Guillen *et al.*, 2015). Gull and Doh (2004) supported the transformation of the modern workplace into a spiritual workplace. Several previous scholars, such as Ashmos and Duchon (2000) and Neal (2000), revealed that the absence of workplace spirituality is one of the prominent reasons for organizational stress, absenteeism and politics. Based on these and other similar research, academic researchers have also started to advocate workplace spirituality to remedy educational stress (see Kumar, 2018; Hasan *et al.*, 2016; Khorshid, 2015). Although these studies established a negative association between workplace spirituality and stress among teachers, most of the previous studies were mostly confined to Western settings. The present study tended to examine the linkage between six Indian dimensions of workplace spirituality (*swadharma*, *lokasangraha alias other's orientation*, sense of community, authenticity, *karma* capital and *gratitude*) and stress. This study also investigates the mediating effect of constructive deviance amid the relationship between workplace spirituality and stress among university teachers. Probably, no other study has yet investigated this proposed mediating role of constructive deviance.

Literature review

Workplace spirituality

Workplace spirituality is defined as "the recognition of an inner life is nourished by meaningful work which takes place in the context of community" (Ashmos and Duchon, 2000). Also, workplace spirituality is "a framework of organizational values evidenced in the culture that promotes employees' experience of transcendence through the work process, facilitating their sense of being connected to others in a way that provides feelings of completeness and joy" (Giacalone and Jurkiewicz, 2003). It tended harmonization of body, mind, heart and spirit at work (Kumar, 2018). Literature lacked a universally accepted

definition of workplace spirituality. In a meta-analysis, Karakas (2010) reported 70 diverse definitions of spirituality at work. Similar to the multiplicity of definitions, researchers concluded the multidimensionality of workplace spirituality (Dehler and Welsh, 2003a, b; Tepper, 2003; Kinjerski and Skrypnek, 2004; Garg *et al.*, 2019).

Mitroff and Denton (1999) elaborated three dimensions of workplace spirituality – meaning in the work–life, virtuous interpersonal relationships and harmony with organizational values. Ashmos and Duchon (2000) identified workplace spirituality as an employee's inner life, meaningful work and a sense of belonging to the community. Sheep (2006) requested another four constructs of workplace spirituality, namely meaningful work, transcendence, spirit–mind integration and experience of the real self. Kinjerski and Skrypnek (2004) discussed six stages of spiritual experience at the workplace – physical, affective, cognitive, interpersonal, spiritual and mystical. Pandey *et al.* (2009) regarded four dimensions, i.e. *swadharma*, *other's orientation*, sense of community and authenticity. These four variables were augmented with two more constructs, "*karma* capital" and "*gratitude*," by Garg (2017b). This study banked upon six Indian constructs of spirituality at the workplace-*swadharm*a, *lokasangraha (other's orientation)*, sense of community, authenticity, *karma* capital and gratitude for three possible reasons. First, since spirituality is a cultural-specific phenomenon, these six native dimensions suit Indian conditions. These six factors are derived from a thorough understanding of ancient texts, scriptures and practices. Second, modern works of literature have also started to accept and examine this Indian conceptualization of spirituality. Researchers like Saxena *et al.* (2020) and Sarkar and Garg (2020) have investigated spirituality at work using these dimensions. And lastly, a reliable and valid scale is available to measure these six dimensions. Garg (2017c) reported acceptable psychometric properties of the combined scale of Indian workplace spirituality.

Swadharma is comprised of two words: "*swa*" and "*dharma*." It highlighted the importance of work allocation according to one's talent, capabilities and interests (Pandey *et al.*, 2009). Modern researchers defined *swadharm*a in terms of meditative and meaningful work. A worker searches for the meaningfulness of life through one's job assignment and work accomplishments (Pandey *et al.*, 2009; Garg and Mehak, 2021). Even Western scholars were assured of the importance of meaningfulness. Scholars like Bailey *et al.* (2018) discussed the importance of the meaningfulness of job assignments. Meditative work is the harmonious alignment of one's body, mind and spirit. Perfect synchronization of these three faculties allows one to become one with the work (McCormick, 1994). *Lokasangraha (other's orientation)* is "coming together for the betterment and welfare of society". According to Sharma (1999), it advocated the binding of the employees to serve society. Brihadaranyaka Upanishad (verse 1.4.14) prayed for a happy, progressive, flourishing, enriching and spiritually productive life for all. Shrimad Bhagwad Gita (Chapter 12, Verse 4) stated, "*Sarva-bhuta-hite ratah*," which means "One should engage in altruistic service for the welfare of fellow human beings." Modern employees aspire to fulfill their physiological, emotional and spiritual needs through their job assignments (Chakraborty, 2016; Pradhan and Jena, 2017a, b; Garg, 2017a).

A sense of community is "a feeling that members have of belongingness, a feeling that members matter to each other as well as to the group and a shared belief that members' needs will be sought after through their commitment to being together" (McMillan and Chavis, 1986). The realization of interdependence and interconnectedness among employees leads to a sense of community (Jurkiewicz and Giacalone, 2004; Garg and Sarkar, 2020). This interconnectedness may be nurtured at the workplace through institutionalizing shared vision, mission and core organizational values (Naylor *et al.*, 1996). A real sense of community could be realized when employees consider themselves an essential part of the organization (Duchon and Plowman, 2005). Authenticity is being true to self and others. It advocated honesty, candidness, truthfulness, nonviolence, nonstealing and celibacy (Pandey *et al.*, 2009).

One of the most revered and ancient Indian scripture, *Mundaka Upanishad* (Mantra 3.1.6), stated: "*Satyamaev jayate*' which means "Only truth triumphs." These values have been highly recommended by all religions, such as Buddhism, Hinduism, Christianity and Jainism. Alignment of thoughts and actions is one of the most desirable traits for success in both personal and professional life.

Karma Capital is one of the special and unique teachings of Indian spiritualism values. The notion of karma capital is "all actions that are done have the power to ordain for their doer's joy or sorrow in future depending upon the action whether it is good or bad" (Mulla and Krishnan, 2006, p. 26). *Shrimad Bhagwad Gita,* Chapter 2, Verse 47, preached performance of duty with hundred percent devotion and sincerity, with no calculation of fear-loss or win-loss. Such selfless and expectancy less dedication to the assigned task is called "*Nishkam Karma*," which is highly desirable for perpetual inner peace and joy (Chopra *et al.*, 2018; Garg *et al.*, 2019). *Gratitude* is thankfulness or feeling of gratitude. Indian spiritual values emphasize the virtue of gratefulness (Garg, 2020). Indian values and worships roughly everything, rivers, trees, animals, air, water, sun, stones, mountains, etc. Gratitude is a divine quality that bestows spiritual enrichment and contentment (Shankar, 2013). Modern positive psychologists started the scientific exploration of gratitude and its well-being implications (Garg and Gera, 2019). Researchers elaborated two essential conditions for experiencing gratefulness – (1) realizations of good done; (2) good has been done by someone else. A feeling of gratefulness for every form of life is the basis of spiritualism in India (*Shrimad Bhagwad Gita*, Chapter 18, and Verse 2).

Constructive deviance

Workplace deviance always gained prime attention from both researchers and practitioners. Several scholars, including Bennett and Robinson (2003), Mitchell and Ambrose (2007), Robinson and Bennett (1995), Bord (1976), Spreitzer and Sonenshein (2004) and Zhang and Bartol (2010), explored types, forms, implications and remedies of deviant behaviors at the workplace. However, most studies conceptualized workplace deviance in a negative connotation, and the positive side of deviance was largely ignored. These research studies argued that any workplace deviation from organizational norms might be destructive and harmful for the organization (Mitchell and Ambrose, 2007; Robinson and Bennett, 1995). However, the 21st century's call for creativity, innovation, entrepreneurship and high performance opened a welcome gate for beneficial deviances at the workplace (Zhang and Bartol, 2010). These highly acceptable and favorable deviances were referred to as "constructive deviance" (Vadera *et al.*, 2013). These deviant behaviors are an intentional disregard for established organizational norms. However, they lead to a better organization and are a blessing in disguise for the organization (Spreitzer and Sonenshein, 2004). According to Vadera *et al.* (2013), a behavior should fulfill three conditions called constructive deviance. These conditions are (1) departure from the customary norms of the group, (2) lead to the benefit of the group and (3) follow hyper norms. Hyper norms are larger societal values like honesty, truthfulness, transparency and integrity. Thus, constructive deviant behaviors may break the organizational rule, but they conform to desirable societal norms (Garg and Saxena, 2019). The present study is based on Vadera *et al.*'s (2013) conceptualization of constructive deviance. Constructive deviance includes behaviors like "creative performance," "issue selling," "extra-role behaviors," "pro-social rule-breaking," "taking charge," "counter-role behavior," "organizational citizenship behavior," "whistleblowing" and "role innovation".

Workplace spirituality, constructive deviance and stress

Numerous studies concluded the beneficial influence of spirituality in business organizations (Quatro, 2004). Researchers have revealed the association of workplace spirituality with

several antecedents of constructive deviance, such as creativity and innovation, courage and empathy. Afsar and Badir (2017) argued that spirituality at the workplace raises the awareness level of employees, which in turn promotes novelty and creativity. Also, "It can lead individuals to experience consciousness at a deeper level, thereby enhancing their intuitive abilities to develop more purposeful and compelling ideas which can increase innovation" (Afsar and Badir, 2017, p. 97). Gall *et al.*'s (2005) explained the association between spirituality and stress using the Transactional Model of Stress and Coping theory. According to the authors, spirituality influences one's way of constructing and give meaning to people and events. Spiritual beliefs and values shape the interpretation of a stressor (Gall *et al.*, 2005). Previous studies have revealed that a spiritual person may attribute a higher purpose to a stressor (Gall, 2003a, b; Saxena *et al.*, 2020). In other words, spirituality induces reframing of the stressors. The same stressor may not have the same level of influence on all employees. A stressor's impact depends on "how an employee takes a particular stressor". Spirituality brings positivity, and a spiritual person is more likely to take a stressor on a healthy note (Saxena *et al.*, 2020). Workplace spirituality leads to a positive mental state and higher levels of happiness and contentment (Mahipalan and Sheena, 2019). Such an elevated psychological state is likely to experience a lesser level of stress.

Following the transactional model, Zou *et al.* (2020) suggested that workplace spirituality enriches employee well-being by enriching personal and social resources. These resources allow an employee to adapt constructively deviant paths to achieve organizational objectives. Garg and Saxena (2020) concluded that constructive deviant behavior reduces stress, giving a chance to vent out anger, frustration and aggression through positive and healthy means. Thus, following the Transactional Model of stress, it is proposed that workplace spirituality may counter work stress through encouraging constructive work behavior. Following the same theoretical premises, the proposed association between six Indian constructs of workplace spirituality, stress and positively deviant behavior are explored here.

Karma capital promotes two approaches to work: (1) Excellence and 100% work efficiency (*Yogah Karamsu Kosalam*, Bhagavad Gita, Chapter 2, verse 50), (2) nonattachment to results (*karmanyevadhikaraste ma phalesukadacana ma karma-phala-heturbhur ma te' sangostvakarmani*, Bhagavad Gita, Chapter 2, verse 47). In simple words, it preaches duty-orientation without calculation of profit or loss, or failure or wins, or sorrow or failure (Garg, 2017c). Such expectation-free and selfless devotion puts a brake on mindless materialistic goal orientation. Feverishness for goals leads to excessive work pressure, work alcoholism, anxiety, depression and stress (Chopra *et al.*, 2018). On the other hand, karma theory provides a balanced approach, which encourages work excellence without feverishness for desirable results. It liberates the mind from the fear of failure that harbingers innovative and courageous activities at the workplace. The Transactional Model of stress may explain employees' indulgence in different forms of constructive deviance like extra-role behavior, organizational citizenship behavior and voice. Based on the model, it is proposed that the experience of karma capital may help an employee counter job stress by inducing employees' desire for work excellence. The adherence to swadharma brings supreme devotion and commitment toward job and job assignment. This utmost dedication may induce positively deviant behavior. Based on the above-discussed theoretical explanations, the following hypothesis is proposed.

H1. Constructive deviance mediates the relationship between karma capital and stress.

Further, researchers and practitioners argued that one of the most prevalent stressors in the modern organization is unauthenticated behavior, including lying, double standards, stereotyping, hypocrite, office politics, favoritism, nepotism, discrimination, harassment and exploitation. These undesirable behaviors not only stressed-out victims, but they do create anxiety for troublers too. Several researchers, including Sheldon *et al.* (1997), Ryan *et al.* (2005)

and Wood *et al.* (2008), concluded a negative linkage between authenticity and stress. "Sense of authenticity reduced depressive-anxious feelings, physical response, and cognition-thought of helplessness" (Ito and Kodama, 2005, p. 24). Following the Transactional model of coping and stress, it is proposed that authenticity may reduce the employees' stress level by developing an open, collaborative, trustworthy, mutual-fulfilling and peaceful organizational climate (Pandey *et al.*, 2016). Such a workplace may encourage desirable deviant behaviors like organizational citizenship behavior, innovative work attitude, etc. Employees high on the authenticity scale are expected to be more faithful to themselves, their work and colleagues (Pandey *et al.*, 2019). Such employees are full of confidence, which is a crucial quality for indulging in risky, challenging and other's oriented activities like constructive deviance Gupta and Singh, is proposed.

H2. Constructive deviance mediates the relationship between authenticity and stress.

A sense of community is based on sharing, caring, helping, relative obligation and commitment to each other (Chawla and Guda, 2012). Organizational community life has been advocated by classical, human relations and neo-human relation schools. The sense of belongingness significantly influences the psychological and mental health of the employees (Pendola and Gen, 2008; Neal and Walters, 2008). Interconnectedness at the workplace creates a buffer against all-pervasive and undesirable workplace evils such as intra- and interteam rivalry, intense and relentless competition, office politics, jealousy and bad mounting. Thus, it makes one more resilient to stress (Ozbay *et al.*, 2007). Further, it provides social support and comfort to an innovative idea or action (Afsar and Badir, 2017). Keys to these highly desirable and advantageous organizational behaviors are more participation and openness for sharing ideas and knowledge. Based on these theoretical foundations, the following hypothesis is proposed.

H3. Constructive deviance mediates the relationship between a sense of community and stress.

Unfulfilled desires and aspirations are significant stressors for the employees. Moreover, modern-day employees lean to satisfy their physiological, psychological, emotional and spiritual at the workplace (Dehler and Welsh, 2003a, b; Bhunia and Das, 2012; Garg *et al.*, 2019). Other's orientation, i.e. "coming together for the welfare of the society," gratifies the employees' higher-level spiritual and psychological needs (Chakraborty, 2016; Pawar, 2009; Pandey *et al.*, 2016; Pradhan and Jena, 2017a, b). The transactional model of stress has stated that spirituality shapes one's orientation toward a stressor. Other's oriented employees seek satisfaction with their spiritual needs through contributing to broader society and the natural environment (Pandey *et al.*, 2019). Such an employee directs its energy and resources for the welfare of others than to fight stressors. All religions commonly rate service to others as the best prayer of God. Others' orientation fills one with utmost satisfaction and happiness, which quickly negates the ill effects of stressful work life. Also, people are ready to move extra miles to serve others. In the pursuit of others' well-being, an employee exhibits organizational citizenship behavior and extra-role behavior. Therefore, based on these references, the following hypothesis is proposed.

H4. Constructive deviance mediates the relationship between other's orientation and stress.

The relationship of swadharma, stress and constructive deviance could be understood with the help of the Person–organization fit paradigm. Incongruence between one's skills and capabilities and requirements of job assignment lead to job stress (Chand and Koul, 2012; Murphy, 1995; Garg and Saxena, 2019). As discussed earlier, swadharma is comprised of meaningful and meditative work. Meaningfulness of work in itself is opposite to the

mismatch of skills and job demands. One considers work meaningful only when it matches his talent or one tends to acquire new skills to accomplish the current job obligations. In both cases, it leads to better person–organization fit, which reduces employees' stress. A higher meaningful score leads to a lower stress level (Lease *et al.*, 2017). Further, modern researchers conceptualize meaningful work as employees' active orientation to an organizational role with an intense craving for the betterment of work environments, processes and contexts (Spreitzer, 1995). This dynamic orientation leads to task motivation and employee engagement (Thomas and Velthouse, 1990), leading to constructively deviant behavior. Purposeful work encourages an employee to acknowledge demanding goals and initiate favorable organizational changes to continue meaningful stay in the organization (Afsar and Badir, 2017), which may again need innovative work behaviors. Similarly, Quatro (2004) concluded a positive linkage between meaningfulness and voluntary extra-role innovative work activities. Based on these premises, the following hypothesis is proposed.

H5. Constructive deviance mediates the relationship between swadharma and stress.

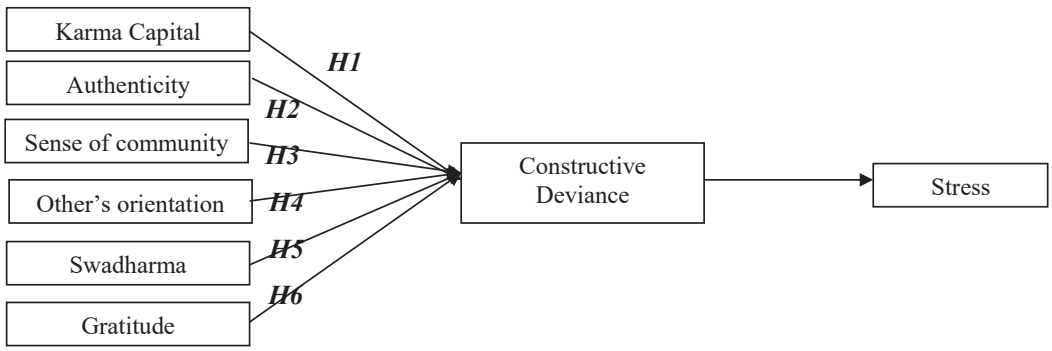
The relationship between gratitude, constructive deviance and stress could be explained in the light of Fredrickson's (2004) broaden-and-build theory of positive emotions. According to broaden-and-build theory, positive emotions such as gratitude "broaden people's momentary thought-action repertoires and build their enduring social resources" (Fredrickson, 2004, p. 147). It means that grateful people do not respond hurriedly in a tit-for-tat manner; instead, they decide their future course of action after exploring different creative options (Garg and Gera, 2019). They think consciously and comprehensively, invariably promoting creativity and novelty in response (Watkins, 2014). Also, gratefulness for the organizations encourages a grateful employee to work for the company's long-term benefit. Such employees do not hesitate to break specific organizational rules for the company's betterment, which may lead to constructive deviant behavior. The broaden-and-build theory further explains that after broadening thought-action repertoires, it builds resources for the employees. Furthermore, one of such desirable resources is good mental health, which is free from any stress, tension or anxiety. Based on this perspective, the following hypothesis is proposed. All hypotheses are represented through a diagram in Figure 1.

H6. Constructive deviance mediates the relationship between gratitude and stress.

Research methodology

Objectives, sample and procedures

The study investigates the mediating effect of constructive deviance amid the relationship between workplace spirituality and stress among Indian university teachers. The sample size



Source(s): Author's own

Figure 1.
Hypothesized model

is 300 teachers employed in various government and private educational institutes across the Delhi-NCR region. Cochran's (1977) following formula is used to calculate the sample size for the present study.

$$N = \frac{z^2 pq}{e^2}$$

Here, the confidence level is 90%, and the margin of error is 5%. Using the same formula, the present study's sample size comes out to be 294. Simple random sampling was used to administer a structured questionnaire. Respondents were approached through e-mails and hard copies. The demographic distribution of respondents is depicted in Table 1.

Measures

Workplace spirituality. Workplace spirituality is measured with the help of the 18-items combined scale of Pandey *et al.* (2009) and Garg (2017b). Garg (2018) explored the reliability and validity testing of this combined questionnaire. This scale accessed six dimensions of Indian workplace spirituality, that is, swadharma, lokasangraha (other's orientation), sense of community, authenticity, karma capital and gratitude on a five-point rating scale ranging from 1 (strongly disagree) to 5 (strongly agree). It includes items such as "my job helps me to understand my life's purpose" (swadharma), "when stuck with a problem, people here feel free to ask for" (sense of community), "people are concerned about the natural environment while working here" (lokasangraha), "I put my 100% in my work" (Karma capital) and "my job makes me understand that I have numerous things to cherish in my life" (gratitude).

Constructive deviance. The 10-item constructive deviance scale developed by Galperin (2012) was used to access constructive deviance among university teachers. It comprised of statements like "Bend or break the rules in order to perform your job" and "I violated company procedures to solve a problem." These statements were rated on a seven-point rating scale. The majority of the previous researchers discussed the theoretical conceptualization of constructive deviance. Since the present study is one of the elementary studies on the organizational utility of constructive deviance, a unidimensional scale was used.

Stress. Cohen's *et al.* (1983) 10-item perceived stress scale was utilized to measure the teacher's stress level. This scale is comprised of statements like "How often have you been upset because of something that happened unexpectedly?" and "How often have you felt that

Table 1. Demographic profile of the sample

Variable	Category	Total	Percentage
Gender	Male	171	57
	Female	129	43
Marital status	Married	201	67
	Unmarried	99	33
Age	Less than 30 years	65	21
	30 to 40 years	107	36
	40 to 50 years	95	32
	More than 50 years	33	11
Experience	0–10 years	122	40
	10–20 years	82	27
	20–30 years	69	23
	More than 30 years	27	9

Source(s): Primary data

you were unable to control the important things in your life?" These 10 statements were rated on a five-point rating scale ranging from 1 (never) to 5 (very often).

Reliability and validity

Internal consistency reliability was examined with the help of Cronbach's alpha, which is the average of all possible split-half correlations. According to George and Mallery (2003), Cronbach's alpha value of 0.70 states good reliability. Table 2 reported satisfactory values of Cronbach's alpha values (*swadharma = 0.887, other's orientation = 0.781*, sense of community = 0.810, authenticity = 0.919, *gratitude* = 0.838, karma capital = 0.737, constructive deviance = 0.712 and stress = 0.890). Average variance extracted (AVE) values were calculated to ascertain convergent validity. Table 2 observed that all AVE values were greater than 0.50, which according to Fornell and Larcker (1981) confirmed convergent validity (*swadharma = 0.624, other's orientation = 0.665*, sense of community = 0.779, authenticity = 0.743, *gratitude* = 0.782, karma capital = 0.814, constructive deviance = 0.763 and stress = 0.800).

Multicollinearity and common method variance (CMV)

Now, the variance inflation factor (VIF) values were captured to access the problem of multicollinearity. VIF values should be less than 5 to conclude the nonexistence of multicollinearity (Kline, 1998). VIF values (*swadharma = 3.12*, *other's orientation = 4.45*, sense of community = 3.30, authenticity = 3.76, *gratitude* = 2.78, karma capital = 3.67) stated that six dimensions of workplace spirituality did not have a statistically significant problem of multicollinearity. The authors used both procedural and statistical medications were used to reduce common method bias. First, prevalidated questionnaires with different rating scales were used to measure workplace spirituality, stress and constructive deviance. Second, teachers were assured of anonymity and academic usage of collected data. Moreover, lastly, Harman's single factor was used to verify the common method biases (Chang *et al.*, 2010). When eighteen statements of workplace spirituality were subjected to exploratory factor analysis, six factors were extracted. Extraction of more than one factor reported the absence of common method variance. The largest factor explained only 23% variance; thus, a single factor does not explain the most variance. It concludes the absence of the issue of common method variance.

Methods and procedure

The collected data are investigated using SPSS 21. The linkage of spirituality at the workplace and stress is examined using correlation and regression analysis. The mediating

Variable	Mean	SD	α	AVE	VIF
Karma capital	4.44	0.760	0.737	0.814	3.67
Authenticity	3.12	1.231	0.919	0.743	3.76
Sense of community	4.54	0.482	0.810	0.779	3.30
Other's orientation	3.10	0.210	0.781	0.665	4.45
Swadharma	3.79	0.873	0.887	0.624	3.12
Gratitude	3.98	0.492	0.838	0.782	2.78
Constructive deviance	5.12	1.083	0.712	0.763	3.70
Stress	3.28	0.627	0.890	0.800	–

Source(s): Primary data. S.D, standard deviation; α, Cronbach's alpha; AVE, average variance extracted; VIF, variance inflation index

Table 2. Descriptive statistics

effect of constructive deviance is examined using Sobel tests and bootstrapping estimates with 5,000 replacements using Hayes's (2013) PROCESS Macro in SPSS using model number 4. PROCESS provides better results than Baron and Kenny's (1986) mediational analysis (Hayes, 2013). It is easy to use and is preprogrammed with several mediation and moderation models. The model of the present study is taken as model number-4 in the PROCESS macro of SPSS.

Result and analysis

Table 2 reported mean, standard deviation, reliability, validity and multicollinearity estimates. These estimates confirmed internal consistency reliability, convergent validity and lack of multicollinearity. Table 3 elaborated on the correlation matrix of all variables of the present study. It was suggested that six dimensions of workplace spirituality were negatively correlated with stress. Also, workplace spirituality was positively correlated with constructive deviance. Since the correlation coefficient determined the degree and direction of the relationship between two variables, regression was used to envisage stress based on six dimensions of workplace spirituality.

Table 4 presented result of the regression model ($S = \beta_0 + \beta_1$ karma capital $+ \beta_2$ authenticity $+ \beta_3$ sense of community $+ \beta_4$ other's orientation $+ \beta_5$ swadharma $+ \beta_6$ gratitude). All regression coefficients were negative and statistically significant $\beta_{KC} = -0.382$, p-value < 0.001; $\beta_A = -0.345$, p-value < 0.001; $\beta_{SC} = -0.504$, p-value < 0.001; $\beta_L = -0.436$, p-value < 0.001; $\beta s = -0.424$, p-value < 0.001; $\beta_K = -0.360$, p-value < 0.001). Table 5 reported that these six dimensions of workplace spirituality could explain 41.9% variation in stress of teachers. And accordingly, following regression model was extracted:-

Table 3. Correlation matrix

Variable	1	2	3	4	5	6	7	8
Karma capital	1							
Authenticity	0.444*	1						
Sense of community	0.312	0.153	1					
Other's orientation	0.376*	0.258*	0.244*	1				
Swadharma	0.302*	0.333*	0.319	0.424*	1			
Gratitude	0.128	0.119	0.453*	0.152	0.471*	1		
Constructive deviance	0.515*	0.412*	0.372*	0.465*	0.381*	0.389*	1	
Stress	−0.606*	−0.528*	−0.478*	−0.579*	−0.646*	−0.610*	−0.232*	1

Source(s): Primary data. *Correlation is significant at the 0.01 level

Table 4. Multiple regression analysis (dependent variable – stress)

Variable	Unstandardized coefficients		Standardized coefficients		
	B	Std. Error	*B*	*t*-value	Sig
Constant	1.028	0.472		4.839	0.000*
Karma capital	−0.413	0.318	−0.382	3.214	0.023*
Authenticity	−0.387	0.082	−0.345	4.427	0.004*
Sense of community	−0.518	0.193	−0.504	1.897	0.047*
Other's orientation	−0.452	0.097	−0.436	6.637	0.000*
Swadharma	−0.440	0.118	−0.424	7.849	0.000*
Gratitude	−0.384	0.382	−0.360	2.764	0.034*

Source(s): Primary data. *Correlation is significant at the 0.01 level

$$S = \beta_0 + 0.382\,\text{Karma Capital} - 0.345\,\text{Authenticity} - 0.504\,\text{Sense of Community} - 0.436\,\text{Other's orientation} - 0.424\,\text{Swadharma} - 0.360\,\text{Gratitude}$$

Table 6 describes the Sobel test and bootstrapping results, which were used to test the mediation effect of constructive deviance. Zero did not lie in the confidence interval (between LLCI and ULCI) in all six mediation models. It states that the mediating effect of constructive deviance is statistically significant. Also, the Sobel test substantiated the statistical significance of the mediation effect of constructive deviance. These results provide support for hypotheses 1–6. All hypotheses are accepted here. Table 7 explained the direct and indirect impact of six dimensions of workplace spirituality on university teachers' work stress. The first column shows the direct effect, and the indirect effect is depicted in the fourth column.

Discussion

This research study examined the relationship between workplace spirituality, stress and constructive deviance. More precisely, the study explored the correlation and regression between six Indian dimensions of workplace spirituality and stress among university teachers. Also, it examined the mediating effect of constructive deviance amid the linkage

Model	*R*	*R* square	Adjusted *R* square	Std. Error of estimate	*F*	Sig
1	0.648	0.419	0.380	0.693	11.839	0.000*

Source(s): Primary data. **Correlation is significant at the 0.01 level

Table 5. Model summary

Proposed mediation model	Sobel test *z*-value	Sobel test Sig	Bootstrapping estimates Effect	SE	LLCI	ULCI	Hypothesis support
KC → CD → S	5.837	0.000*	0.163	0.037	0.173	0.494	H1
A → CD → S	5.483	0.000*	0.194	0.074	0.035	0.103	H2
SC → CD → S	4.736	0.009*	0.293	0.063	0.346	0.685	H3
L → CD → S	3.288	0.017*	0.102	0.028	0.034	0.199	H4
K → CD → S	2.732	0.029*	0.248	0.011	0.023	0.287	H5
S → CD → S	6.636	0.000*	0.171	0.23	0.044	0.058	H6

Source(s): Primary data. **$p < 0.05$, *$p < 0.01$*, KC, karma capital; A, authenticity; SC, sense of community; L, other's orientation; K, gratitude; S, swadharma

Table 6. Mediation analysis (results of Sobel test and bootstrapping)

Proposed mediation model	IV to DV	IV to M	M to DV	IV to DV (with a mediator)
KC → CD → S	−0.375*	0.504*	−0.227*	−0.249*
A → CD → S	−0.330*	0.396*	−0.227*	−0.217*
SC → CD → S	−0.479**	0.357*	−0.227*	−0.276*
L → CD → S	−0.428*	0.440*	−0.227*	−0.289*
K → CD → S	−0.416**	0.366*	−0.227*	−0.131*
S → CD → S	−0.345*	0.372*	−0.227*	−0.187*

Source: Primary data. **$p < 0.05$, *$p < 0.01$*, IV, independent variable; DV, dependent variable; M, mediator, KC, karma capital; A, authenticity; SC, sense of community; L, other's orientation; K, gratitude; S, swadharma

Table 7. Mediation analysis (results of hierarchical regression)

between spirituality and stress. Findings reveal a significant negative association between the six dimensions of workplace spirituality and stress. Thus, these results support that of previous studies conducted by Karakas (2010), Krishnakumar and Robinson (2015), Kumar and Kumar (2014) and Houghton *et al.* (2016). The results also concluded mediation effect of constructive deviance, which means workplace spirituality, promotes constructive deviance that influences the stress level of educators. All six hypotheses of the study are accepted. It signifies that constructive deviance mediates the linkage of stress with six dimensions of workplace spirituality.

These findings could be understood in the background of both the ancient Indian spiritual paradigm and contemporary management thoughts. The modern text defines swadharma in terms of meaning and meditative work. Accordingly, the swadharma of a teacher is teaching, guiding, counseling and enlightening students. Furthermore, a teacher is always ready to perform its "dharma," even if it deviates from routine processes and norms. Contemporary literature on constructive deviance also supports s*wadharma.* Pro-social rule-breaking behavior is "intentional violation of a formal organizational policy, regulation, or prohibition with the primary intention of promoting the organization's welfare or one of its stakeholders" (Morrison, 2006, p. 6). Similarly, proactive behavior is "taking the initiative in improving current circumstances or creating new ones; it involves challenging the status quo rather than passively adapting to present conditions" (Crant, 2000, p. 436). The underlying common dimension between these forms of constructive deviance is the well-being of students or educational institutes. These deviant behaviors help educators perform their duty better, thereby giving immense satisfaction, contentment and happiness to faculty members.

Results also favor the association between gratitude and stress through constructive deviance. Grateful teachers feel morally obliged to return the blessings of nature. Moreover, the best way to reciprocate for an educator is to make the extra effort for transmitting appropriate knowledge to students, even after breaking norms. Organizational citizenship behaviors depicted by academicians include providing individual attention to weak students (it is against the principle of uniform treatment to all), offering additional time to pupils even after class (not a part of duty), taking a class of colleagues in their absence (to keep students engage and to offer additional knowledge) and offering/arranging financial help to support economically weaker students (Somech and Oplatka, 2014; Somech and Ron, 2007). These behaviors look so natural to the teaching profession that nobody calls them positively deviant behaviors. These routine acts of constructive deviance are poised to offer intrinsic satisfaction to the teacher, reducing their stress level. A teacher high on gratitude may feel relieved by providing additional help to students and colleagues.

Authenticity is highly desirable for teachers. Virtues and morals of teachers are highly contagious, as students generally consider their educators as role models. They tend to imitate the ethical behaviors of their teachers. Thus, teachers' authenticity does not only lead to constructive deviant behavior or reduces stress, but it motivates pupils to acquire desirable virtues of truth, honesty, transparency, nonstealing and accountability. An honest and transparent teacher is courageous enough to indulge in positively deviant behavior for the betterment of its student and institutions. It may enrich a teacher with immense contentment of doing goods for others. This contentment indeed reduces their stress and anxiety levels.

Theoretical contributions of the study

The present study offers significant contributions to the literature on workplace spirituality, constructive deviance and stress. First, Indian spiritual values and practices are often criticized for the scarcity of empirical explorations. Although India houses one of the world's largest and oldest spiritual traditions in the form of epics, yoga, pranayama, mediations, etc., they are not readily accepted worldwide due to a lack of scientific pieces of evidence. The studies like the present one are poised to popularize eastern spiritual knowledge in the entire

world. The study also supports Indian dimensions of workplace spirituality. A few dimensions like swadharama, karma capital and lokasangraha are yet to be accepted universally. The present study recommends these Indian dimensions as essential constituents of workplace spirituality. The inclusion of these factors in modern research will make workplace spirituality holistic and comprehensive.

Second, this study contributes to the spirituality literature, which is frequently criticized for the scarcity of research. The shortage of empirical studies restricts the extension, exploration, development and application of the concept of spirituality at the workplace (Gatling *et al.*, 2016; Pawar, 2014). Sheep (2006) also highlighted that spirituality's "inclusion in established research streams is long overdue" (p. 371). By examining the association between six Indian dimensions of spirituality and stress, the present study expands the empirical understanding of the organizational utility of spirituality. It highlights the growing acceptability of spirituality at the workplace. It will also promote studies in specific dimensions of workplace spirituality like gratitude or swadharma. Third, this is one of the pioneer studies that explored the linkage between specific dimensions of workplace spirituality and stress. Six dimensions of workplace spirituality are derived from previous empirical and theoretical studies and are based on traditional wisdom and knowledge of great Indian epics and scriptures.

Fourth, the findings of the study cross-validated the claim of previous researchers. Daniel (2015) suggested that workplace spirituality reduces stress among respondents of the USA and Mexico. Other researchers like Eklund (2009), Kumar (2018), Hasan *et al.* (2016), Khorshid (2015), Piryaei and Zare (2013) and Tevichapong (2012) also reported a negative relationship between spirituality at work and stress. But, the majority of these studies were conducted in Western settings. This study is based on the Indian conceptualization of workplace spirituality. The present study validates the findings of previous studies in eastern contexts.

And, the study also unlocked the "black-box" (Gardner *et al.*, 2001) of the mechanism through which workplace spirituality influences stress. The study revealed that workplace spirituality promotes constructive deviant behaviors, which reduces stress among teachers. This is one of the first studies that examined the mediating effect of constructive deviance. Thus, the finding enriches the literature of constructive deviance too. Also, it highlights the utility of promoting constructive deviance at the workplace. The studies like the present one may elicit more acceptability to positively deviant behavior in modern organizations.

The study contributes to the literature of positive organizational scholarship in two possible ways. By exploring constructive deviance, the study strengthens our understanding of organizational psychology. The majority of previous studies have discussed workplace deviance in a negative connotation. The present study may promote further exploration of antecedents and impacts of constructively deviant behavior. And, stress is one of the most explored topics of personal psychology. Previous scholars have recommended various interventions to deal with stress. Based on the Indian conceptualization of spirituality, the study offers another line of direction to psychologists and researchers. Positive psychologists may derive more practical interventions from traditional Indian wisdom. A blend of ancient Indian knowledge and modern scientific understanding of psychology can greatly benefit positive organizational revolution.

Practical implications of the study

Several practical implications could be derived from the findings of the present study. The findings of this study encourage academic leaders to realize the importance of institutionalizing workplace spirituality for effective management of the lecturer's stress. The most meaningful and meditative work for teachers is to assist students. Thus, it seems like nonacademic work, role ambiguity and role overload create stress and anxiety among

teachers. Academic managers are suggested to inspire a work environment where teachers could find meaning in these activities (Caudron, 1997; Guinot *et al.*, 2014). Such a dreamed educational institute develops with time by ensuring teachers' decision-making participation and two-way open communication between teachers and administrators. Academic administrators of various educational institutes may use specific spiritual interventions to make teachers realize the importance of workplace spirituality. Open discussion on karma capital, gratitude or swadharma could facilitate a better understanding of spiritual values and their benefits. It could be complemented with tailor-made seminars and workshops on breathing techniques and *asanas*. Guided meditation helps in the internalization of the real meaning of spirituality. A sense of community could easily be facilitated by organizing formal functions like social gatherings, award distribution ceremonies and informal events like picnics and movies. Values of other's orientation (collective welfare of others) could be cherished through encouraging teachers, students and other stakeholders to engage in any community-building practices. Charity programs, community services, visits to hospitals, orphanages, old age homes and other similar events may help teachers and students appreciate the notion of other's orientation.

Limitations and scope for future studies

This study has several limitations. First, the generalizability of the study remains a concern. This study is based on a sample of 300 teachers residing in the Delhi-NCR region only. Future researchers may validate the results of this study using pan-India data. The stress level of other working professionals was studied in the light of workplace spirituality and constructive deviance. Second, data related to all three variables were collected from a single point simultaneously. It may lead to reverse causality and common method variance. Although appropriate remedies were proposed, future researchers should use double source data, longitudinal study or qualitative research design. Third, six dimensions of workplace spirituality reported significant intercorrelations among themselves, which is acceptable given the nature of constructs and study. Moreover, VIF values negated the problem of multicollinearity of these variables. Fourth, this study is based on a one-dimensional scale of constructive deviance and stress. Future researchers could examine specific constructive deviant behaviors such as pro-social behavior, organizational citizenship behavior, whistle-blower behavior and issue selling.

Conclusion

The study theoretically proposed and empirically examined the association between workplace spirituality and stress among Indian university teachers. It also explored the mediating effect of constructive deviance amid the relationship between workplace spirituality and stress. The findings supported the initial proposition that six dimensions of workplace spirituality, i.e. karma capital, authenticity, sense of community, lokasangraha, gratitude and swadharama, could reduce teachers' stress. The study also offered an insight into the mechanism through which spirituality reduces stress. The study's findings provided appropriate suggestions to policymakers to combat the menace of stress among the teaching fraternity.

References

Afsar, B. and Badir, Y. (2017), "Workplace spirituality, perceived organizational support and innovative work behavior", *Journal of Workplace Learning*, Vol. 29 No. 2, pp. 95-109, doi: 10.1108/jwl-11-2015-0086.

Aggarwal, R. (2012), "Self-efficacy as predictor of occupational stress among", *Academic Faculties of Punjab University and Guru Nanak Dev UniversityIndian Journal of Psychological Science*, Vol. 3 No. 1, pp. 49-61.

Arens, A.K. and Morin, A.J.S. (2016), "Relations between teachers' emotional exhaustion and students' educational outcomes", *Journal of Educational Psychology*, Vol. 108 No. 6, pp. 800-813.

Ashmos, D.P. and Duchon, D. (2000), "Spirituality at work: a conceptualization and measure", *Journal of Management Inquiry*, Vol. 9 No. 2, pp. 134-145.

Bailey, C., Yeoman, R., Madden, A., Thompson, M. and Kerridge, G. (2018), "A review of the empirical literature on meaningful work: progress and research agenda", *Human Resource Development Review*, Vol. 18 No. 1, pp. 83-113, doi: 10.1177/1534484318804653.

Baron, R.M. and Kenny, D.A. (1986), "The moderator–mediator variable distinction in social psychological research: conceptual, strategic and statistical considerations", *Journal of Personality and Social Psychology*, Vol. 51 No. 6, pp. 1173-1182.

Bennett, R.J. and Robinson, S.L. (2003), "The past, present, and future of workplace deviance research", in Greenberg, J. (Ed.), *Organizational Behavior: The State of the Science*, 2nd ed., Erlbaum, Mahwah, NJ.

Bhunia, A. and Das, A. (2012), "Explore the impact of workplace spirituality on motivations for earnings management: an empirical analysis", *International Journal of Scientific and Research Publications*, Vol. 2 No. 2, pp. 1-5.

Bhuvaneshwari, M. (2013), "A case study on psychological and physical stress undergone by married working women", *Journal of Business and Management*, Vol. 14 No. 6, pp. 38-44.

Bord, R.J. (1976), "The impact of imputed deviant identities in structuring evaluations and reactions", *Sociometry*, Vol. 39, pp. 108-115.

Caudron, S. (1997), "The search for meaning at work", *Training and Development*, Vol. 51 No. 9, pp. 24-27.

Chakraborty, S.K. (2016), *Human Values and Ethics: Achieving Holistic Excellence*, Icfai University Press, Hyderabad.

Chand, P. and Koul, H. (2012), "*Workplace spirituality, organizational emotional ownership and job satisfaction as moderators* in *coping with job stress", Paper Presented at the International Conference on Humanities, Economics and Geography, Bankok*, 17–18 March.

Chang, S.J., van Witteloostuijn and Eden, L. (2010), "From the editors: common method variance in international business research", *Journal of International Studies*, Vol. 41 No. 1, pp. 178-184.

Chaplain, R.P. (2008), "Stress and psychological distress among trainee secondary teachers in England", *Educational Psychology*, Vol. 28 No. 2, pp. 195-209.

Chopra, K.N., Sharma, N. and Behura, A. (2018), "Optimization of financial resources for growth of spiritual learning (nishkam karma yoga—desire less action) for stress management in business corporate sector", *International Journal of Accounting and Finance Studies*, Vol. 1 No. 2, p. 142, doi: 10.22158/ijafs.v1n2p142.

Chawla, V. and Guda, S. (2012), "Workplace spirituality as a precursor to Relationship-Oriented selling characteristics", *Journal of Business Ethics*, Vol. 115 No. 1, pp. 63-73, doi: 10.1007/s10551-012-1370-y.

Cochran, W.G. (1977), *Sampling Techniques*, 3rd ed., John Wiley & Sons, New York.

Cohen, S., Kamarck, T. and Mermelstein, R. (1983), "A global measure of perceived stress", *Journal of Health and Social Behavior*, Vol. 24 No. 4, pp. 386-396.

Crant, J.M. (2000), "Proactive behavior in organizations", *Journal of Management*, Vol. 26 No. 1, pp. 435-462.

Daniel, J.L. (2015), "Workplace spirituality and stress: evidence from Mexico and US", *Management Research Review*, Vol. 38 No. 1, pp. 29-43, doi: 10.1108/mrr-07-2013-0169.

Dehler, G.E. and Welsh, M.A. (2003a), "The experience of work: spirituality and the new work-place", in Giacalone, R.A. and Jurkiewicz, C.L. (Eds), *Handbook of Workplace Spirituality and Organizational Performance*, M.E. Sharpe, Armonk, NY, pp. 108-122.

Dehler, G.E. and Welsh, M.A. (2003b), "The handbook of workplace spirituality and organizational performance", in Giacalone, R.A. and Jurkiewics, C.L. (Eds), NY7 M.E. Sharpe, pp. 2108-2122.

Duchon, D. and Plowman, D.A. (2005), "Nurturing the spirit at work: impact on work unit performance", *The Leadership Quarterly*, Vol. 16, pp. 807-833.

Eklund, N. (2009), "Sustainable workplaces, retainable teachers", *Phi Delta Kappan*, Vol. 91 No. 2, pp. 25-27.

Fornell, C. 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), "Gratitude, like other positive emotions, broadens and builds", in Emmons, R.A. and McCullough, M.E. (Eds), *The Psychology of Gratitude*, Oxford University Press, pp. 145-166.

French, J.R.P. Jr, Caplan, R.D. and Harrison, R.V. (1982), *The Mechanisms of Job Stress and Strain*, Wiley, London.

Gall, T.L. (2003a), "Religious and spiritual attributions in older adults' adjustment to illness", *Journal of Psychology and Christianity*, Vol. 22 No. 3, pp. 210-222.

Gall, T.L. (2003b), "The role of religious resources for older adults coping with illness", *The Journal of Pastoral Care and Counseling*, Vol. 57 No. 2, pp. 211-224.

Gall, T.L., Charbonneau, C., Clarke, N.H., Grant, K., Joseph, A. and Shouldice, L. (2005), "Understanding the nature and role of spirituality in relation to coping and health: a conceptual framework", *Canadian Psychology/Psychologie Canadienne*, Vol. 46 No. 2, pp. 88-104.

Galperin, B.L. (2012), "Exploring the nomological network of workplace deviance: developing and validating a measure of constructive deviance", *Journal of Applied Social Psychology*, Vol. 42 No. 12, pp. 2988-3025.

Gardner, T., Moynihan, L., Park, H.J. and Wright, P. (2001), "Beginning to unlock the black box in the HR firm performance relationship: the impact of HR practices on employee attitudes and employee outcomes", *Cornell University, School of Industrial and Labor Relations, Center for Advanced Human Resource Studies*, Vol. 1 No. 12, Ithaca, NY, pp. 1-45.

Garg, N. (2017a), "Workplace spirituality and employee well-being: an empirical exploration", *Journal of Human Values*, Vol. 23 No. 2, pp. 1-19.

Garg (2017b), "Workplace spirituality and organizational commitment: an empirical exploration", *Purushartha- A Journal of Management, Ethics and Spirituality*, Vol. 10, pp. 8-24.

Garg, N. (2017c), "Workplace spirituality and organizational performance in Indian context: mediating effect of organizational commitment, work motivation and employee engagement", *South Asian Journal of Human Resources Management*, Vol. 4, pp. 191-211.

Garg, N. (2018), "Promoting organizational performance in Indian insurance industry: the roles of workplace spirituality and organizational citizenship behaviour", *Global Business Review*, Vol. 21 No. 2, pp. 1-16.

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, doi: 10.1108/IJEM-07-2019-0253.

Garg, N. and Gera, S. (2019), "Gratitude and leadership in higher education institutions: exploring the mediating role of social intelligence among teachers", *Journal of Applied Research in Higher Education*, Vol. 12 No. 5, pp. 915-926, doi: 10.1108/JARHE-09-2019-0241.

Garg, N. and Mehak (2021), "Gratitude resentment and appreciation scale (GRAT-16): analyzing psychometrics properties in the Indian context", *Current Psychoogy*. doi: 10.1007/s12144-021-02169-6.

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, doi: 10.1108/JOEPP-03-2020-0041.

Garg, N. and Saxena, A. (2019), "Promoting constructive deviance as an antidote to organizational stress", *Analyzing Workplace Deviance in Modern Organizations*, IGI Global, Ahmedabad, Gujarat, pp. 139-158.

Garg, N. and Saxena, A. (2020), "Analyzing the interrelation between workplace spirituality and Constructive Deviance", *Asian Journal of Business Ethics*, Vol. 9 No. 1, pp. 121-141, 5.

Garg, N., Punia, B.K. and Jain, A. (2019), "Workplace spirituality and job satisfaction: exploring mediating effect of organization citizenship behaviour", *Vision*, Vol. 23 No. 3, pp. 287-296.

Gatling, A., Kim, J.S. and Milliman, J. (2016), "The relationship between workplace spirituality and hospitality supervisors' work attitudes", *International Journal of Contemporary Hospitality Management*, Vol. 28 No. 3, pp. 471-489, doi: 10.1108/ijchm-08-2014-0404.

George, D. and Mallery, P. (2003), *SPSS for Windows Step by Step: A Simple Guide and Reference. 11.0 Update*, 4th ed., Allyn & Bacon Publication, Boston, Massachusetts.

Ghania, M., Charr, P. and Verma, A. (2014), "Stress among Indian teachers", *Journal of Academics*, Vol. 34 No. 1, pp. 12-21.

Giacalone, R.A. and LJurkiewicz, C. (2003), "Toward a science of workplace spirituality", in Giacalone, R.A. and Jurkiewicz, C.L. (Eds), *The Handbook of Workplace Spirituality and Organizational Performance*, M.E. Sharpe, Armonk, NY.

Guillen, M., Ferrero, I. and Hoffman, W.M. (2015), "The neglected ethical and spiritual motivations in the workplace", *Journal of Business Ethics*, Vol. 128 No. 4, pp. 803-816.

Guinot, J., Chiva, R. and Roca-Puig, V. (2014), "Interpersonal trust, stress and satisfaction at work: an empirical study", *Personnel Review*, Vol. 43 No. 1, pp. 96-115.

Gull, G. and Doh, J. (2004), "The 'transmutation' of the organization: toward a more spiritual workplace", *Journal of Management Inquiry*, Vol. 13 No. 2, pp. 128-139.

Harmsen, R., Helms-Lorenz, M., Maulana, R. and Veen, K. (2018), "The relationship between beginning teachers' stress causes, stress responses, teaching behaviour and attrition", *Teachers and Teaching*, Vol. 24 No. 6, pp. 626-643.

Hassan, M., bin Nadeem, A. and Akhter, A. (2016), "Impact of workplace spirituality on job satisfaction: mediating effect of trust", *Cogent Business and Management*, Vol. 3 No. 1, 1189808, doi: 10.1080/23311975.2016.1189808.

Hayes, A.F. (2013), *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*, The Guilford Press, New York, NY.

Houghton, J.D., Neck, C.P. and Krishnakumar, S. (2016), "The what, why, and how of spirituality in the workplace revisited: a 14-year update and extension", *Journal of Management, Spirituality and Religion*, Vol. 13 No. 3, pp. 177-205.

Hanif, R., Tariq, S. and Nadeem, M. (2011), "Personal and job related predictors of teacher stress and job performance among school teachers", *Pakistan Journal of Commerce and Social Sciences*, Vol. 5 No. 2, pp. 319-329.

Hooftman, W.E., Mars, G.M.J., Janssen, B., de Vroome, E.M.M. and van den Bossche, S.N.J. (2015), *Nationale Enquête Arbeidsomstandigheden 2014. Methodologie en globale resultaten [National Survey Working Conditions 2014. Methodology and Global Results]*, (Research Report No. 525236), Retrieved from TU Delft Online Library, available at: http://publications.tno.nl/publication/34616775/NPtzyl/hooftman-2015-nationale.pdf.

Ito, M. and Kodama, M. (2005), "The relationships of sense of authenticity with stress-responses and stress-copings", *Japanese Journal of Health Psychology*, Vol. 18, pp. 24-34.

Jennings, P.A. and Greenberg, M.T. (2009), "The prosocial classroom: teacher social and emotional competence in relation to student and classroom outcomes", *Review of Educational Research*, Vol. 79 No. 1, pp. 491-525.

Jurkiewicz, C.L. and Giacalone, R.A. (2004), "A values framework for measuring the impact of workplace spirituality on organizational performance", *Journal of Business Ethics*, Vol. 49 No. 2, pp. 129-142.

Karakas, F. (2010), "Spirituality and performance in organizations: a literature review", *Journal of Business Ethics*, Vol. 94 No. 1, pp. 89-106.

Khorshid, S. (2015), "The effects of workplace spirituality and teacher's self-efficacy beliefs on his/her job satisfaction", *International Journal of Psychology*, Vol. 9 No. 1, pp. 64-103.

Kinjerski, V.M. and Skrypnek, B.J. (2004), "Defining spirit at work: finding common ground", *Journal of Organizational Change Management*, Vol. 17 No. 1, pp. 26-42.

Kline, R.B. (1998), *Principles and Practice of Structural Equation Modeling*, The Guilford Press, New York, NY.

Klusmann, U., Richter, D. and Lüdtke, O. (2016), "Teachers' emotional exhaustion is negatively related to students' achievement: evidence from a large-scale assessment study", *Journal of Educational Psychology*, Vol. 108 No. 8, pp. 1193-1203.

Krishnakumar, S. and Robinson, M.D. (2015), "Maintaining an even keel: an affect-mediated model of mindfulness and hostile work behavior", *Emotion*, Vol. 15, pp. 579-589.

Kumar, S. (2018), "A study of perceived workplace spirituality of school teachers", *Psychological Thought*, Vol. 11 No. 2, pp. 212-223, doi: 10.5964/psyct.v11i2.298.

Kumar, V. and Kumar, S. (2014), "Workplace spirituality as a moderator in relation between stress and health: an exploratory empirical assessment", *International Review of Psychiatry*, Vol. 26 No. 3, pp. 344-351.

Lease, S.H., Ingram, C.L. and Brown, E.L. (2017), "Stress and health outcomes: do meaningful work and physical activity help?", *Journal of Career Development*, Vol. 46 No. 3, pp. 251-264, doi: 10.1177/0894845317741370.

Manabete, D.S.S., John, C.A., Makinde, A.A. and Duwa, S.T. (2016), "Job stress among school administrators' and teachers in Nigerian secondary schools and technical colleges", *International Journal of Education, Learning and Development*, Vol. 4 No. 2, pp. 9-24.

Mahipalan, M. and Sheena, S. (2019), "Workplace spirituality, psychological well-being and mediating role of subjective stress", *International Journal of Ethics and Systems*, Vol. 35 No. 4, pp. 725-739, doi: 10.1108/ijoes-10-2018-0144.

McCormick, D. (1994), "Spirituality and management", *Journal of Managerial Psychology*, Vol. 9 No. 6, pp. 5-8.

McMillan, D.W. and Chavis, D. (1986), "Sense of community: a definition and theory", *Journal of Community Psychology*, Vol. 14 No. 1, pp. 6-23.

Mitchell, M.S. and Ambrose, M.L. (2007), "Abusive supervision and workplace deviance and the moderating effects of negative reciprocity beliefs", *Journal of Applied Psychology*, Vol. 92 No. 4, pp. 1159-1168, doi: 10.1037/0021-9010.92.4.1159.

Mitroff, I.I. and Denton, A.E. (1999), "A study of spirituality in the workplace", *Sloan Management Review*, Vol. 40 No. 4, pp. 83-93.

Morrison, E.W. (2006), "Doing the job well: an investigation of pro-social rule breaking", *Journal of Management*, Vol. 32 No. 1, pp. 5-28.

Mulla, Z.R. and Krishnan, V.R. (2006), "Karma yoga: a conceptualization and validation of the indian philosophy of work", *Journal of Lndian Psychology*, Vol. 24 Nos 1 & 2, pp. 26-43.

Murphy, L.R. (1995), "Occupational stress management: current status and future directions", in Cooper, C.L. and Rousseau, D.M. (Eds), *Trends in Organizational Behavior*, John Wiley & Sons, Vol. 2, pp. 1-14.

Naylor, T., Willimon, W. and Osterberg, R. (1996), "The search for community in the workplace", *Business and Society Review*, Vol. 97 No. 1, pp. 42-47.

Neal, J. (2000), "Andre delbecq: spiritual friend and mentor", *Journal of Management, Spirituality and Religion*, Vol. 17 No. 1, pp. 22-36.

Neal, S. and Walters, S. (2008), "Rural be/longing and rural social organizations: conviviality and Community-Making in the English countryside", *Sociology*, Vol. 42 No. 2, pp. 279-297, doi: 10.1177/0038038507087354.

Newberry, M. and Allsop, Y. (2017), "Teacher attrition in the USA: the relational elements in a Utah case study", *Teachers and Teaching*, Vol. 23 No. 5, pp. 1-18.

Ozbay, F., Johnson, D.C., Dimoulas, E., Morgan, C.A., Charney, D. and Southwick, S. (2007), "Social support and resilience to stress: from neurobiology to clinical practice", *Psychiatry (Edgmont (Pa. : Township))*, Vol. 4 No. 5, pp. 35-40.

Pandey, A., Gupta, R.K. and Arora, A.P. (2009), "Spiritual climate of business organizations and its impact on customers' experience", *Journal of Business Ethics*, Vol. 88 No. 2, pp. 313-332.

Pandey, A., Gupta, R.K. and Kumar, P. (2016), "Spiritual climate and its impact on learning in teams in business organizations", *Global Business Review*, Vol. 17 No. 3, p. 1S-14S.

Pandey, A., Gupta, V. and Gupta, R. K. (2019), "Spirituality and innovative behaviour in teams: examining the mediating role of team learning", *IIMB Management Review*, Vol. 31 No. 2, pp. 116-126, doi: 10.1016/j.iimb.2019.03.013.

Parry, W., Kumar, S. and Awasthi, P. (2016), "Stress among teachers: a theoretical examination", *International Journal of Indian Psychology*, Vol. 3 No. 4, pp. 88-97.

Paul, M., Jena, L. and Sahoo, K. (2020), "Workplace spirituality and workforce agility: a psychological exploration among teaching professionals", *Journal of Religion and Health*, Vol. 59 No. 1, pp. 135-153.

Pawar, B.S. (2009), "Workplace spirituality facilitation: a comprehensive model", *Journal of Business Ethics*, Vol. 90 No. 3, pp. 375-386, doi: 10.1007/s10551-009-0047-7.

Pawar, B.S. (2014), "Leadership spiritual behaviors toward subordinates: an empirical examination of the effects of a leader's individual spirituality and organizational spirituality", *Journal of Business Ethics*, Vol. 122 No. 3, pp. 439-452, doi: 10.1007/s10551-013-1772-5.

Pendola, R. and Gen, S. (2008), "Does 'Main Street' promote sense of community?: a comparison of San Francisco neighborhoods", *Environment and Behavior*, Vol. 40 No. 4, pp. 545-574.

Piryaei, S. and Zare, R. (2013), "Workplace spirituality and positive work attitudes: the moderating role of individual spirituality", *Indian Journal of Economics and Development*, Vol. 1 No. 4, pp. 91-97.

Pradhan, R.K. and Jena, L.K. (2017a), "Workplace spirituality in Indian organizations: construction of reliable and valid measurement scale", *Business: Theory and Practice*, Vol. 18 No. 1, pp. 43-53.

Pradhan, S. and Jena, L.K. (2017b), "Effect of abusive supervision on employee's intention to quit and the neutralizing role of meaningful work in indian IT organizations", *International Journal of Organizational Analysis*, Vol. 25 No. 5, pp. 825-838, doi: 10.1108/ijoa-01-2017-1107.

Quatro, A. (2004), "New age or old age: classical management theorey and traditional organized religion as underpinnings of the contemporary organizational spirituality movement", *Human Resource Development Review*, Vol. 3 No. 3, pp. 228-249.

Ramberg, J., Låftman, S., Fransson, E. and Modin, B. (2018), "School effectiveness and truancy: a multilevel study of upper secondary schools in Stockholm", *International Journal of Adolescence and Youth*, Vol. 24 No. 2, pp. 185-198.

Ravichandran, R. and Rajendran, R. (2007), "Perceived sources of stress among the teachers", *Journal of the Indian Academy of Applied Psychology*, Vol. 33 No. 1, pp. 133-136.

Robinson, S.L. and Bennett, R.J. (1995), "A typology of deviant workplace behaviors: a multidimensional scaling study", *Academy of Management Journal*, Vol. 38 No. 2, pp. 555-572, doi: 10.2307/256693.

Ronfeldt, M., Loeb, S. and Wyckoff, J. (2013), "How teacher turnover harms student achievement", *American Educational Research Journal*, Vol. 50 No. 1, pp. 4-36.

Ryan, A.M., Patrick, H. and Shim, S.O. (2005), "Differential profiles of students identified by their teacher as having avoidant, appropriate, or dependent Help-Seeking tendencies in the Classroom", *Journal of Educational Psychology*, Vol. 97 No. 2, pp. 275-285, doi: 10.1037/0022-0663.97.2.275.

Sarkar, A. and 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*, Vol. 31 No. 5, pp. 709-728.

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.

Shankar, R.S.S. (2013), *Celebrating Silence*, Sri Sri Publication Trust, Banglore.

Sharma, S. (1999), "Corporate Gita: lessons for management, administration and leadership", *Journal of Human Values*, Vol. 5 No. 2, pp. 51-67.

Sheep, M.L. (2006), "Nurturing the whole person: the ethics of workplace spirituality in a society of organizations", *Journal of Business Ethics*, Vol. 664 No. 4, pp. 357-369.

Sheldon, K.M., Ryan, R.M., Rawsthorne, L.J. and Ilardi, B. (1997), "Trait self and true self: cross-role variation in the Big-Five personality traits and its relations with psychological authenticity and subjective well-being", *Journal of Personality and Social Psychology*, Vol. 73 No. 6, pp. 1380-1393, doi: 10.1037/0022-3514.73.6.1380.

Shen, B., McCaughtry, N., Martin, J., Garn, A., Kulik, N. and Fahlman, M. (2015), "The relationship between teacher burnout and student motivation", *British Journal of Educational Psychology*, Vol. 85 No. 4, pp. 519-532.

Skaalvik, E. and Skaalvik, S. (2015), "Job satisfaction, stress and coping strategies in the teaching profession—what do teachers say?", *International Education Studies*, Vol. 8 No. 3, pp. 181-192.

Somech, A. and Oplatka, I. (2014), *Organizational Citizenship Behavior in Schools*, 1st ed., Taylor & Francis, London, doi: 10.4324/9781315866956.

Somech, A. and Ron, I. (2007), "Promoting organizational citizenship behavior in schools: the impact of individual and organizational characteristics", *Educational Administration Quarterly*, Vol. 43 No. 1, pp. 38-66, doi: 10.1177/0013161x06291254.

Spreitzer, G.M. (1995), "Psychological, empowerment IN the workplace: dimensions, measurement and validation", *Academy of Management Journal*, Vol. 38 No. 5, pp. 1442-1465, doi: 10.2307/256865.

Spreitzer, G.M. and Sonenshein, S. (2004), "Toward the construct definition of positive deviance", *American Behavioral Scientist*, Vol. 47 No. 6, pp. 828-847, doi: 10.1177/0002764203260212.

Tepper, B.J. (2003), "Organizational citizenship behavior and the spiritual employee", in Giacalone, R.A. and Jurkiewicz, C.L. (Eds), *Handbook of Workplace Spirituality and Organizational Performance*, Armonk: M.E. Sharpe, pp. 181-190.

Tevichapong, P. (2012), *Individual Spirit at Work and its Relationship with Employee Work Attitudes and Organizational Outcomes: An Empirical Examination in Corporate Thailand*, (Unpublished doctoral dissertation), Aston University, England.

Thomas, K.W. and Velthouse, B.A. (1990), "Cognitive elements of empowerment: an "interpretive" model of intrinsic task motivation", *The Academy of Management Review*, Vol. 15 No. 4, p. 666, doi: 10.2307/258687.

Vadera, A.K., Pratt, M.G. and Mishra, P. (2013), "Constructive deviance in organizations integrating and moving forward", *Journal of Management*, Vol. 39 No. 5, pp. 1221-1276.

Watkins, P.C. (2014), *Gratitude and the Good Life: Toward a Psychology of Appreciation*, Springer Science.

Wood, A.M., Linley, P.A., Maltby, J., Baliousis, M. and Joseph, S. (2008), "The authentic personality: a theoretical and empirical conceptualization and the development of the Authenticity Scale", *Journal of Counseling Psychology*, Vol. 55 No. 3, pp. 385-399, doi: 10.1037/0022-0167.55.3.385.

Zhang, X. and Bartol, K.M. (2010), "Linking empowering leadership and employee creativity: the influence of psychological empowerment, intrinsic motivation, and creative process engagement", *Academy of Management Journal*, Vol. 53 No. 1, pp. 107-128, doi: 10.5465/amj.2010.48037118.

Zou, W., Zeng, Y., Peng, Q., Xin, Y., Chen, J. and Houghton, J.D. (2020), "The influence of spiritual leadership on the subjective well-being of Chinese registered nurses", *Journal of Nursing Management*, Vol. 28 No. 6, pp. 1432-1442, doi: 10.1111/jonm.13106.

Further reading

Chakraborty, D. and Chakraborty, S.K. (2006), "The 'nishkam karma' principle: it's relevance to effectiveness and ethics", *IIMB Management Review*, Vol. 18, pp. 115-125.

Garg, N. (2015), "Readiness of Indian Inc. for modern HRM practices", *IUP Journal of Organizational Behaviour*, Vol. 14 No. 1, pp. 58-76.

Gupta, V. and Singh, S. (2014), "Psychological capital as a mediator of the relationship between leadership and creative performance behaviors: empirical evidence from the indian R&D sector", *The International Journal of Human Resource Management*, Vol. 25 No. 10, pp. 1373-1394, doi: 10.1080/09585192.2013.870311.

Hasan, N., Mitschke, D. and Ravi, K. (2018), "Exploring the role of faith in resettlement among Muslim Syrian refugees", *Journal of Religion and Spirituality in Social Work: Social Thought*, Vol. 37 No. 3, pp. 223-238.

Pradhan, R.K. and Jena, L.K. (2016), "Employee performance at workplace: conceptual model and empirical validation", *Business Perspectives and Research*, Vol. 5 No. 1, pp. 69-85, doi: 10.1177/2278533716671630.

Rego, A., Sousa, F., Marques, C. and Cunha, M.P. (2012), "Authentic leadership promoting employees' psychological capital and creativity", *Journal of Business Research*, Vol. 65 No. 3, pp. 429-437.

Sweetman, D., Luthans, F., Avey, J.B. and Luthans, B.C. (2010), "Relationship between positive psychological capital and creative performance", *Canadian Journal of Administrative Sciences/Revue Canadienne Des Sciences de l'Administration*, Vol. 28 No. 1, pp. 4-13, doi: 10.1002/cjas.175.

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**

Environmental Science and Pollution Research
https://doi.org/10.1007/s11356-021-18049-2

REVIEW ARTICLE

Review on fabrication methodologies and its impacts on performance of dye-sensitized solar cells

Geetam Richhariya[1] · Bhim Charan Meikap[2] · Anil Kumar[3,4]

Received: 21 August 2021 / Accepted: 7 December 2021
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

Abstract

This review highlights and summarizes the impact of different fabrication processes on the efficiency of dye-sensitized solar cells (DSSCs). Energy conversion efficiency of cell depends upon semiconductor, sensitizer, electrolyte, and counter electrode. Efficiency of DSSCs can be enhanced by properly selecting the optimum significance of various parameters of fabrications process. Major challenges of these solar cells are non-vegetal, noxious, extreme sensitizers. Application of natural dyes in this field plays a significant role. An optimized CdSe-TiO_2 photoanode showed a power conversion efficiency (PCE) of 13.29% and short circuit current density of 15.30 mA cm^{-2} for the DSSC. Power conversion efficiency of 3.26% was achieved by using TTO electrode for DSSC device that is ascribed to the improved electrical and optical properties due to doping with Ta element. Absorbance of betalain was shown in the visible range of 530–535 nm for betanin while 450–559 nm for anthocyanin pigment. The natural dyes are economical, readily available, and environmentally friendly. This compilation would be beneficial for researchers working on dye solar cell.

Keywords Dye-sensitized solar cell · Working electrode · Natural dye · Electrolyte · Counter electrode

Introduction

All over the world, non-renewable energy based power plants such as thermal power plants, nuclear power plants, diesel and gas turbine power plants, etc. are used to generate electricity. These fossil fuels pollute the environment due to carbon emissions (Polo et al. 2004). On the other hand, the application of renewable energy sources results in clean surroundings (Awasthi et al. 2020) and a healthy environment across the universe (Sharma et al. 2015). There are various renewable energy sources such as wind energy, solar energy, and hydro energy (Fukurozaki et al. 2013). Solar energy is the more convenient renewable energy source due to simple principles and easy construction of conversion devices. Photovoltaic is a device that converts solar energy into electricity (Kumar et al. 2015). There are three generations of solar cells: (i) crystalline solar cell, (ii) thin-film solar cell, and (iii) dye-sensitized solar cell (Ludin et al. 2014). A dye-sensitized solar cell (DSSC) is particularly interested in research and industries due to the novel and emerging solar cell generation. These solar cells have different parts, photosensitizer, photoelectrode, liquid electrolyte, and counter electrode, as shown in Fig. 1.

Highlights
- Different fabrication methodologies of dye-sensitized solar cells are described.
- Ideal features of fabrication parameters of DSSCs are explained.
- Various natural sensitizers for a healthy environment are discussed.
- Performance evaluation between natural and synthetic dye is described.

Responsible Editor: Philippe Garrigues

✉ Anil Kumar
anilkumar76@dtu.ac.in

[1] Energy Centre, Maulana Azad National Institute of Technology, Bhopal 462051, India

[2] Department of Chemical Engineering, Indian Institute of Technology, Kharagpur 721302, India

[3] Department of Mechanical Engineering, Delhi Technological University, Delhi 110 042, India

[4] Centre for Energy and Environment, Delhi Technological University Delhi, Delhi 110 042, India

Published online: 03 January 2022

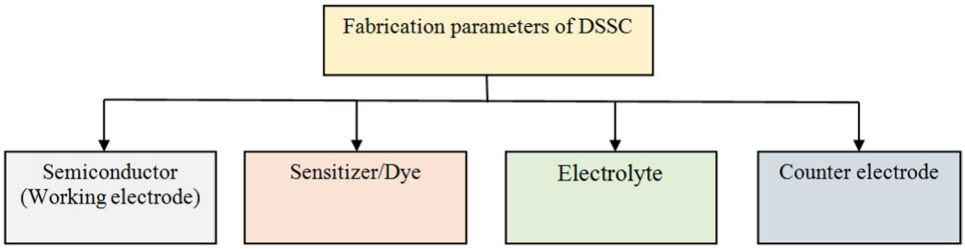


Fig. 1 Fabrication parameters of DSSC

Photosensitizer converts incident solar radiation into electric voltage. Photoelectrode is made up of a nanocrystalline semiconductor that absorbs solar radiation. During illumination, electrons move from sensitizer to the working electrode (semiconductor). Liquid electrolyte consists of a redox couple that regenerates the dye. Some researchers have been examined the factors that influence DSSC performance, resulting in good alternatives. The porphyrin-dyed ZnPorph-L-H2Porph(ZnP-H2P) and metal-free organic dye solar cell were used to make co-sensitized DSSC (Sharma et al. 2014). Co-sensitized (mixed) dye solar cells C (Zn{Porph}-L-H_2{Porph}/DC) and D (Zn{Porph}-L-H_2{Porph}) were obtained by immersing the working electrode in separate solutions of sensitizers in varying sequences. Power conversion efficiency of C and D was observed as 6.16% and 4.80%, respectively. Three different types of dye solar cells were prepared by treating titanium dioxide electrodes in three different ways, untreated, post-treated, and pre- and post-treated, through titanium tetrachloride ($TiCl_4$) solution. Pre- and post-treated working electrodes showed the best results and gave 5.1% conversion efficiency. This novel working electrode raises the thickness of TiO_2, thus increasing the current density of the solar cell. Therefore, the thickness of TiO_2 raises cell performance (Sedghi and Miankushki 2013). Nitrogen-doped TiO_2 (NSND-TiO_2) nanoparticle was prepared using modified flame spray pyrolysis (FSP) equipment. NSND-TiO_2 was formed by spraying ammonia water above the flame in which TiO_2 particles were made. Addition of nitrogen elements to TiO_2 showed a decreased energy band gap to 2.90eV and recorded 6.03% power conversion efficiency (Huo et al. 2014). Naturally occurring dyes available in various kinds of roots, leaves, fruits, and flowers show a viable environmentally friendly alternative and can be employed in DSSCs. Various natural dyes named *Anethum graveolens*, parsley, arugula, *Spinacia oleracea*, and green algae to fabricate DSSCs were treated using two different methods, i.e., before and after drying raw materials. It has been found that some DSSCs using after drying method showed better performance and others revealed better results using before drying method. The *Spinacia oleracea* extract gave the highest efficiency after the drying method was recorded as 0.29% (Taya et al. 2013). Benzonitrile-based electrolytes have low vapor pressure, thus causing more extended stability to the solar cells using N179 dye (Latini et al. 2014). Five aldimine derivatives were prepared by condensing the suitable amine with salicylaldehyde (m1–m4) and 4-aminobenzoic acid with 2-thiophene carboxaldehyde (m5). Highest efficiency of 0.575% was obtained with 2-(2-hydroxybenzylideneamino) benzoic acid-based dye solar cell (Batniji et al. 2014).

The objectives of this review are (i) to describe different fabrication methodologies of dye-sensitized solar cell while analyzing the impact of various technical procedures on the photovoltaic performance of DSSC; (ii) to discuss various ideal features of performance parameters for DSSCs; (iii) to discuss the detailed description of various natural sensitizers for the healthy environment; and (iv) to discuss performance evaluation between natural and synthetic DSSC. This study will be fruitful for the researcher to understand the construction and functioning of natural DSSC. Also, it provides knowledge in the field of natural dyes in the environment.

Performance parameters of DSSCs

Open-circuit voltage, short circuit current, fill factor, maximum voltage, and maximum current are the various distinct performance parameters of dye solar cell. All the parameters of DSSC discussed are as follows (Fig. 2):

Open-circuit voltage

Open-circuit voltage is the difference in potential between the two terminals in the cell under light illumination when the circuit is open. It is influenced by the fermi level

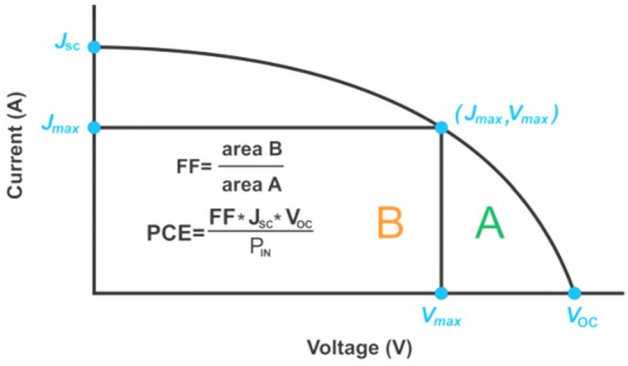


Fig. 2 Performance parameters of DSSC (Sharma et al. 2020)

(semiconductor) and the level of dark current. It is calculated when the current through the DSSC is zero (open circuit). Open-circuit voltage is dependent on the cell temperature. At the open terminals of DSSC, the voltage is mentioned as open-circuit voltage. The following is the expression for open-circuit voltage (Takagi et al. 2013):

$$V_{OC} = V_t \ln\left\{\left(\frac{I_{SC}}{I_O}\right) + 1\right\} \tag{1}$$

Short circuit current

Short circuit current density (I_{sc}) is the photocurrent per unit area (mA/cm^2) when an illuminated DSSC is short-circuited. It relies on various factors such as light absorption, light intensity, regeneration of the oxidized dye, and injection efficiency. It is strongly associated with incident photon conversion efficiency and theoretical values of the current density can be determined from the IPCE spectrum. At the short circuit terminals of DSSC, the current is mentioned as short circuit current. When the cell temperature rises, short circuit current also rises (Takagi et al. 2013). The expression for short circuit current is as follows (Lee et al. 2005):

$$I_{SC} = I + I_O\left\{\exp\left(\frac{V}{V_t}\right) - 1\right\} \tag{2}$$

Fill factor

Fill factor measures the ideality of the DSSC and is defined as the ratio of the maximum power output per unit area to the product of open-circuit voltage and current density. Several factors can affect the fill factor, such as a high internal resistance value that will offer a low fill factor and a reduced overall efficiency. It is expressed as follows (Zhou et al. 2007):

$$FF = \frac{V_m I_m}{V_{OC}\, I_{SC}} \tag{3}$$

Efficiency

Overall solar energy to the electrical conversion efficiency of a dye solar cell is defined as the maximum value of the cell output divided by the incident light power. It is estimated by measuring photocurrent density at short circuit (J_{sc}), the open-circuit photovoltage (V_{oc}) at open-circuit terminals, and intensity of the incident light and fill factor of the cell (FF), as shown in Eq. 4. Since it is reliant on all the three factors under standard conditions (Mbonyiryivuze et al. 2015), it is of great significance to optimize each one of them for high overall efficiency (Nan et al. 2017):

$$\eta = FF\left(\frac{V_{oc}\, I_{SC}}{Incident\ Solar\ Power}\right) \tag{4}$$

Fabrication elements

Structure of DSSC is shown in Fig. 3, displaying all components of the cell. Structure of DSSC is varied from the first and second generations of the solar cell. DSSC consists of specially four parts: (i) working electrode (WE); (ii) sensitizer; (iii) electrolyte; (iv) counter electrode (Gratzel 2003). Optimization of all the components is of immense meaning to enhance the overall efficiency of the cell. Working substrate is made up of semiconductor material deposited on the transparent conductive oxide. In general, metal oxides with wide band gaps are used as materials for working electrodes. Titanium dioxide, zinc oxide, and tin oxide have been extensively used in dye-sensitized solar cells (Hagfeldt et al. 2010). Semiconductor metal oxides nanoparticles form nanoporous and semi-transparent working electrodes. To sensitize working electrode, sensitizer is used that enhances semiconductor conductivity. On the surface of the working electrode, a layer of dye molecules is liable for electron harvest. Ruthenium-based dyes such as N719 (known as black dye) (Nazeeruddin et al. 1999) and N3 have been extensively employed as synthetic dyes for DSSCs (Nazeeruddin et al. 1993). N719 or N3 based DSSCs have been shown the highest efficiency of the cell. Electrolyte is filled between the working and counter electrode to complete the flow of electrons. Iodide/triiodide (I^-/I_3^-) redox couple has been widely preferred because of its appropriate

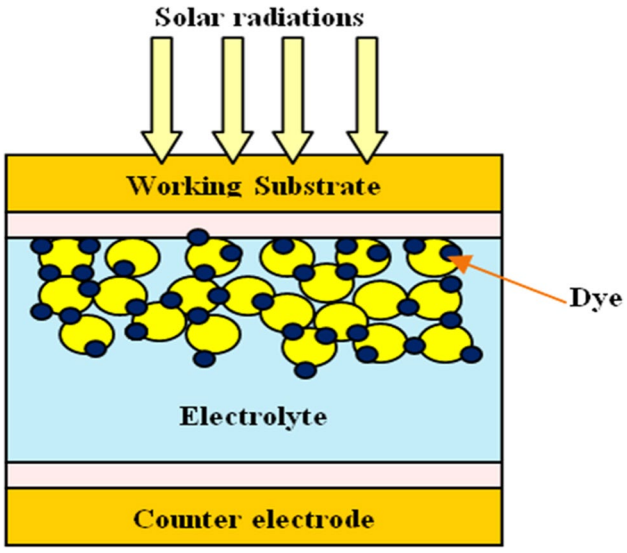


Fig. 3 Constructional parts of a dye-sensitized solar cell (Rawal et al. 2015)

redox potential (Daeneke et al. 2011). Other redox couples such as Br^-/Br_3^- (Wang et al. 2005), $SCN^-/(SCN)_3^-$, and $SeCN^-/(SeCN)_3^-$ (Oskam et al. 2001) have been striking due to their redox potentials closer to the ground state of the dyes, which provide a higher output voltage. Counter electrode is used to catalyze the regeneration of I^-. Platinum is mostly used as a counter electrode due to its better performance. Efficient regeneration of iodide enhances the diffusion rate of iodide and thus increases the reduction rate of the oxidized dye.

Working substrate

Working electrode (anode) consists of a mesoporous extensive bandgap semiconductor layer. Titanium dioxide is the most common material used as a semiconductor for DSSC. Nanoparticle layer of titanium dioxide is sintered on the working electrode to enhance electronic conduction. Mesoporous layer is deposited on a transparent conducting oxide (TCO). The fluorine-doped tin oxide (FTO) and indium-doped tin oxide are mainly transparent conducting oxide. Plastic foils (Richhariya and Kumar 2016) and metal sheets (Heo et al. 2009) are also used as materials for the working substrate of DSSC (Jun et al. 2007). Merits and failures of different substrates are shown in Table 1.

DSSC efficiency can be increased by using suitable functionalized semiconducting metal oxides having quantum dots, organic conjugated polymers, etc. Photoanodes are fabricated using semiconductor oxides having wide bandgap. Such metal oxides can absorb almost all the incident sunlight and have stability against photo-corrosion (Kumar et al. 2020). An optimized CdSe-TiO_2 photoanode showed a power conversion efficiency (PCE) of 13.29% and short circuit current density of 15.30 mA cm^{-2} for the DSSC, as shown in Fig. 4 (Bhattacharya and Datta 2020).

A facile spray pyrolysis method was used to deposit a thin film of Ta-doped SnO_2 (TTO). Resistance stability of sheet resistance was relatively better up to 400°C representing the appropriateness of TTO electrode to provide as conducting

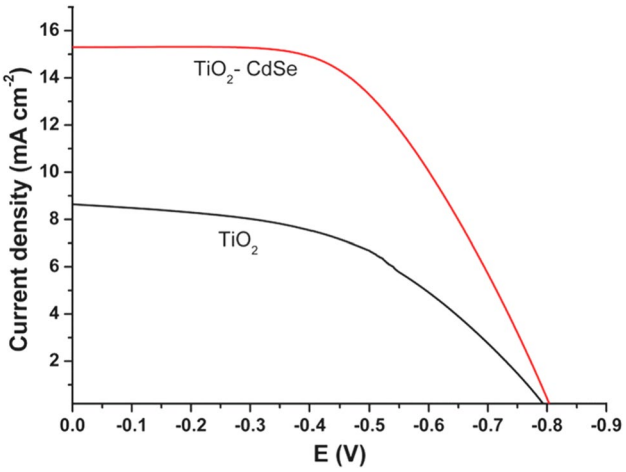


Fig. 4 J-V curve for CdSe-TiO_2-based DSSC (Bhattacharya and Datta 2020)

photoanode electrode in DSSC. XRD study showed the crystallinity, phase purity, and texture properties of the thin film (Fig. 5). Power conversion efficiency of 3.26% was achieved using the TTO electrode for DSSC device ascribed to the improved electrical and optical properties due to doping with Ta element (Ramarajan et al. 2020).

Materials of working substrate

Titanium dioxide is the best semiconductor for DSSC due to its suitable properties that match the solar cell requirement. Conversion efficiency of 2.61% was achieved by employing two-step hydrothermal method for synthesizing TiO_2 materials. Paste was prepared by grinding 6.0g of TiO_2 with ethanol and water, and then mixed with terpineol and ethyl cellulose after sonication (40). Photoanode with 23μm thickness was prepared using shear-exfoliated graphene, causing less electron recombination and giving cell efficiency of 8.9%. Nanofiber-based photoanode causes a higher electron diffusion coefficient (Lo and Leung 2019). Two grams of

Table 1 Merits and failures of working electrodes

Working electrode	Merits	Failures	Reference
TCO	-Transparent in nature	-Uneconomical	(Weerasinghe et al. 2013)
	-Greater conductivity	-Breakable	(Hashmi et al. 2011)
		-Hard construction	(Narayan 2012)
Polymers-based substrate	-Flexibility of cell	-Less efficiency	(Han et al. 2015)
	-Suitable for roll to roll fabrication technique	-Less temperature tolerance	
Metallic substrate	-Very high electrical conductivity	-Rusty in nature	(Yugis et al. 2015)
		-Uneconomical	(Law et al. 2005)
			(Yune et al. 2013)

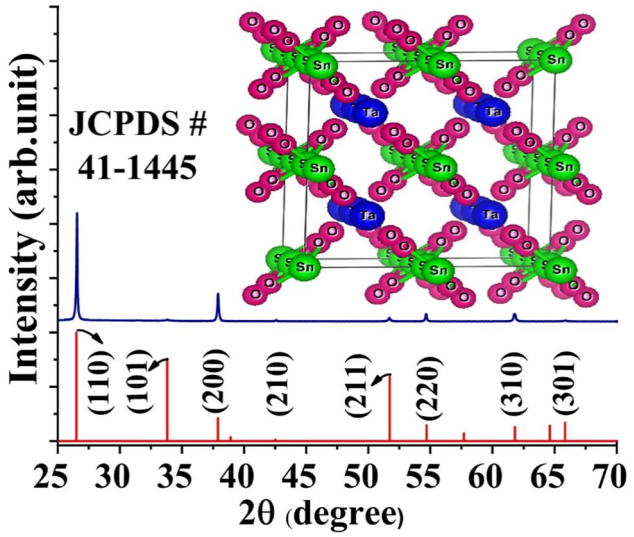


Fig. 5 XRD pattern of the TTO thin film (Ramarajan et al. 2020)

titania powder was placed in a beaker and 6mL citric acid of 0.1M, 0.2mL polyethylene glycol, 0.2mL titanium (IV) isopropoxide, and 0.1mL non-ionic surfactant Triton X-100 were added to prepare the solution. Table 2 reveals that dyes extracted using acetone acquired every parameter of DSSC superior to the other four solvents. Therefore, that experiment verified the need for solvent effect and showed that the solvent having less polarity, being as acetone, was more able to extract an effective fraction of red amaranth dye than the other four solvents (Uddin et al. 2015).

TiO_2 working substrate was synthesized using a hydrothermal gel method employing polyvinylpyrrolidone (PVP) by adding 100mL of deionized water in 10mL of titanium tetrabutanolate under vigorous agitation at room temperature. Colloid was mixed with 0.8g of polyethylene glycol, 1ml of OP emulsifier, and stoichiometric PVP (0.0270, 0.0405, 0.0540, 0.0810, 0.108, and 0.135) and subsequently concentrated at 80°C. The resultant titanium dioxide nanocrystallites were calculated as TiO_2-1 wt%, TiO_2-1.5 wt%, TiO_2-2 wt%, TiO_2-3 wt%, TiO_2-4 wt%, and TiO_2-5 wt%, respectively. The TiO_2-0 wt% was considered zero PVP. Compared with dense titanium dioxide working substrate, the power conversion efficiency was reported as 9.86%, as shown in Table 3 (Hu et al. 2014).

Table 3 Performance parameters of resultant dye solar cell (Hu et al. 2014)

Working substrate	V_{oc} (V)	J_{sc} (mA/cm^2)	FF	η (%)
TiO_2-0 wt%	0.729	12.82	0.686	6.41
TiO_2-1 wt%	0.740	11.92	0.729	3.43
TiO_2-1.5 wt%	0.779	17.99	0.704	9.86
TiO_2-2 wt%	0.746	17.63	0.728	9.57
TiO_2-3 wt%	0.743	15.18	0.671	7.57
TiO_2-4 wt%	0.750	11.35	0.724	7.30
TiO_2-5 wt%	0.763	11.76	0.720	6.46

Table 4 Performance parameters of silver-coated DSSCs (Peng et al. 2013)

Working electrode	Deposition time (min)	V_{oc} (V)	J_{sc} (mA/cm^2)	FF	η (%)
Silver-coated TiO_2	0	0.720	12.44	0.66	5.97
	5	0.724	12.98	0.68	6.46
	10	0.735	13.55	0.69	6.86
	15	0.730	12.76	0.66	6.18
	30	0.730	12.22	0.57	5.15

TiO_2 colloidal dispersions synthesize hydrothermally at 200°C by using titanium isopropoxide gels and acetic acid under non-ionic surfactant. Particle size varying from 15 to 20nm showed the cell's average efficiency as 5.2% (Muniz et al. 2011). TiO_2-based working electrode was fabricated using silver coating photo-deposition with distinct light durations. In a solution of 5×10^{-4} mol/L $AgNO_3$, TiO_2 working electrode was dipped in distinct durations of 5, 10, 15, and 30 min, and ultraviolet radiations irradiated the solutions. This novel fabrication enhanced the conversion efficiency of the DSSC 6.86% in deposition time of 10 min with fill factor of 69.4% (Table 4). IPCE spectra of dye solar cell using TiO_2 and TiO_2-Ag-coated electrode are shown in Fig. 6. There is higher IPCE in the case of the silver-coated electrode (Peng et al. 2013). Anatase TiO_2 was synthesized using a hydrothermal process in two steps using amine ligands and achieved a conversion efficiency of 2.61% for the smallest nanoparticles (Phonkhokkong et al. 2016).

Table 2 Performance parameters of DSSC using different solvents (Uddin et al. 2015)

Working electrode	Solvent	V_{oc} (V)	J_{sc} (mA/cm^2)	P_{max} (mW)	FF	η (%)
TiO_2	Water (room temp)	0.452	0.17	0.043	0.56	0.04
	Boiling water	0.450	0.47	0.117	0.55	0.12
	Ethanol	0.390	0.27	0.066	0.63	0.07
	Methanol	0.435	0.42	0.107	0.59	0.11
	Acetone	0.475	0.65	0.213	0.69	0.21

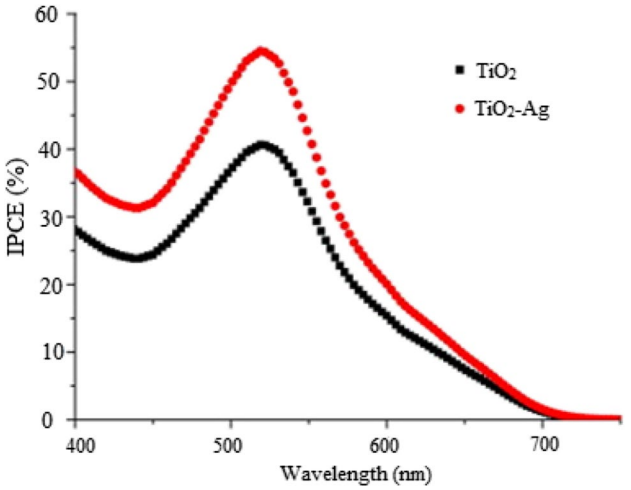


Fig. 6 IPCE spectra of dye solar cell using TiO_2 and TiO_2-Ag-coated electrode (Peng et al. 2013)

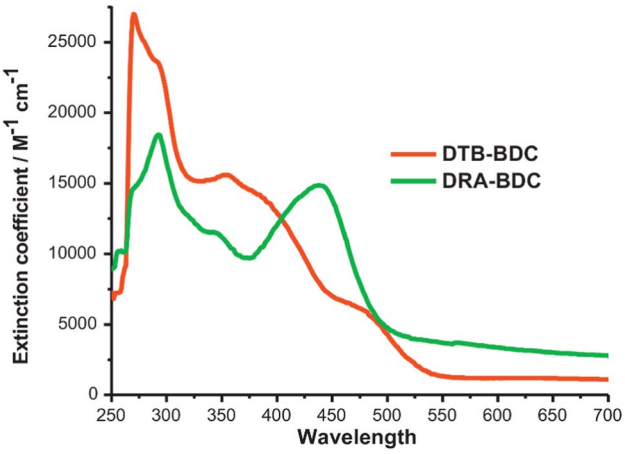


Fig. 7 Absorption spectra of organic dyes (Moustafa et al. 2021)

Sensitizer

A layer of the sensitizer is adsorbed onto the surface of the semiconductor by chemical bonding. The primary role of the sensitizer is to absorb the incident light, infuse the excited electron into the semiconductor, and turn it into regenerated by the redox couple in the electrolyte. Dye (sensitizer) is the heart of dye-sensitized solar cells, making the DSSC different from other solar cells. Metal complex sensitizers, metal-free organic sensitizers, and natural sensitizers are the most common varieties of sensitizers in DSSC. Natural sensitizers are environmentally friendly and easily extractable using a simple extraction process (Calogero et al. 2012). Two new organic dyes were DRA-BDC and DTB-BDC, with electron acceptors as rhodanine-3-acetic acid/thiobarbituric acid while electron donor as N,N-butyldicarbazole, which have been used as dye for DSSC. DRA-BDC and DTB-BDC revealed higher absorption peaks at 440 and 370 nm, respectively, in UV-visible spectra of dyes (Fig. 7). Novel sensitizer has a larger energy gap between its LUMO and conduction band of TiO_2, leading to increased current density ($2.46 mA/cm^2$), thus contributing to the enhancement of the PCE (1.16%) (Moustafa et al. 2021).

Two novel compounds derived from the phenyltetrazole system, 5-(4-decyloxyphenyl)tetrazole (LTz-4) and N,N-diethyl-4-{((2′-nitro-4′-tetrazoyl)phenyl)diazenyl}aniline (SD - 6), were used as co-adsorbents for DSSC. Highest open-circuit voltage (0.71V) and PCE (9.20%) were observed in the case of dye HD-14 (ruthenium complex) with co-adsorbent DCA (Fig. 8). High value of photocurrent density can be due to the co-adsorbent, favoring the capture of photons and encouraging the introduction of electrons into semiconductors (Silva et al. 2020).

Natural dye was extracted from the daisy flowers family (*Leucanthemum vulgare*), namely yellow daisy, purple daisy, and wine daisy, containing luteolin flavonoid used to fabricate DSSCs. The yellow daisy-, wine daisy-, and purple daisy-based DSSC revealed PCE of 0.6%, 0.4%, and 0.8%, respectively (Fig. 9) (Ferreira et al. 2020).

Materials of sensitizer

There are various methods used for the fabrication of dye for DSSC using different natural dyes. Extracts from *Dianthus barbatus*, *Lepidium sativum*, and *Raphanus raphanistrum* were employed as dyes in titanium dioxide nano-powder-based dye solar cell. Highest efficiency and fill factor of 0.15% and 0.48 were observed from *Dianthus barbatus* (Researcher et al. 2016). One gram of rosella and blue pea flower was taken in 100mL of two different solvents at various temperatures (25, 50, 50, 70, and 100°C). The observed conversion efficiency of DSSCs using extracts of rosella, blue pea, and mixed dye extracted at 100°C was 0.37%, 0.05%, and 0.15%, respectively (Table 5) (Hamadanian et al. 2014).

Anthocyanin was extracted from maqui berry by drying 0.5g of maqui berries at 45°C and dipped in 5mL of distilled water. At 750 and 1500mg of anthocyanin/L, the extracted concentration of maqui was examined and power conversion efficiency was obtained as 0.14 % and 0.19%, respectively (Table 6) (Ozuomba et al. 2013).

Natural dyes from *Reseda luteola*, *Berberis integerrima*, *Panica granatum*, Pleniflora, *Consolida orientalis*, *Reseda gredensis*, *Clematis orientalis*, *Adonis flammea*, *Salvia sclarea*, and *Consolida ajacis* plants were extracted using Soxhlet extractor. Identifying suitable natural dyes for improving V_{OC} without resulting dye degradation may result in additional improvement of cell performance. The

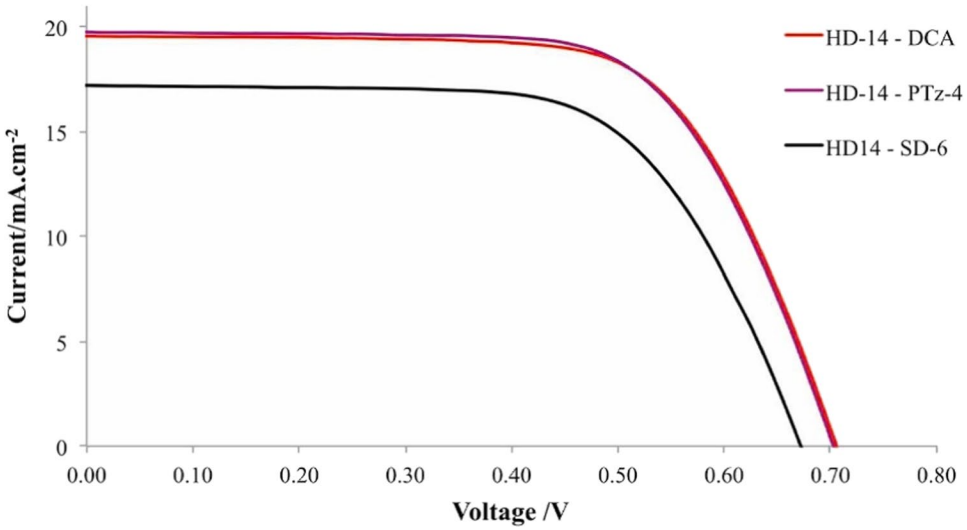


Fig. 8 J-V characteristics for HD-14 (ruthenium complex) with co-adsorbent DCA (Silva et al. 2020)

maximum energy conversion of DSSC using Reseda luteola extract is 0.22%. (Table 7) (Phinjaturus et al. 2016).

Two dye solar cells were fabricated using anthocyanin extracted from *Hibiscus sabdariffa* and plain (un-dyed) cells, which reported an energy conversion efficiency of 0.58% for dyed cells while 0.03% for un-dyed solar cell. Husk, cob, and silk of purple corn were extracted in acetone, ethanol, and DI waters. Maximum conversion efficiency of 1.06% was obtained from husk extracted from acetone, as given in Table 8 (Ananth et al. 2015).

TiO_2-based DSSC was sensitized using natural dye formed by *Pterocarpus marsupium* stem bark. One hundred milliliters of distilled water was added in 250g of *Pterocarpus marsupium* stem bark (cut pieces) for a duration of 14 h, and filtered out. In this concentrated solution, distilled water of 250mL and isopropanol of 15mL were added. Initial pH value was recorded as 8.75. Nitric acid was added drop wise to raise the pH value up to 2, and the solution was vigorously stirred. After that, the titanium isopropoxide solution of 5mL was mixed drop wise until white precipitation was formed. Solution was then heated up to 80°C for 2–3 h. At room temperature, the solution was going to foraging for 1 h. The resultant white precipitate was left to dry at 150°C for 15–20 h to form fine particles of pure titanium dioxide. Table 9 displays the conversion efficiency of DSSC using pre-dyed titanium nanoparticles as 0.49% (Jia et al. 2018).

Novel X type D-$(\pi\text{-A})_2$ organic dyes JX1 and JX2 were developed and co-sensitized with porphyrin dyes JP1 and JP3 for dye solar cells. JX1-based DSSC showed photo-conversion efficiency (PCE) of 3.67%, while JX2-based DSSC showed PCE of 4.63%. Co-sensitized dye solar cell (JP3+JX2) showed the highest efficiency of 8.08% (Zeng et al. 2010).

Electrolyte

Electrolyte regenerates the sensitizer of dye solar cells. Redox couple present in the electrolyte is an essential feature of the electrolyte for better cell performance. The function of electrolytes is to transfer the electron between two electrodes. Active transport of electrons within the dye solar cell is essential, and many different redox systems have been explored. Iodide triiodide is the most common liquid electrolyte. However, severe problems such as electrolyte volatility can lead to long-term stability concerns because of complications in sealing the device. Transform of the volatile electrolyte into a nonvolatile ionic liquid (Choi et al. 2008) has been recognized to be victorious (Gorlov and Kloo 2008), and high efficiencies and excellent stabilities have been reported (Tian et al. 2000). Numerous other iodine-free redox mediators have been experienced. Both complexes (Wang et al. 2010) and organic redox couples have been verified to be potential alternatives (Ayalew and Ayele 2016). Electrolyte configuration can be further changed to enhance the performance of the cell. DSSC has been fabricated using gel polymer electrolytes (GPEs) with polyacrylonitrile (PAN)-based polymer with variant number of tetrabutylammonium iodide (TBAI) salt and iodine. The S3 electrolyte-based DSSC showed the highest PCE of 3.45%, with V_{oc} of 582 mV and J_{sc} of 12.9 mA cm^{-2} (Fig. 10). High PCE of DSSC (S3 electrolyte) over a visible wavelength range shows superior photon harvesting efficiency (Chowdhury et al. 2020).

Gel polymer electrolyte (GPE) has been prepared using different lithium iodide concentrations (LiI). Solution was contained PEO and PVA in same quantity, ethylene carbonate (EC), tetrabutylammonium iodide (TBAI), dimethyl sulfoxide (DMSO), and iodine crystals (I_2). DSSC fabricated

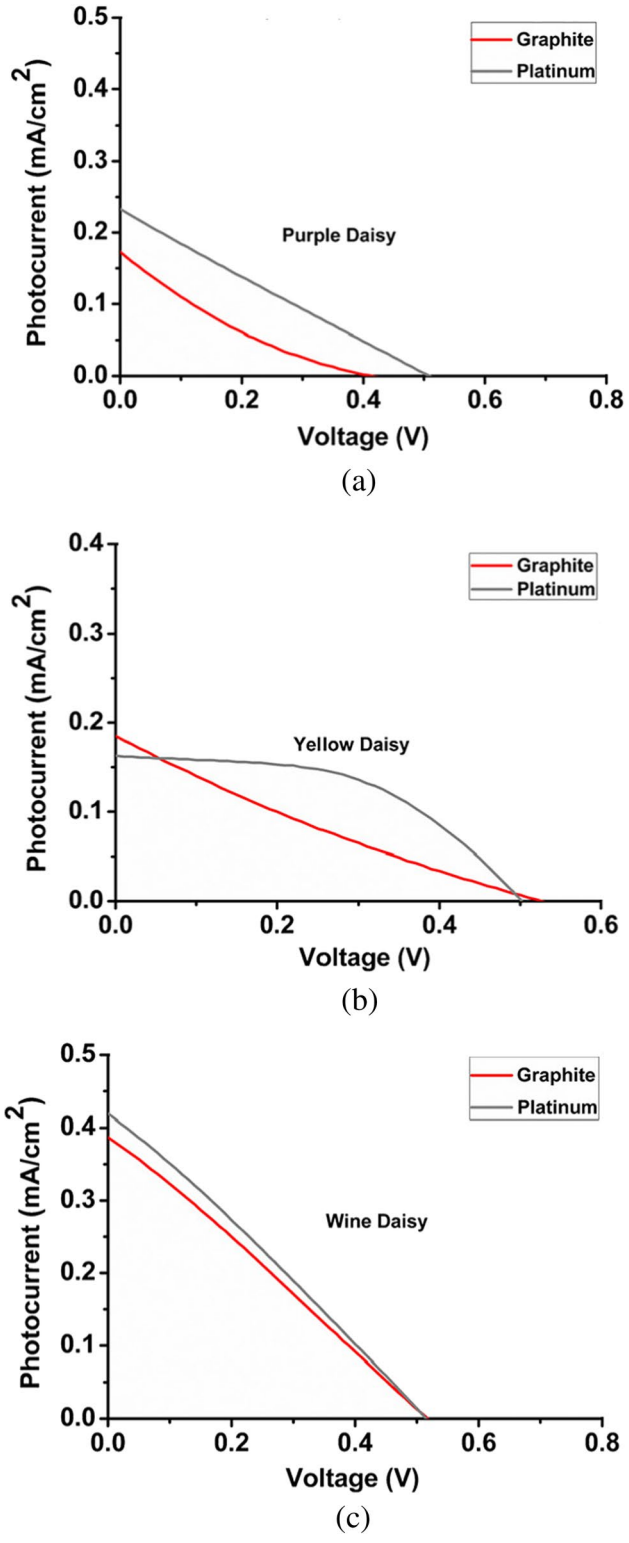


Fig. 9 J-V curves for dyes: **a** purple; **b** yellow; **c** wine (Ferreira et al. 2020)

Table 5 Performance parameters of DSSCs using natural extracts (Hamadanian et al. 2014)

Dye	V_{oc} (V)	J_{sc} (mA cm^{-2})	FF	η (%)
Mixed dye	382	0.82	0.47	0.15
Rosella	404	1.63	0.57	0.37
Blue pea	372	0.37	0.33	0.05

Table 6 Performance parameters of the DSSCs using maqui berries at concentrations of 750mg/L and 1500mg/L (Ozuomba et al. 2013)

Anthocyanin concentration (mg/L)	V_{oc} (V)	J_{sc} (mA/cm^2)	FF	η (%)
Maqui (1500mg/L)	0.44	0.44	0.54	0.19
Maqui (750mg/L)	0.43	0.37	0.49	0.14

Table 7 Performance parameters of natural dyes extracts of various plants (Phinjaturus et al. 2016)

Natural dye	V_{oc} (mV)	I_{sc} (mA)	FF	η (%)
Reseda luteola	0.64	0.54	0.63	0.22
Berberis integerrima	0.56	0.004	0.53	0.01
Panica granatum Pleniflora	0.62	0.40	0.61	0.15
Consolida orientalis	0.60	0.56	0.53	0.18
Reseda gredensis	0.53	0.14	0.71	0.07
Consolida ajacis	0.55	1.68	0.65	0.60
Adonis flammea	0.59	0.40	0.66	0.16
Salvia sclarea	037	0.10	0.54	0.02
Clematis orientalis	0.42	0.22	0.49	0.05

Table 8 Performance parameters using different solvents (Ananth et al. 2015)

Solvent	Natural dye	V_{oc} (V)	J_{sc} (mA/cm^2)	FF	η (%)
Acetone	Husk	0.46	3.57	0.64	1.06
Acetone	Cob	0.48	3.42	0.62	1.01
Acetone	Silk	0.48	3.25	0.62	0.96
Ethanol	Husk	0.43	2.91	0.61	0.76
Ethanol	Cob	0.45	3.21	0.61	0.89
Ethanol	Silk	0.47	2.19	0.64	0.65
DI water	Husk	0.40	2.13	0.59	0.50
DI water	Cob	0.42	2.29	0.56	0.53
DI water	Silk	0.47	1.46	0.61	0.43

from GPE with 1.34 wt% LiI gave the highest efficiency of 6.26%, as shown in Fig. 11 (Teo et al. 2018).

Table 9 Performance parameters using *Euphorbia cotinifolia* leaves and *Acanthus sennii* chiov flower-based dye solar cell (Jia et al. 2018)

Natural dye	V_{oc} (V)	J_{sc} (mA/cm^2)	FF	η (%)
Euphorbia cotinifolia leaves	0.436	0.642	0.48	0.136
Acanthus sennii chiov flower	0.507	0.491	0.60	0.150

Materials of electrolyte

Usually, iodide triiodide redox couple is generally used in the fabrication of dye solar cells. Evaporation and long-term stability are the main issues in front of liquid electrolytes. Gel polymer electrolytes were synthesized using distinguished weights (0%, 2%, 4%, 5%, 6%, and 8%) of acetamide in poly(ethylene oxide) (PEO) with LiI/I_2. Gel electrolyte was formed by mixing 0.53g of polyethylene oxide with acetonitrile and propylene carbonate (20:1) stirred for 2h continuously. After that, 0.2g of LiI, 0.04g of I_2, and different wt% of acetamide (0%, 2%, 4%, 5%, 6%, and 8%) were added and dropping was continued for 2 h; homogeneous sol was evaporated at 80°C to get gel polymer electrolyte. The highest cell efficiency 9.01% was observed with 5%wt of acetamide, as discussed in Table 10 (Pavithra et al. 2015).

Single-step methodology was employed for preparing counter electrodes using copper sulfide nanoparticle that was treated with two distinct electrolytes, Co(II)/(III) bipyridine and Fe(II)/(III) ferrocene-based liquid electrolyte. Ferrocene electrolyte was prepared by mixing 0.1M of ferrocene and 0.05M of ferrocenium tetrafluoroborate in propylene carbonate. The sol was kept

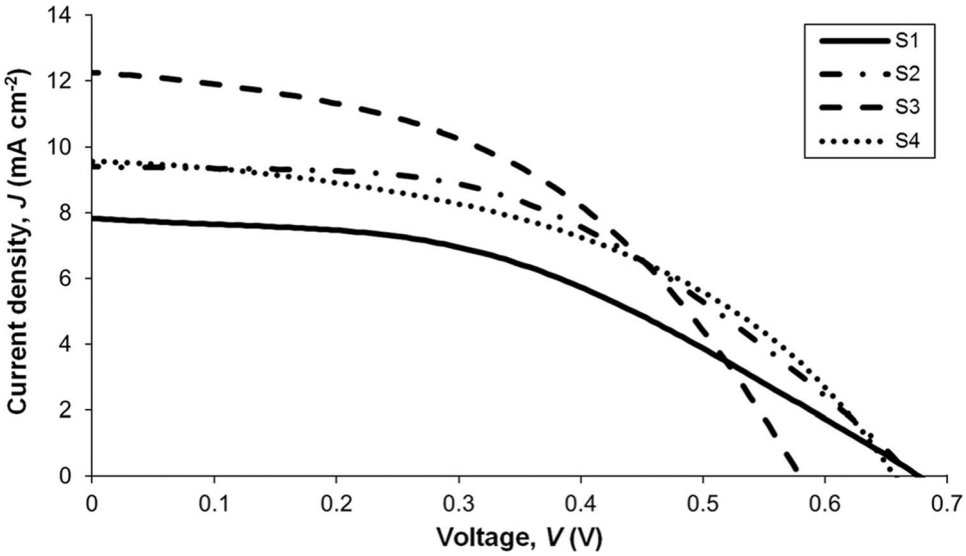


Fig. 10 J-V characteristics of DSSCs using different electrolytes (Chowdhury et al. 2020)

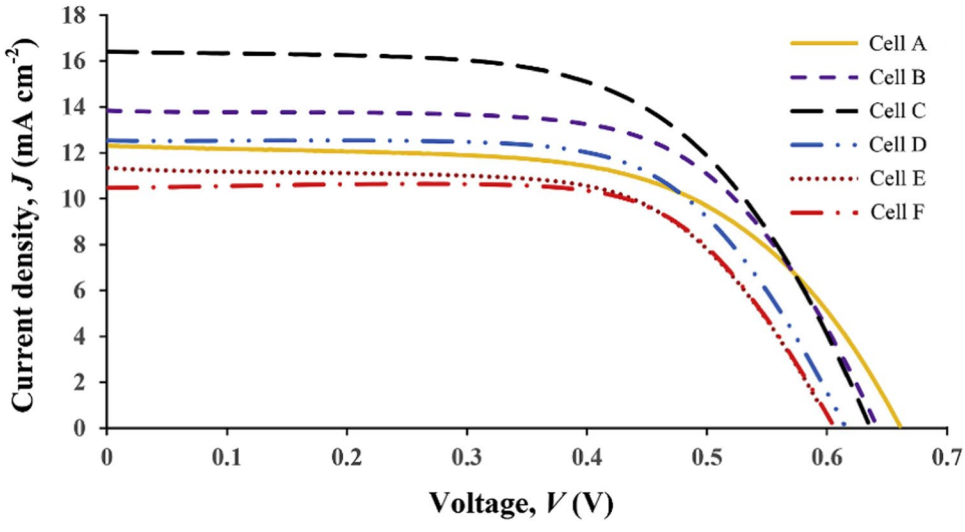


Fig. 11 J-V characteristics of DSSCs using PVA-PEO-based GPEs consisting of different ratios of TBAI and LiI salts (Teo et al. 2018)

Table 10 Performance parameters using acetamide with different weights (Pavithra et al. 2015)

Acetamide (wt%)	Electrolyte	V_{oc} (V)	J_{sc} (mA/cm^2)	FF	η (%)
-	Liquid	0.68	5.87	0.62	2.93
0	A	0.74	12.04	0.54	5.70
2	B	0.74	16.78	0.44	6.40
4	C	0.75	18.59	0.48	7.89
5	D	0.74	19.81	0.52	9.01
6	E	0.76	15.08	0.45	6.04
8	F	0.76	11.38	0.56	5.74

Table 11 Performance parameters using biopolymer electrolyte (Singh et al. 2014)

Electrolyte	V_{oc} (V)	J_{sc} (mA/cm^2)	η (%)
Biopolymer electrolyte	0.56	5.68×10^{-4}	0.63

under the nitrogen atmosphere, thus avoiding degradation of ferrocenium ion in contact with air. A high limiting current (11.8mAcm^{-2}) was observed by the ferrocene/ferrocenium redox couple (Congiu et al. 2016). Ionic liquid electrolyte enhances the thermal durability of DSSC with an energy conversion efficiency of 2.06% and a fill factor of 0.49 (Ito and Takahashi 2012). An electrolyte was developed by adding 0.172g of arrowroot powder in 20 mL of double-distilled water (D2) and stirred for 30 min at 70°C. On the other hand, potassium iodide was taken in D2. Finally, two solutions were mixed drop-by-drop and stirred continuously to produce biopolymer electrolyte, and 0.63% efficiency observed in this was observed in Table 11 (Singh et al. 2014).

Different electrolytes, liquid, solid-state, and quasi-solid-state electrolytes used in DSSC and concluded quasi-solid-state electrolytes, were more suitable for providing the best performance of the cell. Table 12 shows the performance parameters of different electrolytes (Wu et al. 2008). Photoconversion efficiency of 2.06% was achieved using heteroleptic copper(I) dyes and homoleptic copper(I)/(II) redox shuttles. It shows the potential for all copper-based dye solar cells (Karpacheva et al. 2018).

Counter electrode

The function of the counter electrode in dye solar cells is to regenerate the electrolyte. Counter electrode consists of a conducting layer on a glass (TCO) or a plastic substrate. Catalytic material is used for resulting fast chemical reactions in the DSSC, thus reducing transition time. A layer of platinum is commonly coated on the substrate for the effective regeneration of the redox couple. High catalytic activity towards the iodide/triiodide redox reaction (Pettersson et al. 2007) is the most significant advantage of using platinum as a counter electrode. In electrolytes, platinum is also chemically stable (Kay and Gratzel 1996). Various catalytic materials have been tested in dye solar cells, such as carbon (Saito et al. 2002), platinum (Bay et al. 2006), and graphene (Kay and Gratzel 1996). Increasing an extremely efficient double-featured electrode material is imperative to the identity charged integrated electronics that unite the energy conversion unit of the dye-sensitized solar cell (DSC) and the energy storage unit of the supercapacitor. Nitrogen-doped carbon sphere (NCS) was incorporated to encourage electrode activity. A power energy conversion efficiency of 8.64% was recorded, much higher than the conventional solid carbon sphere (SCS), as shown in Fig. 12 (Wang et al. 2020).

A structured configuration (n-MWCNT-TiO_2/N_3/MWCNT) was used for DSSC. In this configuration, expensive Pt was replaced by economical MWCNT that ensures less costly DSSC. Absorbance spectrum of dye is shown in Fig. 13. Presence of MWCNT in n-MWCNT-TiO_2 nanocomposite encourages the introduction of electron initiated because of dye to conduction band MWCNT-TiO_2 nanocomposite semiconductor (Younas et al. 2019a).

Pristine perfect and defective graphene nanosheets (GNSs) were formed and applied as substrates for catalytic activity against the reduction reaction of T_2 to T^-. It was observed that pristine GNSs could accomplish high reactivity to T_2/T^- through an exothermic adsorption reaction (Tontapha et al. 2019). Admirable electrocatalytic properties and economics are major necessities for the appropriate counter electrodes of dye-sensitized solar cells. DSSCs with the optimal PEDOT-Ni_2P-3 electrode and PEDOT-CO_2P-3 electrode displayed the power conversion efficiency of 7.14% and 6.85%, respectively (Table 13) (Di et al. 2019).

DSSC using tungsten oxide/titanium oxide nanocomposites (nWO_3-TiO_2) with three dissimilar WO_3-TiO_2 (1%WO_3-TiO_2/N719/MWCNT) displayed an efficiency improvement of about

Table 12 Performance parameters of DSSC using various electrolytes (Wu et al. 2008)

Electrolyte	V_{oc} (V)	J_{sc} (mA/cm^2)	FF	η (%)
0.5 M NaI, 0.05 M I_2 in 30 vol % NMP, and 70 vol % GBL mixed solvents (liquid A)	0.632	14.78	0.61	5.73
Liquid b: adding 0.4 M pyridine into liquid A (liquid B)	0.705	13.34	0.71	6.70

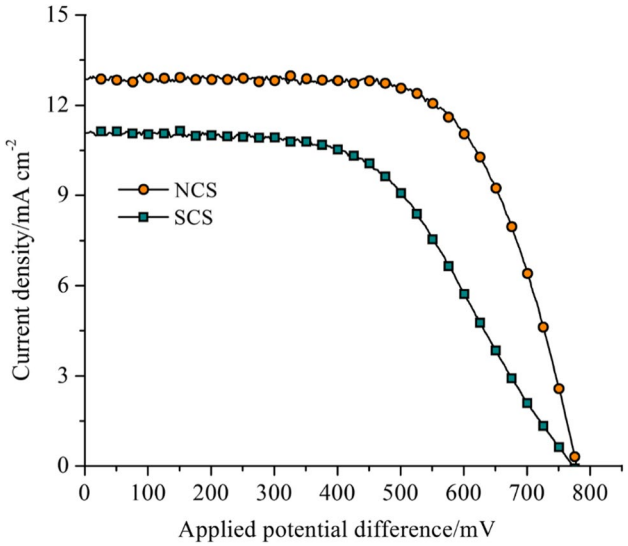


Fig. 12 J-V characteristics of nitrogen-doped carbon sphere-based DSSC (Wang et al. 2020)

40% as corresponding to the conventional TiO_2/dye/Pt solar cell as shown in Fig. 14 (Younas et al. 2019a).

Due to its remarkably low price, polyaniline (PANI) is a striking option to platinum counter electrode for dye-sensitized solar cell. Though, deposition of PANI film up to fluorine-doped tin oxide glass surface is too complicated. DSSC using Gr/PANI and Pt counter electrode displayed photoconversion efficiency of 3.58% and 3.97%, respectively (Fig. 15) (Shahid et al. 2019).

Materials of counter electrode

Counter electrode can be fabricated by using different fabrication methods. One-step hydrothermal method was used to prepare nickel sulfide hollow spheres as a counter electrode using 0.1g nickel sulfide (NiS) and 0.025 g polyethylene glycol powder in 1mL absolute ethanol solution stirred in an agate mortar to make a paste. NiS counter electrode-based DSSC showed power conversion efficiency of 6.90% (Chen et al. 2009). Novel platinum-free counter electrode named poly-3-methylthiophene (P3MT) was fabricated using an electrochemical deposition method. At room temperature, the P3MT electro-deposition was taken out in acetonitrile solution containing 0.1 M 3-methylthiophene (3MT) and 0.1M tetrabutylammonium tetrafluoroborate. An efficiency of 2.76% with a fill factor of 0.50 was reported for P3MT counter electrode-based DSSC (Yang et al. 2014). Carbon was generated in the presence of argon from the graphitization of glucose at high temperatures. Counter electrode was prepared using 1g of carbon and 0.12g of PVP solution in a mortar pestle and the mixture was carried on pot mill for 2 days for homogenous slurry. Conversion efficiency was observed as 3.63% for carbon cathode (Table 14) (Torabi et al. 2014).

Table 13 Performance parameters of PEDOT-Ni_2P-3- and PEDOT-CO_2P-3-based DSSC (Di et al. 2019)

Counter electrode	V_{oc} (V)	J_{sc} (mAcm^{-2})	FF	PCE (%)
PEDOT-CO_2P-3	0.75	13.61	0.68	6.85
PEDOT-Ni_2P-3	0.75	13.91	0.68	7.14

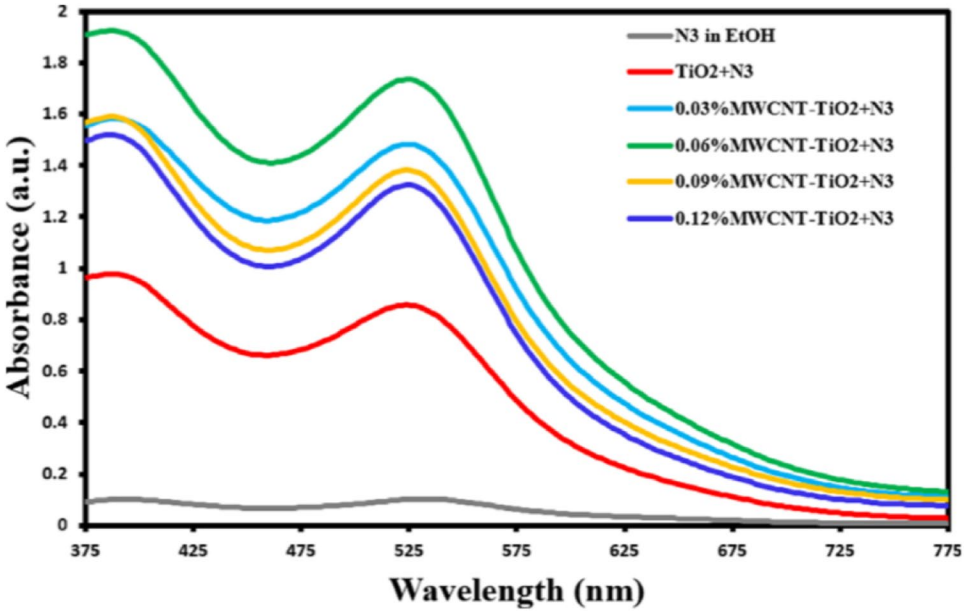


Fig. 13 UV-Vis spectra of N3 dye in n-MWCNT-TiO_2/N3, TiO_2/N3, and EtOH films (Younas et al. 2019b)

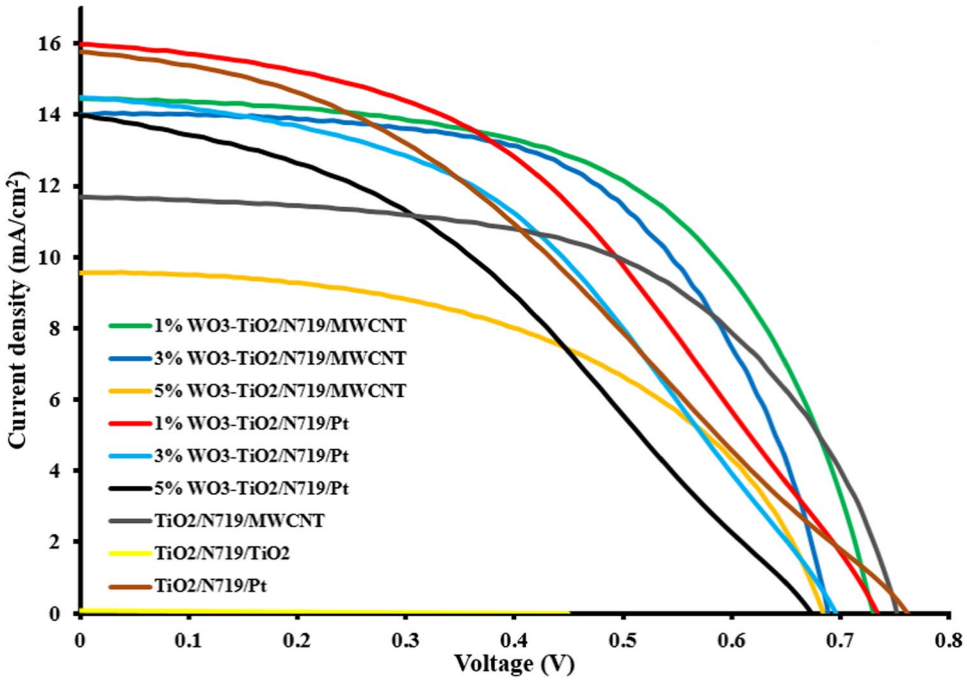


Fig. 14 J-V curves for nWO_3-TiO_2-based DSSC (Younas et al. 2019b)

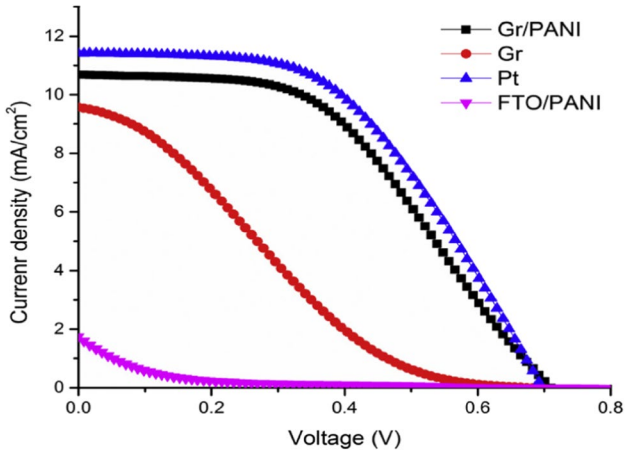


Fig. 15 J-V curves of Gr/PANI, Gr, FTO/PANI, and Pt counter electrode-based DSSCs (Shahid et al. 2019)

H_2-reduced carbon counter electrode showed the best performance of DSSC as compared to the conventional cell. Counter electrode was prepared using H_2 reduction technique. In this process, carbon was heated in a mixture of H_2-Ar gas at 450°C for half an hour. The highest power conversion efficiency was found 7.7%, with a fill factor of 0.69% and an open-circuit voltage of 0.65V, as shown in Table 15 (Kumar and Bharg 2015).

Hummers methodology was used to synthesize graphene oxide (GO) and obtained solar-reduced graphene oxide (SRGO) by reducing GO under focused sunlight. Before spraying substrate, 1mg of GO (or SRGO) was scattered in 1mL of isopropanol and then the solution was ultrasonicated for half an hour. SRGO-DSSC has shown better results (Table 16) (Takada et al. 2015).

4. Natural vs synthetic dye

Recent development in natural photosensitizer

Natural dye solar cells have tremendous advantages over synthetic dyes-based solar cells. Environmental friendly, economic, and facile synthesis chooses fabrication of dye

Table 14 Performance parameters comparison of DSSC using nickel sulfide and platinum counter electrode

Counter electrode	V_{oc} (V)	J_{sc} (mA/cm^2)	FF	η (%)	Ref.
Nickel sulfide	0.71	15.26	0.63	6.90	(Chen et al. 2009)
Platinum	0.72	15.11	0.62	6.75	
Platinum	0.7	15.29	0.47	5.06	(Yang et al. 2014)
P3MT	0.63	8.88	0.50	2.76	
Carbon	0.77	8.30	0.50	3.63	(Torabi et al. 2014)
Platinum	0.72	12.07	0.60	5.82	

Table 15 Performance parameters of DSSC under initial and after durability test (Kumar and Bharg 2015)

Counter electrode	V_{oc} (V)	J_{sc} (mA/cm^2)	FF	η (%)
After 1000 h; carbon A	0.64	17.0	0.62	6.8
After 1000 h; carbon B	0.65	17.0	0.69	7.7
Initial carbon A	0.71	15.1	0.72	7.4
Initial carbon B	0.71	14.6	0.72	7.5

Table 16 Performance parameters of DSSC using GO and SRGO (Takada et al. 2015)

Counter electrode—DSSC	V_{OC} (V)	J_{SC} (mA/cm^2)	FF	η (%)
GO-DSSC	0.69	10.83	0.45	3.38
SRGO-DSSC	0.72	12.20	0.44	3.96

solar cells using natural dyes. Performance of dye solar cell can be understood by evaluating open-circuit voltage (V_{oc}), short circuit current density (J_{sc}), fill factor (FF), and efficiency of the cell (η). Various photoelectrochemical parameters of dye solar cell with different natural sensitizers are given in Table 17.

Natural resources like *Crocus sativus* (Saffron), *Allium cepa* L (red onion), Malva sylvestris (Mallow), and oregano (Origanum vulgare) were used to prepare natural dye for DSSC. The efficiency of 0.54% was observed by red onion L. Existence of carbonyl and hydroxyl groups enables these dyes to bind to semiconductor layer (Jalali et al. 2020). Betalain and anthocyanin pigments were extracted from prickly pear and mulberry, respectively. UV-Vis spectroscopy (Fig. 16) shows absorbance in the visible range of 530–535 nm for betanin while 450–559 nm for anthocyanin pigment (Obi et al. 2020).

The effect of different organic solvents, temperatures, and pH levels for the extraction of dye from *Areca catechu was studied*. The most favorable state of dye extraction was recorded at 80°C, pH 10, and ethanol as extraction solvent, as shown in Fig. 17 (Al-Alwania et al. 2020).

Natural and synthetic dyes

The most important field of the DSSC compared to the other types of solar cells is applying the dye. It is well determined that the energy gap size of the semiconductors examines the absorption frequency of light in the solar cells. A vital purpose for using the dyes in the DSSC is to explore the absorption spectra on the visible light because the visible light has about 96% energy of the sunlight (Phonkhokkong et al. 2016). Absorption spectra of dye solar cells are determined by grouping the photoelectrode's nanoparticles, e.g., titanium dioxide and sensitizer, where dyes can assist dye solar cell in exploring their absorption spectra. Synthetic dyes synthesized from a complex methodology. On the other hand, natural dyes are extracted from natural resources which results in economic feasibility. Natural and synthetic dyes are evaluated based on economics, environmental aspects, the methodology used, performance, stability, and absorbance, as given in Table 18.

Performance evaluation between natural and synthetic

Natural and synthetic dyes are compared based on performance and fabrication parameters, as shown in Table 19.

Table 17 Recent development in natural photosensitizer

Year	Author	Sensitizer	V_{oc} (V)	I_{sc} (mA/cm^2)	FF	η (%)	Reference
2016	Maurya et al.	C. haematocephala	0.37	0.25	0.70	0.06	(Nagavolu et al. 2016)
2016	Cerda et al.	Maqui	0.06	0.00002	0.47	0.006	(Maurya et al. 2016)
2015	Lim et al.	*Canarium odontophyllu*m	0.419	3.54	0.59	0.68	(Cerda et al. 2016)
2015	Latif et al.	Olive leaves	0.59	0.85	0.33	0.17	(Lim et al. 2015)
2015	Godibo et al.	*Bougainvillea spectabilis*	0.5	1.11	0.58	0.325	(Latif et al. 2015)
2014	Hug et al.	*Bixa orellana* L.	0.57	1.1	0.59	0.37	(Godibo et al. 2015)
2014	Reddy et al.	Mangosteen	0.67	2.69	0.63	1.17	(Hug et al. 2013)
2014	Reddy et al.	Shisonin	0.53	4.80	0.51	1.13	(Mehmood et al. 2014)
2014	Chien et al.	Crude anthocyanin	0.612	2.25	0.70	0.965	(Chien and Hsu 2014)
2013	Chien et al.		0.572	3.027	0.67	1.162	(Chien and Hsu 2013)
2013	Park et al.	Gardenia yellow	0.55	0.43	0.64	0.16	(Park et al. 2012)
2013	Susanti et al.	Tamarillo	0.513	0.321	-	0.037	(Susanti et al. 2014)
2013	Esteban et al.	1A2G		1.47	0.47	0.423	(Esteban and Enriquez 2013)
2013	Kumara et al.	*Canarium odontophyllum* (CMB)	0.385	9.8	0.46	0.59	(Kumara et al. 2013)

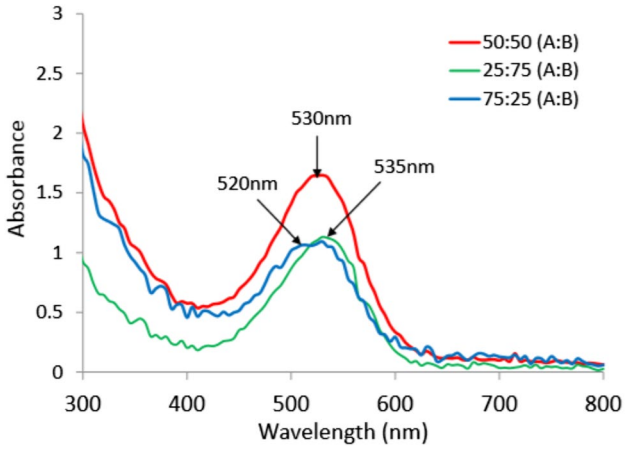


Fig. 16 Absorbance spectra of different mix concentrations of anthocyanin (**a**) and betalain (**b**) (Obi et al. 2020)

Conclusion and recommendations

This article summarizes various factors affecting performance of the DSSC with critical review. The various fabrication parameters (semiconductor, sensitizer, electrolyte, and counter electrode) affect the performance of the DSSC.

The following concluding remarks are drawn based on the above review:

- There are various fabrication methods employed for DSSC. The use of titanium dioxide showed the best results in semiconductor material, while mangosteen used as natural dye gave the highest power conversion efficiency. Also, the quasi-solid-state electrolyte used in DSSC was more suitable for DSSC performance. Conversion efficiency of 7.5% was achieved by carbon as a counter electrode.
- Shear-exfoliated graphene-based photoanode with 23µm thickness gave conversion efficiency of 8.9%, resulting in lesser electron recombination.
- Extracts of *Dianthus barbatus* gave an efficiency of 0.05%, with a fill factor of 0.48 showing progress in dye solar cells using natural dyes.
- Porphyrin dyes co-sensitized with new X organic dyes showed excellent enhancement in conversion efficiency (8.08%) compared to individual dye (4.63%).
- Dye-sensitized solar cells with copper(I) dyes and copper(I)/(II) redox shuttles achieved an efficiency of 2.06%, showing the potential of all copper-based DSSCs.
- An optimized CdSe-TiO_2 photoanode showed a power conversion efficiency (PCE) of 13.29% and short circuit current density of 15.30 mA cm^{-2} for the DSSC.
- TTO electrode-based DSSC gave power conversion efficiency of 3.26% ascribed to the improved electrical and optical properties due to doping with Ta element.

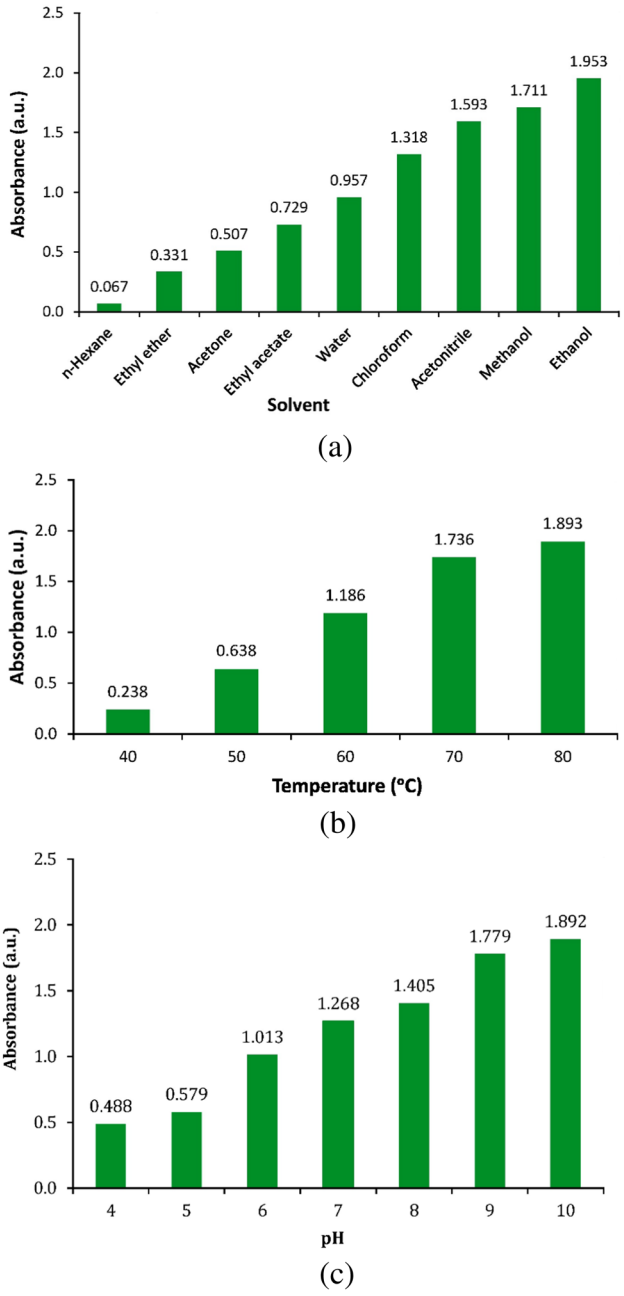


Fig. 17 Effect of **a** solvent, **b** temperature, and **c** pH on the absorbance of the dye (Al-Alwania et al. 2020)

- Absorbance of betalain was shown in the visible range of 530–535 nm for betanin while 450–559 nm for anthocyanin pigment.

Significant issues in the progress of dye solar cells are less efficient and have poor stability. The following are the various issues regarding the efficiency and stability of dye solar cell:

- Viscous nature of electrolyte used in DSSC.

Table 18 Synthetic and natural dyes

Constraint	Dyes	
	Natural dye	Synthetic dye
Ecological effects	Environmentally friendly due to natural availability and entire biodegradation (104)	Toxicity results in harmful effects on the environment. Ruthenium complexes consist of heavy metal that is dangerous to the atmosphere
Economic feasibility	These dyes extracted from natural resources are economically feasible (Oskam et al. 2001)	Nobel metal complex dyes need chemical synthesis which causes uneconomical (Nazeeruddin et al. 1999)
Methodology and performance	Simple extraction procedure makes it less costly. Siahkooti fruit extract gave an efficiency of 0.32% (Adel et al. 2015)	Complicated and costly synthesis. DSSC based on synthetic dyes reveals higher efficiency and favorable photoelectrochemical characteristics. Y123 gave an efficiency of 6.9% (Kavan et al. 2016)
Stability	Degradation of dye in the presence of sunlight radiation results poor stability	Excellent chemical stability, therefore long life of such cells (Mehmood et al. 2017)
Absorption in solar spectrum	*Erythrina variegata* showed maximum absorption at 451–492 (Hao et al. 2005)	Strong charge transfer absorption in the broad range of the visible spectrum. Porphyrin dye revealed absorption (400–650 nm) (Campbell et al. 2007)

Table 19 Performance evaluation of various DSSCs using different fabrication parameters

Dye	Performance parameter				Fabrication parameters			Ref.
	V_{oc} (V)	J_{sc} (mA/cm^2)	FF	η (%)	Semiconductor	Electrolyte	Counter electrode	
Natural dyes								
Turmeric	0.568	0.65	0.44	0.16	TiO_2	Potassium iodide and iodine	Carbon	(Ruhane et al. 2017)
Hemigraphis colorata	0.279	0.18	0.13	0.0065	TiO_2	15.9672g of 1-butyl-3-methylimidazolium iodide, 1.3385g of lithium iodide, 6.7605g of 4-tert-butyl 1 pyridine, 1.269g of iodine and 1.1816g of guanidine thiocyanate in 100 mL acetonitrile	Platinum	(Nandakumar et al. 2017)
Henna	0.40	4.9	-	1.08	TiO_2	0.8 g KI, 10 mL acetonitrile and added by 0.127 g I_2	Graphite	(Sathyajothi et al. 2017)
Synthetic dye								
$(Ru(bpy)_3)^{2+}$	0.534	0.76	0.55	0.22	TiO_2	A mixture of 0.6M 1-propyl-3-methylimidazolium iodide (Aldrich), 0.03M iodine (Aldrich), 0.1M guanidine thiocyanate and 0.5M 4-tert-butylpyridine in acetonitrile	Platinum	(Zalas et al. 2017)
$(Ru(dcbH_2)(bpy)(9AA)Cl)^{2+}$	0.58	3.54	0.46	0.95	TiO_2	$0.05molL^{-1}$ $I_2/0.5molL^{-1}LiI/0.5molL^{-1}$ pyridine in acetonitrile and 3-methyl-2-oxazolidinone (9:1 v/v)	Platinum	(Silva et al., 2016)
N3	0.71	12.59	0.64	5.65				
N719	0.69	10.83	0.45	3.38	TiO_2	0.1 M lithium iodide and 0.05 M iodine in acetonitrile	Graphene oxide	(Teo et al. 2018)
N3	0.687	1.53	0.67	2.366	TiO_2	0.1 M lithium iodide and 0.05 M iodine in acetonitrile	Platinum	(Mali et al. 2012)

- Weak performance of sensitizers in the near to infra-red region (NIR) of solar spectrum.
- Fast degradation of sensitizer resulting in less life of DSSC.
- The following recommendations can improve the efficiency and stability of dye solar cell:
- Structure of sensitizer—enhancement in dye structure to provide high efficiency in NIR region of the solar spectrum.
- Viscosity of electrolyte—electron mobility can be improved by a less viscous electrolyte that enhances the stability of the DSSC.

- Morphology of working substrate—dark current can be decreased by the proper morphology of the working substrate.
- Expensive platinum-based counter electrode can be replaced by numerous materials such as conducting polymers, carbonaceous materials, sulfides, and oxides.
- Flexible working electrode is very useful for curved surfaces. The lesser weight and low cost promote the application of flexible electrodes rather than transparent conducting oxide glasses.

Nomenclature *DSSC*: dye-sensitized solar cell; *FF*: fill factor; *FSP*: flame spray pyrolysis; I_{sc}: short circuit photocurrent (mA); P_{max}: maximum value power (mW); J_{sc}: photocurrent current density (mA/cm^2); *PCE*: power conversion efficiency (%); *PVP*: polyvinylpyrrolidone; V_{oc}: open-circuit voltage of DSSC (V); V_m: maximum value of voltage corresponds to maximum power (mV or V); η: efficiency (%)

Acknowledgements The authors are highly thankful to the Centre for Energy and Environment, Delhi Technological University, and Energy Centre, Maulana Azad National Institute of Technology, Bhopal, for providing the basic facility for this review compiling.

Author contribution Geetam Richhariya: methodology, writing (original draft), and investigation. Bhim Charan Meikap: conceptualization and supervision. Anil Kumar: conceptualization, writing (review and editing), visualization, and supervision.

Availability of data and materials Here, critical analysis has been done with reference to earlier research work. This is a kind of comprehensive review. Hence, there is no data used.

Declarations

Ethics approval and consent to participate Not applicable

Consent for publication Not applicable

Competing interests The authors declare no competing interests.

References

Adel R, Abdallah T, Moustafa YM, Al Sabagh AM, Talaat H (2015) Effect of polymer electrolyte on the performance of natural dye sensitized solar cells. SuperlatticMicrostruc 86:62–67. https://doi.org/10.1016/j.spmi.2015.07.024

Al-Alwania MAM, Hassimi AH, Al-Shorgani NKN, Al-Mashaan ABSA (2020) Natural dye extracted from Areca catechu fruits as a new sensitizer for dye-sensitised solar cell fabrication: optimisation using D-Optimal design. Mater Chem Phys 240:122204. https://doi.org/10.1016/j.matchemphys.2019.122204

Ananth S, Vivek P, Solaiyammal T, Murugakoothan P (2015) Pre dye treated titanium dioxide nano particles sensitized by natural dye extracts of Pterocarpus marsupium for dye sensitized solar cells. Optik 126:1027–1031. https://doi.org/10.1016/j.ijleo.2015.02.066

Awasthi A, Shukla AK, Shukla KN, Porwal D, Richhariya G (2020) Review on sun tracking technology in solar PV system. Energy Rep 6:392–405. https://doi.org/10.1016/j.egyr.2020.02.004

Ayalew WA, Ayele DW (2016) Dye sensitized solar cells using natural dye as light-harvesting materials extracted from Acanthus sennii chiovenda flower and Euphorbia cotinifolia leaf. J Sci: Adv Mater Device 1:488–494. https://doi.org/10.1016/j.jsamd.2016.10.003

Batniji AY, Morjan R, Abdel-Latif MS (2014) Aldimine derivatives as photosensitizers for dye-sensitized solar cells. Turk J Phys 38:86–90

Bay L, West K, JensenW B, Jacobsen T (2006) Electrochemical reaction rates in a dye-sensitised solar cell the iodide/tri-iodide redox system. Sol Energy Mater Sol Cells 90(3):341–351. https://doi.org/10.1016/j.solmat.2005.04.040

Bhattacharya A, Datta J (2020) Wide-low energy coupled semiconductor layers of TiO_2- CdX boosting the performance of DSSC. Sol Energy 208:674–687. https://doi.org/10.1016/j.solener.2020.08.024

Calogero G, Yum JH, Sinopoli A, Marco GD, Gratzel M, Nazeeruddin MK (2012) Anthocyanins and betalains as light harvesting pigments for dye sensitized solar cells. Sol Energy 86:1563–1575. https://doi.org/10.1016/j.solener.2012.02.018

Campbell WM, Jolley KW, Wagner P, Wagner K, Walsh PJ, Gordon KC (2007) Highly efficient porphyrin sensitizers for dye sensitized solar cells. J Phys Chem C 111:11760–11762. https://doi.org/10.1021/jp0750598

Cerda B, Sivakumar R, Paulraj M (2016) Natural dyes as sensitizers to increase the efficiency in sensitized solar cells. J Phys Conf Ser 720:1–5. https://doi.org/10.1088/1742-6596/720/1/012030

Chen J, Li K, Luo YA (2009) A flexible carbon counter electrode for dye-sensitized solar cells. Carbon 47:2704–2708. https://doi.org/10.1016/j.carbon.2009.05.028

Chien CY, Hsu BD (2013) Optimization of the dye sensitized solar cell with anthocyanin as photosensitizer. Sol Energy 98:203–211. https://doi.org/10.1016/j.solener.2013.09.035

Chien CY, Hsu BD (2014) Performance enhancement of dye-sensitized solar cells based on anthocyanin by carbohydrates. Sol Energy 108:403–411. https://doi.org/10.1016/j.solener.2013.09.035

Choi H, Baik C, Kang SO, Ko J, Kang MS, Nazeeruddin MK, Gratzel M (2008) Highly efficient and thermally stable organic sensitizers for solvent-free dye-sensitized solar cells. Chem, Int Ed 47(2):327–330. https://doi.org/10.1002/ange.200703852

Chowdhury FI, Buraidah MH, Arof AK, Mellander BE, Noor IM (2020) Impact of tetrabutylammonium, iodide and triiodide ions conductivity in polyacrylonitrile based electrolyte on DSSC performance. Sol Energy 196:379–388. https://doi.org/10.1016/j.solener.2019.12.033

Congiu M, Neto ON, Marco DML, Dini D, Graeff CFO (2016) Cu_2-_xS films as counter-electrodes for dye solar cells with ferrocene-based liquid electrolytes. Thin Solid Films 61:22–28. https://doi.org/10.1016/j.tsf.2016.05.033

Dadkhah M, Niasari MS, Mir N (2014) Synthesis and characterization of TiO_2 nanoparticles by using new shape controllers and its application in dye sensitized solar cells. J Ind Eng Chem 20:4039–4044. https://doi.org/10.1016/j.jiec.2014.01.003

Daeneke T, Kwon T, Holmes AB (2011) High-efficiency dye-sensitized solar cells with ferrocene-based electrolytes. Nat Chem 3:211–215. https://doi.org/10.1038/nchem.966

Di Y, Jia S, Li N, Hao C, Zhang H, Hu S, Liu H (2019) Electrocatalytic films of PEDOT incorporating transition metal phosphides as efficient counter electrodes for dye sensitized solar cells. Sol Energy 189:8–14. https://doi.org/10.1016/j.solener.2019.07.039

Esteban ACMS, Enriquez EP (2013) Graphene anthocyanin mixture as photo sensitizer for dye sensitized solar cell. Sol Energy 98:392–399. https://doi.org/10.1016/j.solener.2013.09.036

Ferreira FC, Babu RS, Barros ALF, Raja S, Conceicao LRB, Mattoso LHC (2020) Photoelectric performance evaluation of DSSCs using the dye extracted from different color petals of Leucanthemum vulgare flowers as novel sensitizers. Spectrochim Acta Part A: Mol Biomol Spectrosc 233:118198. https://doi.org/10.1016/j.saa.2020.118198

Fukurozaki SH, Zilles R, Sauer IL (2013) Energy payback time and CO_2 emissions of 1.2 kWp photovoltaic roof top system in Brazil. Int J Smart Grid Clean Energy 2:164–169. https://doi.org/10.12720/sgce.2.2.164-169

Godibo DJ, Anshebo ST, Anshebo TY (2015) Dye sensitized solar cells using natural pigments from five plants and quasisolid state electrolyte. J Braz Chem Soc 26:92–101. https://doi.org/10.5935/0103-5053.20140218

Gorlov M, Kloo L (2008) Ionic liquid electrolytes for dye-sensitized solar cells. Dalton Trans 20:2655–2666. https://doi.org/10.1039/B716419J

Gratzel M (2003) Review: dye-sensitized solar cells. J PhotochemPhotobiol C: Photochem Rev 4:145–153. https://doi.org/10.1016/S1389-5567(03)00026-1

Hagfeldt A, Boschoo G, Sun L (2010) Dye sensitized solar cells. Chem Rev 110:6595–6663. https://doi.org/10.1021/cr900356p

Hamadanian M, Ghomi JS, Hosseinpour M, Masoomi R, Jabbari V (2014) Uses of new natural dye photo sensitizers in fabrication of high potential dye-sensitized solar cells (DSSCs). Mater Sci Semicond Process 27:733–739. https://doi.org/10.1016/j.mssp.2014.08.017

Han HG, Weerasinghe HC, Kim KM, Kim JS (2015) Ultrafast Fabrication of flexible dye sensitized solar cells by ultrasonic spray coating technology. Mater Sci Eng 5:1–9. https://doi.org/10.1038/srep14645

Hao S, Wu J, Huang Y, Lin J (2005) Natural dyes as photo sensitizers for dye sensitized solar cell. Sol Energy 80:209–214. https://doi.org/10.1016/j.solener.2005.05.009

Hashmi G, Miettunen K, Peltola T, Halme J (2011) Review of materials and manufacturing options for large area flexible dye solar cells. Renew Sust Energ Rev 15:3717–3732. https://doi.org/10.1016/j.rser.2011.06.004

Heo JH, Jung KY, Kwak DJ (2009) Fabrication of titanium doped indium oxide films for dye-sensitized solar cell application using reactive RF magnetron sputter method. IEEE Trans Plasma Sci 37:1586–1592. https://doi.org/10.1109/TPS.2009.2023477

Hu B, Tang Q, He B, Lin L, Chen H (2014) Mesoporous TiO_2 anodes for efficient dye-sensitized solar cells: an efficiency of 9.86% under one sun illumination. Power Sourc 267:445–451. https://doi.org/10.1016/j.jpowsour.2014.05.119

Hug H, Bader M, Mair P, Glatzel T (2013) Biophotovoltaic natural pigments in dye sensitized solar cells. Appl Energy 115:216–225. https://doi.org/10.1016/j.apenergy.2013.10.055

Huo J, Hu Y, Jiang H, Hou X (2014) Continuous flame synthesis of near surface nitrogen doped TiO_2 for dye sensitized solar cells. J Chem Eng 258:163–170. https://doi.org/10.1016/j.cej.2014.07.026

Ito S and Takahashi K (2012). Fabrication of monolithic dye sensitized solar cell using ionic liquid electrolyte. Int J Photoenergy 6. https://doi.org/10.1155/2012/915352

Jalali T, Arkian P, Golshan M, Jalali M, Osfouri S (2020) Performance evaluation of natural native dyes as photosensitizer in dye-sensitized solar cells. Opt Mater 110:110441. https://doi.org/10.1016/j.optmat.2020.110441

Jia HL, Chen YC, Ji L, Lin XL, Guana MY, Yang Y (2018) Cosensitization of porphyrin dyes with new X type organic dyes for e-cient dye-sensitized solar cells. Dyes Pigments 163:589–593. https://doi.org/10.1016/j.dyepig.2018.12.048

Jun Y, Kim J, Kang MG (2007) A study of stainless steel based dye-sensitized solar cells and modules. Sol Energy Mater Sol Cells 91:779–784. https://doi.org/10.1016/j.solmat.2007.01.007

Karpacheva M, Malzner FJ, Wobill C, Buttner A, Constable EC, Housecroft, Constable CE (2018) Dye-sensitized solar cells with copper(I) dyes and copper(I)/(II) redox shuttles. Dyes Pigments 156:410–416. https://doi.org/10.1016/j.dyepig.2018.04.033

Kavan L, Liska P, Zakeeruddin SM, Michael G (2016) Low temperature fabrication of highly efficient, optically transparent (FTO-free) graphene cathode for co-mediated dye-sensitized solar cells with acetonitrile free electrolyte solution. Electrochim Acta 195:34–42. https://doi.org/10.1016/j.electacta.2016.02.097

Kay A, Gratzel M (1996) Low cost photovoltaic modules based on dye sensitized nanocrystalline titanium dioxide and carbon powder. Sol Energy Mater Sol Cells 44:99–117. https://doi.org/10.1016/0927-0248(96)00063-3

Kumar R, Bharg P (2015) Fabrication of a counter electrode using glucose as carbon material for dye sensitized solar cells. Mater Sci Semicond Process 40:331–336. https://doi.org/10.1016/j.mssp.2015.06.009

Kumar A., Richhariya G., Sharma A.2015. Solar photovoltaic technology and its sustainability. Energy Sustain through Green Energy, edited by Atul Sharma, Sanjay Kumar Kar. Chapter 1. 3-25. Springer. ISBN: 978-81-322-2336-8. https://doi.org/10.1007/978-81-322-2337-5_1

Kumar DK, Kriz J, Bennett N, Chen B, Upadhayaya H, Reddy KR, Sadhu V (2020) Functionalized metal oxide nanoparticles for efficient dye-sensitized solar cells (DSSCs): a review. Mater Sci Energy Technol 3:472–481. https://doi.org/10.1016/j.mset.2020.03.003

Kumara NTRN, Ekanayake P, Lim A (2013) A layered co-sensitization for enhancement of conversion efficiency of natural dye sensitized solar cells. J Alloys Compd 581:186–191. https://doi.org/10.1016/j.jallcom.2013.07.039

Kumara NTRN, Lim A, Lim CM, Petra MI, Ekanayake P (2017) Recent progress and utilization of natural pigments in dye sensitized solar cells: a review. Renew Sust Energ Rev 78:301–317. https://doi.org/10.1016/j.rser.2017.04.075

Latif MSA, Abuiriban MB, El-Agez TM, Taya SA (2015) Dye sensitized solar cells using dyes extracted from flowers, leaves, parks, and roots of three trees. Int J Renew Energy Res 5:294–298

Latini A, Aldibaja FK, Cavallo C, Gozzi D (2014) Benzonitrile based electrolytes for best operation of dye sensitized solar cells. J Power Sources 269:308–316. https://doi.org/10.1016/j.jpowsour.2014.06.154

Law M, Greene LE, Johnson JC, Saykally R, Yang P (2005) Nano wire dye sensitized solar cell. Nat Mater 4:455–459. https://doi.org/10.1038/nmat1387

Lee WJ, Lee DY, Song JS, Min BK (2005) Effect of process parameters on the efficiency of dye sensitized solar cells. Met Mater Int 11:465–471. https://doi.org/10.1007/bf03027496

Lim A, Kumara NTRN, Tan AL, Mirza AH, Chandrakanthi RLN, Petra MI (2015) Potential natural sensitizers extracted from the skin of Canarium odontophyllum fruits for dye-sensitized solar cells. Spectroc Acta Part A: Mol BiomolSpectros 13:596–602. https://doi.org/10.1016/j.saa.2014.11.102

Lo KSK, Leung WWF (2019) Dye-sensitized solar cells with shear-exfoliated graphene. Sol Energy 180:16–24. https://doi.org/10.1016/j.solener.2018.12.077

Ludin NA, Al-Alwani Mahmoud AM, Mohamad AB, Kadhum AAH, Sopian K, Abdul Karim NS (2014) Review on the development of natural dye photo sensitizer for dye sensitized solar cells. Renew Sust Energ Rev 31:386–396. https://doi.org/10.1016/j.rser.2013.12.001

Mali SS, Betty CA, Bhosale PN, Patil PS (2012) Eosin-Y and N3-dye sensitized solar cells (DSSCs) based on novel nanocoral TiO_2: a comparative study. Electrochim Acta 59:113–120. https://doi.org/10.1016/j.electacta.2011.10.043

Maurya IC, Neetu, Gupta AK, Srivastava P, Bahadur L (2016) Callindrahaematocephata C and Peltophorum pterocarpum flowers as natural sensitizers for TiO_2 thin film based dye-sensitized solar cells. Opt Mater 60:270–276. https://doi.org/10.1016/j.optmat.2016.07.041

Mbonyiryivuze A, Omollo I, Ngom BD, Mwakikunga, Dhlamin SM, Park E, Maaza M (2015) Natural dye sensitizer for gratzel cells: Sepia melanin. Phys Mater Chem 3:1–6. https://doi.org/10.12691/pmc-3-1-1

Mehmood U, Rahman S, Harrabi K, Hussein IA, Reddy BVS (2014). Review article: recent advances in dye sensitized solar cells. Adv Mater Sci Eng. 1-13. https://doi.org/10.1155/2014/974782

Mehmood U, Al-Ahmed A, Al-Sulaiman FA, Malik MI, Shehzad F, Khan AH (2017) E-ect of temperature on the photovoltaic performance and stability of solid state dye sensitized solar cells: a review. Renew Sust Energ Rev 79:946–959. https://doi.org/10.1016/j.rser.2017.05.114

Moustafa S, Abusaif A, Fathy M, Abu-Saied MA, Elhenawy AA, Kashyout AB, Mohamed R, Yousry S, Ammara A (2021) New carbazole-based organic dyes with different acceptors for dye-sensitized solar cells: Synthesis, characterization, DSSC fabrications and density functional theory studies. J Mol Struct 1225:129297. https://doi.org/10.1016/j.molstruc.2020.129297

Muniz EC, Goes MS, Silv JJ, Varela JA, Joanni E (2011) Synthesis and characterization of mesoporous TiO_2 nanostructured films prepared by a modified sol–gel method for application in dye solar cells. Ceram Int 37:1017–1024. https://doi.org/10.1016/j.ceramint.2010.11.014

Nagavolu C, Susmitha K, Raghavender M, Giribabu L, Mills CA, Silva SRP (2016) Pt-free spray coated reduced graphene oxide counter electrodes for dye sensitized solar cells. Sol Energy 137:143–147. https://doi.org/10.1016/j.solener.2016.08.002

Nan H, Shen HP, Wang G, Xie SD, Yang GJ, Lin H (2017) Studies on the optical and photoelectric properties of anthocyanin and chlorophyll as natural co-sensitizers in dye sensitized solar cell. Opt Mater 73(2017):172–178. https://doi.org/10.1016/j.optmat.2017.07.036

Nandakumar VG, Suresh S, Sreekala CO, Sudheer SK, Pillai MVP (2017) Hemigraphiscolorata as a natural dye for solar energy conversion. Mater Today: Proc 4:4358–4365. https://doi.org/10.1016/j.matpr.2017.04.006

Narayan MR (2012) Review: dye sensitized solar cells based on natural photosensitizers. Renew Sust Energ Rev 16:208–215. https://doi.org/10.1016/j.rser.2011.07.148

Nazeeruddin MK, Kay A, Rodicio I (1993) Conversion of light to electricity by cis-X2bis(2,2'-bipyridyl-4,4'-dicarboxylate) ruthenium(II) charge-transfer sensitizers (X =Cl-, Br-,I-,CN-, and SCN-) on nanocrystalline titanium dioxide electrodes. J Am Chem Soc 115:6382–6390. https://doi.org/10.1021/ja00067a063

Nazeeruddin MK, Zakeeruddin SM, Baker HR (1999) Acid-base equilibria of (2,2'-Bipyridyl-4,4'-dicarboxylic acid)ruthenium(II) complexes and the effect of protonation on charge transfer sensitization of nanocrystalline titania. Inorg.Chem. 38:6298–6305. https://doi.org/10.1021/ic990916a

Obi K, Frolova L, Fuierer P (2020) Preparation and performance of prickly pear (Opuntia phaeacantha) and mulberry (Morus rubra) dye-sensitized solar cells. Sol Energy 208:312–320. https://doi.org/10.1016/j.solener.2020.08.006

Oskam G, Bergeron BV, Meyer GJ (2001) Pseudohalogens for dye sensitized TiO_2 photoelectrochemical cells. J Phys Chem B 105:6867–6873. https://doi.org/10.1021/jp004411d

Ozuomba JO, Okoli LU, Ekpunobi AJ (2013) The performance and stability of anthocyanin local dye as a photosensitizer for DSSCs. Adv Appl Sci Res 4:60–69 ISSN: 0976-8610

Park KH, Kim TY, Park JY, Jin EM, Yim SH, Choi DY, Lee JW (2012) Adsorption characteristics of gardenia yellow as natural photo sensitizer for dye sensitized solar cells. Dyes Pigments 96:595–601. https://doi.org/10.1016/j.dyepig.2012.10.005

Pavithra N, Asiri AM, Anandan SF (2015) Fabrication of dye sensitized solar cell using gel polymer electrolytes consisting poly(ethylene oxide)-acetamide composite. J Power Sources 286:346–353. https://doi.org/10.1016/j.jpowsour.2015.03.160

Peng W, Zeng Y, Gong H, Leng YQ, Yan YH, Hu W (2013) Silver coated TiO_2 electrodes for high performance dye-sensitized solar cells. Solid State Electron 89:116–119. https://doi.org/10.1016/j.sse.2013.07.011

Pettersson H, Gruszecki T, Bernhard R, Haggman L, Gorlov M, Boschloo G, Edvinsson T, Kloo L, Hagfeldt A (2007) The monolithic multicell: a tool for testing material components in dye-sensitized solar cells. Prog Photovolt 15(2):113–121. https://doi.org/10.1002/pip.713

Phinjaturus K, Maiaugree W, Suriharn B, Pimanpaeng S, Amornkitbamrung V, Swatsitang E (2016) Dye sensitized solar cells based on purple corn sensitizers. Appl Surf Sci 380:101–107. https://doi.org/10.1016/j.apsusc.2016.02.050

Phonkhokkong T, Thongtem T, Thongtem S, Phuruangrat A, Promnopas W (2016) Synthesis and characterization of TiO_2 nanopowders for fabrication of dye sensitized solar cells. Digest J Nanomater Biostruct 1:81–90 ISSN: 1842-3582

Polo AS, Itokazu MK, Iha NYM (2004) Metal complex sensitizers in dye-sensitized solar cells. Coord Chem Rev 248:1343–1361. https://doi.org/10.1016/j.ccr.2004.04.013

Ramarajan R, Purushothamreddy N, Dileep RK, Kovendhan M, Veerappan G, Thangaraju K, Joseph DP (2020) Large-area spray deposited Ta-doped SnO_2 thin film electrode for DSSC application. Sol Energy 211:547–559. https://doi.org/10.1016/j.solener.2020.09.042

Rawal N, Vaishaly AG, Sharma H, Mathew B (2015) Dye sensitized solar cells: the emerging technology. Energy Power Eng Sci 2:46–52. https://doi.org/10.12966/epes.05.03.2015

Researcher J.L.E, Hunter R, Rubilar M, Pavez B, Morales E, Torres S (2016) Development of dye sensitized solar cells based on naturally extracted dye from the maqui berry (Aristotelia chilensis). Opt Mater 60:411–417. https://doi.org/10.1016/j.optmat.2016.08.021

Richhariya G, Kumar A (2016) A Review on performance affected parameters for dye sensitized solar cell. Energy Secur Sustain, CRC Press, Boca Raton, pp 93–107 Print ISBN: 978-1-4987-5443-9, eBook ISBN: 978-1-4987-5444-6

Ruhane TA, Islam MT, Rahaman MS, Bhuiyan MMH, Islam JMM, Newaz MK, Khan KA, Khan MA (2017) Photo current enhancement of natural dye sensitized solar cell by optimizing dye extraction and its loading period. Optik 149:174–183. https://doi.org/10.1016/j.ijleo.2017.09.024

Saito Y, Kitamura T, Wada Y, Yanagida S (2002) Application of poly (3,4-ethylenedioxythiophene) to counter electrode in dye-sensitized solar cells. Chem Lett 10:1060–1061. https://doi.org/10.1021/jp303958r

Sathyajothi S, Jayavel R, Dhanemozhi AC (2017) The fabrication of natural dye sensitized solar cell (DSSC) based on TiO_2 using henna and beetroot dye extracts. Mater Today: Pro 4:668–676. https://doi.org/10.1016/j.matpr.2017.01.071

Sedghi A, Miankushki HN (2013) Influence of $TiCl_4$ treatment on structure and performance of dye sensitized solar cells. Jpn J Appl Phys 52:1–5. https://doi.org/10.7567/JJAP.52.075002

Shahid MU, Mohamed NM, Muhsan AS, Bashiri R, Shamsudin AE, Zaine SNA (2019) Few-layer grapheme supported polyaniline (PANI) film as a transparent counter electrode for dye-sensitized solar cells. Diamond Relat Mater 94:242–251. https://doi.org/10.1016/j.diamond.2019.03.009

Sharma GD, Zervaki GE, Angaridis PA, Vatikioti A, Gupta KSV, Gayathri T, Nagarjuna P (2014) Stepwise co-sensitization as a useful tool for enhancement of power conversion efficiency of dye sensitized solar cells: the case of an unsymmetrical porphyrin dyad and a metal free organic dye. Org Electron 15:1324–1337. https://doi.org/10.1016/j.orgel.2014.03.033

Sharma A, Srivastava J, Kumar A (2015) A comprehensive overview of renewable energy status in India. Environ Sustain 2015:91–105. https://doi.org/10.1007/978-81-322-2056-5_5

Sharma K, Sharma V, Sharma SS (2020) Dye sensitized solar cells: fundamentals and current status. Nanoscale Res Lett 13:381. https://doi.org/10.1186/s11671-018-2760-6

Silva MDSDP, Diogenes ICN, Carvalho IMMD, Zanoni KPS, Amaral RCM, Iha NY (2016) Novel heteroleptic ruthenium complexes for dye sensitized solar cells. J Photochemphtobiol A:Chem 314:75–80. https://doi.org/10.1016/j.jphotochem.2015.08.012

Silva LD, Sanchez M, Freeman HS (2020) New tetrazole based dyes as efficient co-sensitizers for DSSCs: Structure-properties relationship. Org Electron 87:105964. https://doi.org/10.1016/j.orgel.2020.105964

Singh R, Bhattacharya B, Rhee HW, Singh PK (2014) New biodegradable polymer electrolyte for dye sensitized solar cell. Int J Electrochem Sci 9:2620–2630 ISSN 1452-3981

Susanti D, Nafi M, Purwaningsih H, Fajarin R, Kusuma GE (2014) The preparation of dye sensitized solar cell (DSSC) from TiO_2 and tamarillo extract. Proc Chem 9:3–10. https://doi.org/10.1016/j.proche.2014.05.002

Takada H, Obana Y, Sasaki R, Kuribayashi M, Kanno M, Zhu C (2015) Improved durability of dye sensitized solar cell with H_2 reduced carbon counter electrode. J Power Sources 274:1276–1282. https://doi.org/10.1016/j.jpowsour.2014.10.148

Takagi K, Magaino S, Saito H, Aoki T, Aoki D (2013) Measurements and evaluation of dye-sensitized solar cell performance. J PhotochemPhotobiol C: Photochem Rev 14:1–12. https://doi.org/10.1016/j.jphotochemrev.2012.08.003

Taya SA, El-Agez TM, El-Ghamri HS, Latif MSA (2013) Dye sensitized solar cells using fresh and dried natural dyes. Int J Mater Sci Appl 2:37–42. https://doi.org/10.11648/j.ijmsa.20130202.11

Teo LP, Tiong TS, Buraidah MH, Arof AK (2018) Effect of lithium iodide on the performance of dye sensitized solar cells (DSSC) using poly(ethylene oxide) (PEO)/poly (vinyl alcohol) (PVA) based gel polymer electrolytes. Opt Mater 85:531–537. https://doi.org/10.1016/j.optmat.2018.09.026

Tian H, Jiang X, Yu Z, Kloo L, Hagfeldt A, Sun L, Angew. (2000) Efficient organic dye sensitized solar cells based on an iodine free electrolyte. Chem Int Ed 49(40):7328–7331. https://doi.org/10.1002/anie.201003740

Tontapha S, Aroon WS, Promgool T, Kanokmedhakul S, Maiaugree W, Swatsitang E, Homrahad V, Amornkitbumrung V (2019) Electrocatalytic activity of disulfide/thiolate with graphene nanosheets as an efficient counter electrode for DSSCs: a DFT study. Mater Today Commun 22:100742. https://doi.org/10.1016/j.mtcomm.2019.100742

Torabi N, Behjat A, Jafari F (2014) Dye sensitized solar cells based on porous conjugated polymer counter electrodes. Thin Solid Films 573:112–116. https://doi.org/10.1016/j.tsf.2014.11.034

Uddin J, Islam JMM, Karim E, Khan SMM, Akhter S (2015) Preparation and characterization of dye sensitized solar cell using natural dye extract from Red Amaranth (Amaranthus sp.) as sensitizer. Int J Thin Film Sci Tech 4:41–146. https://doi.org/10.12785/ijtfst/040212

Wang ZS, Sayama K, Sugihara H (2005) Efficient eosin y dye-sensitized solar cell containing Br^-/Br^{3-}electrolyte. J Phys Chem B 109:22449–22455. https://doi.org/10.1021/jp053260h

Wang M, Chamberland N, Breau L, Moser JE, Baker HR (2010) An organic redox electrolyte to rival triiodide/iodide in dye-sensitized solar cells. Nat Chem 2(5):385–389. 32. https://doi.org/10.1038/nchem.610

Wang H, Gao J, Zhu J, Ma JY, Zhou H, Xiao J, Wu M (2020) Design bifunctional nitrogen doped flexible carbon sphere electrode for dye-sensitized solar cell and supercapacitor. Electrochim Acta 334:135–582. https://doi.org/10.1016/j.electacta.2019.135582

Weerasinghe HC, Huang F, Cheng YB (2013) Fabrication of flexible dye sensitized solar cells on plastic substrates. Nano Energy 2:174–189. https://doi.org/10.1016/j.nanoen.2012.10.004

Wu J, Lan Z, Hao S, Li P, Lin J, Huang M, Fang L, Huan Y (2008) Progress on the electrolytes for dye-sensitized solar cells. Pure Appl Chem 80:2241–2258. https://doi.org/10.1351/pac200880112241)

Yang X, Luo J, Zhou L, Yang B, Zuo X, Li G (2014) A novel Pt-free counter electrode for dye-sensitized solar cells: nickel sulfide hollow spheres. Chem Mater Let 136:241–244. https://doi.org/10.1016/j.matlet.2014.08.053

Younas M, Gondal MA, Dastageer MA, Harrabi K (2019a) E-cient and cost-e-ective dye-sensitized solar cells using MWCNT-TiO_2 nanocomposite as photoanode and MWCNT as Pt-free counter electrode. Sol Energy 188:1178–1188. https://doi.org/10.1016/j.solener.2019.07.009

Younas M, Gondal MA, Dastageer MA, Baig U (2019b) Fabrication of cost effective and efficient dye sensitized solar cells with WO_3-TiO_2 nanocomposites as photoanode and MWCNT as Pt-free counter electrode. Ceram Int 45:936–947. https://doi.org/10.1016/j.ceramint.2018.09.269

Yugis AR, Mansa RF, Sipaut CS (2015) Review on metallic and plastic flexible dye sensitized solar cell. Mater Sci Eng 78:1–7. https://doi.org/10.1088/1757-899X/78/1/012003

Yune JH, Karatchevtseva I, Triania G, Wagner K, Officer DJ (2013) A study of TiO_2 binder-free paste prepared for low temperature dye-sensitized solar cells. Mater Res 28:488–496. https://doi.org/10.1557/jmr.2012.354

Zalas M, Gierczyk B, Bossi A, Mussini PR, Klein M, Pankiewicz R (2017) The influence of anchoring group position in ruthenium dye molecule on performance of dye-sensitized solar cells. Dyes Pigments 150:335–346. https://doi.org/10.1016/j.dyepig.2017.12.029

Zeng W, Cao Y, Bai Y, Wang Y, Shi Y, Zhang M, Wang F, Pan C, Wang P (2010) Efficient dye sensitized solar cells with an organic photosensitizer featuring orderly conjugated ethylene dioxythiophene and dithienosiloleblocks. Chem Mater 22:1915–1925. https://doi.org/10.1021/cm9036988

Zhou W, Yang W, Fang Z (2007) A novel model for photovoltaic array performance prediction. Appl Energy 84:1187–1198. https://doi.org/10.1016/j.apenergy.2007.04.006

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:03/Issue:11/November-2021 Impact Factor- 6.752 www.irjmets.com

REVIEW RESEARCH PAPER NEUROMORPHIC COMPUTING AND APPLICATIONS

Hitesh Dureja*1, Yash Garg*2, Dr. Rishu Chaujar*3, Bhavya Kumar*4

*1UG Student (2K20/EP/055, Dept. Of Applied Physics, Delhi Technological University)

*2UG Student (2K20/EP/121, Dept. Of Applied Physics, Delhi Technological University)

*3Professor (Delhi Technological University)

*4Ph.D. Student (Delhi Technological University)

ABSTRACT

A neuro-inspired computing chip emulates the structure and operation of the biological brain and represents an innovative and assured approach to the development of intelligent computing[1]. When processing Artificial Intelligence workloads, these neuro-inspired computing chips are expected to provide advantages in power efficiency and computing power over traditional systems[3]. Over the past few years, a spread of neuro-inspired computing chips have been developed[11]. Several neuro-inspired innovations have been incorporated into these chips at various levels, from the hardware to the circuitry to the architecture[7]. This is still at a beginning stage in the development of neuro-inspired computing chips, so exploring the hurdles and opportunities for the field is crucial[10].

The origins of neuro-inspired computing chips and recent progress in the domain is studied by us[5]. We classify four critical metrics for deciding the performance of the fragments: computing density, energy efficiency, computing accuracy, and learning capability[8]. We then explore the challenges and co-design principles of developing large-scale chips based on non-volatile memory (NVM)[2]. We also address the future electronic design automation (EDA) toolchain and propose a technological roadmap to develop large-scale neuro-inspired computing chips[1].

I. INTRODUCTION

Human brains are capable of processing extensive amounts of information while consuming little energy. In response to a need, the brain turns up computation, but it immediately returns to a baseline. As far as silicon-based computers are concerned, such efficiencies have never been achieved. Massive amounts of electricity are required to process huge amounts of data[4].

Deep neural networks (DNNs) and chips have achieved considerable improvements in accuracy on a variety of large-scale classification tasks, some even surpassing human performance[6]. Nevertheless, in order to achieve better training efficiency, the DNNs model's parameters rise exponentially, resulting in hundreds of millions of parameters and large training datasets stored in the memory. In the traditional, von Neumann-based computer architecture, the data needs to be moved back and forth between memory, and the processor, resulting in limited hardware energy efficiency for these machine learning workloads[14,10,8].

The concept of adaptive parallel processing in biological neural networks (BioNNs) using neuro-inspired computing is proposed to eliminate the energy-intensive and inefficient transmission of von Neumann-based platforms.

The two main paradigms of neuro-inspired computing are digital-bit-encoded artificial neural networks (ANNs) and spike-timing-encoded spiking neural networks (SNNs). Deep neural networks (DNN) and convolutional neural networks (CNNs), as well as recurrent neural networks (RNNs), have been successful in logical computing, machine vision, intelligent search, and automatic driving.[2,3]

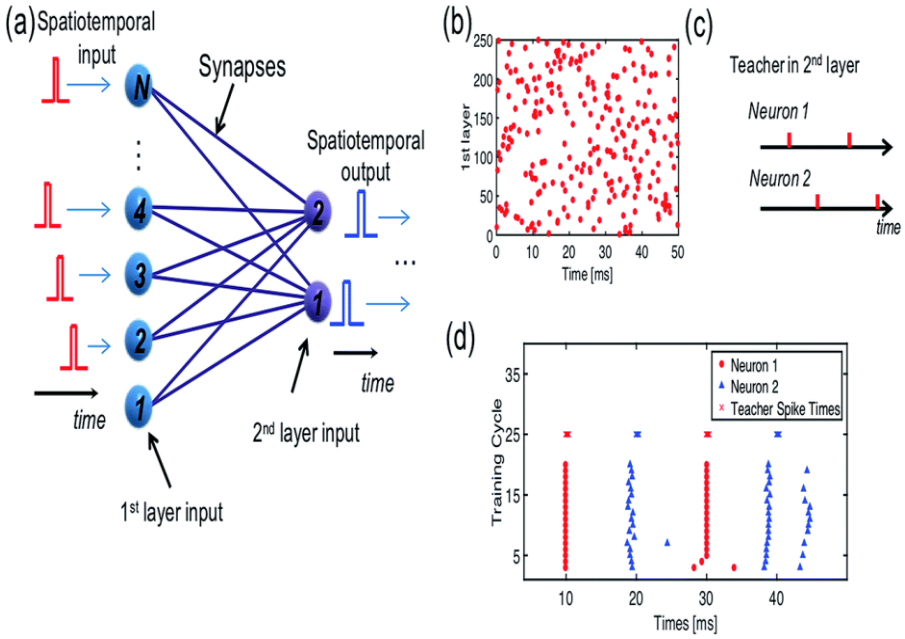
Algorithms of the brain. Human brains are dynamic, reconfigurable systems composed of neurons that are interconnected through synaptic connections[12]. For physicists, it exhibits many fascinating phenomena such as energy minimization or entropy minimization, phase transitions, criticality, self-oscillation chaos, synchronization, stochastic resonance, and many more[1]. Since the beginning, physicists have participated in theoretical efforts to explain the brain's algorithms, making contributions to both the fields of computer science and computational neuroscience. Statistical physics, nonlinear dynamics, and complex systems theory helped

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:03/Issue:11/November-2021 Impact Factor- 6.752 www.irjmets.com

cast light on neural mechanisms that allow learning. Hopfield networks and Boltzmann machines, which inherit from Ising spin systems, are the most famous ones, but many others have been proposed, exploiting, in particular, nonlinear dynamics for computing[3,8,10].

II. PHYSICS FOR NEUROMORPHIC COMPUTING

Neuromorphic computing influences the brain to create energy-efficient hardware for information processing, capable of highly sophisticated tasks. Systems built with regular electronics obtain gains in pace and energy by mimicking the dispersed topology of the brain. Scaling up such arrangements and improving their energy usage, performance, and speed by several orders of magnitude requires a revolution of hardware. Neuromorphic computing could be greatly enhanced by incorporating more physics in algorithms and nanoscale materials. We review impressive results that leverage physics to improve the computing capabilities of artificial neural networks, using resistive switching elements, photonics, spintronics, and other technologies. We study the pathways that could affect these approaches to maturity towards low-power, miniaturized chips that could infer and acquire in real-time.



The performance conditions of NVM devices for neuro-inspired computing chips are largely dependent on particular systems and applications. The number of analog states determines the weight tuning precision. It has been reported that precision of at least eight equivalent bits is required for training a relatively big neural network45, such as ResNet46.

Current electronics are not enough. In the brain, neurons which can roughly be viewed as performing processing -- have straight access to memory, supported by synapses. Current electronics, on the contrary, intrinsically separate memory and computing into discrete physical units, between which data must be transferred back and forth. This "von Neumann bottleneck" is an issue for artificial intelligence algorithms, which require reading substantial amounts of data at every step, performing complex operations on this data, and then writing the results back to memory. It slows down computing and considerably enhances the energy loss for learning and inference.

The standard model in neuromorphic computing is, therefore, to take inspiration from the topology of the brain to build circuits formed of physical neurons interconnected by physical synapses that perform memory in-situ, in a non-volatile way, thus drastically cutting the need to move data around the circuit and providing huge gains in speed and energy efficiency. This is unfortunately complicated by using Complementary Metal Oxide Semiconductor (CMOS) technology alone. Dozens of transistors are needed to imitate each neuron, and additional outside memories are required to execute synapses. CMOS-based artificial neurons and synapses are typically several micrometers wide. The number of physical neurons and synapses that can be blended into a CMOS chip is inherently limited by the chip area. This is problematic because the performance of neural

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:03/Issue:11/November-2021 Impact Factor- 6.752 www.irjmets.com

networks increases with the number of neurons and synapses: ideal image recognition algorithms today involve millions of neurons and synapses on an average. Large numbers of neurons and synapses can be achieved by gathering chips. In addition, the whole system becomes heavy, and much energy is wasted in the interconnects. Nanodevices that can imitate important features of neurons and synapses at the nanoscale, such as non-linearity, memory, and learning, are required to build low-power chips comprising several millions of neurons and synapses.

Finally, it is difficult to achieve a high degree of interconnection between neurons using CMOS technology only. The brain highlights an average of 10,000 synapses per neuron. Such connectivity is impossible to reproduce with current electronics. CMOS technology is mostly confined to two dimensions (2D), fanout is limited, and it is difficult to efficiently and fairly supply energy to components in the circuit. On the contrary, the brain is tri-dimensional (3D), neuron axons and dendrites provide high fan-in/fan-out, and blood efficiently distributes energy to the entire system.

III. SPIKING NEURAL UNIT (SNN)

In spiking neurons, the input is filtered in some way, usually a low pass filter, and they fire when a state variable exceeds a threshold. The spike train is calculated using Dirac delta functions where tk is the spike time. We wish to express learning in an SNN as minimization of a loss function across an extensive number of training samples, similar to old machine learning. A learning process includes finding sets of synaptic weights that allow scattered representation to be performed as well as decreasing the sum of all scattered coding losses in the sparse coding case. Learning in an SNN naturally proceeds online, where training samples are sent to the network sequentially.

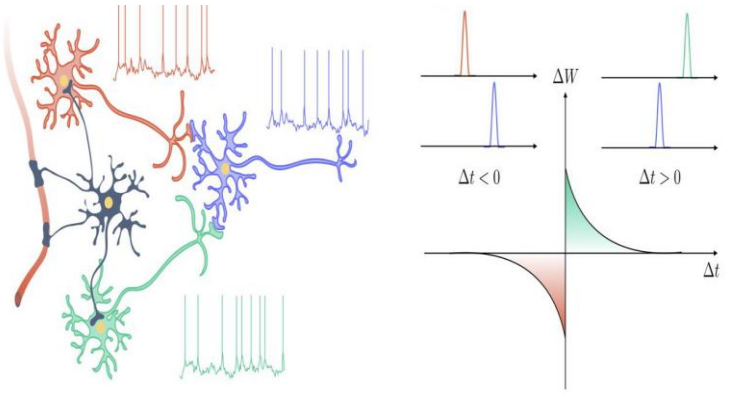
$$\dot{v}_i(t) = -\frac{1}{\tau_v} v_i(t) + u_i(t) - \theta_i \sigma_i(t)$$

$$u_{i(t)} = \sum_{j \neq i} w_{i,j} (\alpha_u * \sigma_j)(t) + b_i$$

IV. NEUROSCIENCE CONCEPTS AND HARDWARE

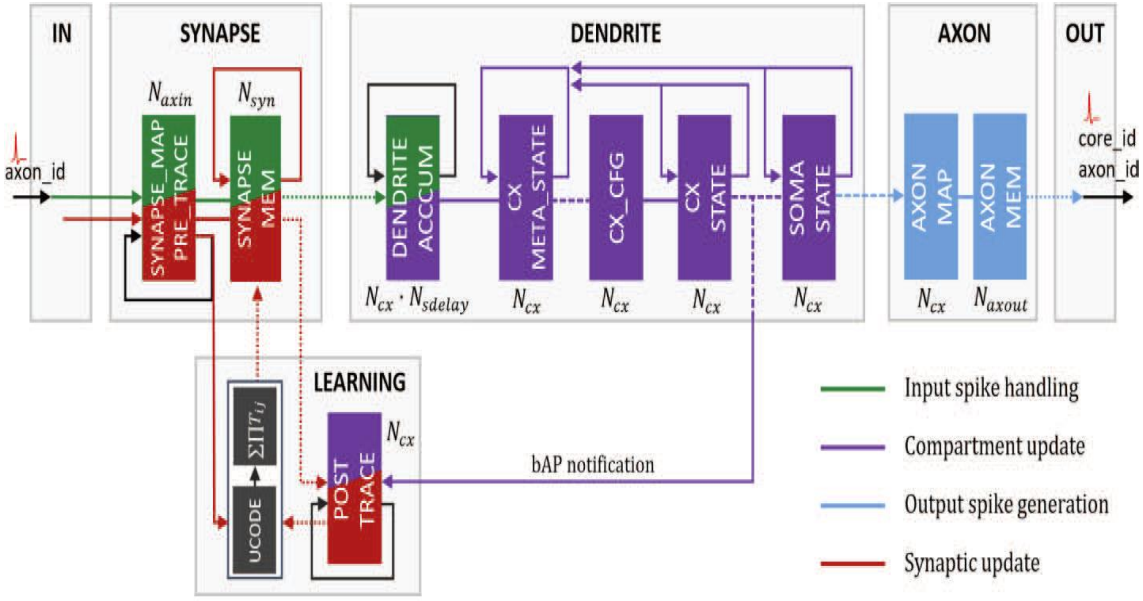
Artificial Intelligence algorithms cannot compete with the complexity of the brain. Biological neurons are more than nonlinear functions. A spike is leaky, feature memory is stochastic, and it can oscillate and synchronize. Several functional slots integrate signals coming from different areas over a space that is spatially extended.

Biological synapses are more than analog weights. Some synapses transmit only a fraction of the spikes they receive, so they can be extremely stochastic. Often overlooked brain components play a critical role and are important to imitate: dendrites appear to be able to execute very complex computations, whereas astrocytes are involved in neuronal regulation. Artificial neural networks can benefit from all these attributes and thus are more appealing to implement.



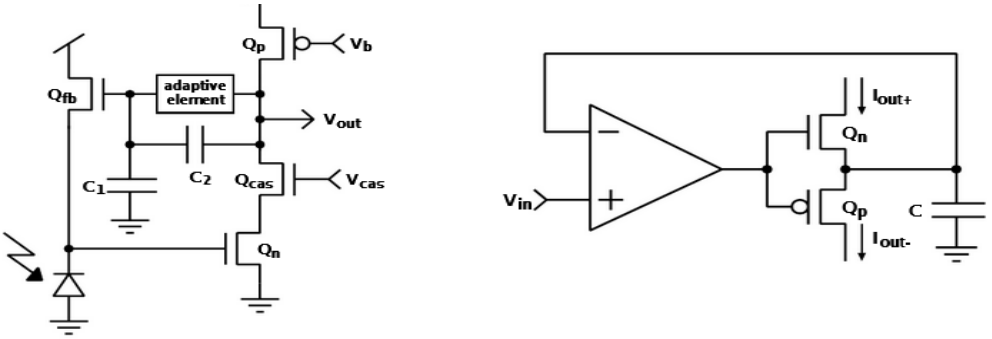
Through the related physical effects and intrinsic qualities of materials, all of these ideas have great potential if they can be realized at very low energy. The purpose of this research field was to exploit the exponential dependence of transistor leakage current on voltage. In the last decade, a huge variety of other physical phenomena, depicted, have been used to imitate interesting qualities of synapses and neurons.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:03/Issue:11/November-2021 Impact Factor- 6.752 www.irjmets.com



V. LOGIC GATES AND APPLICATIONS

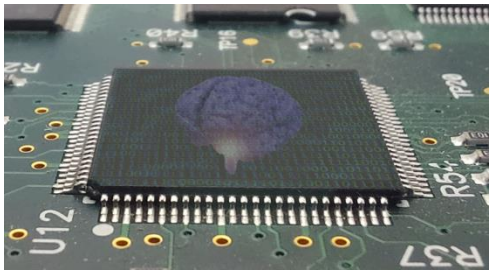
A human brain performs synaptic operations every second. Therefore, it has an efficiency of about 3x1014 operations per joule. Despite the slow and noisy parts in our heads and bodies, the human brain can perform complex computations in real-time. In comparison with the computing efficiency of a digital microprocessor, it can be observed that the brain is no less than seven orders of magnitude more systematic . Additionally, neural networks, based on the real brain, have a great deal of connectivity, which is expected to contribute to the high efficiency of biological systems due to their fault tolerance. The neuron achieves the equivalent of a logical OR operation on the excitatory inputs. By interpreting pulse behavior as a logical value of 1, we can realize how the OR gate functions by using neurons with two excitatory inputs and the output feedback as a piece of inhibitory information. If the excitation ceases, the neuron returns to its comfortable state, which corresponds to a logical value of '0'.



VI. NEURO-INSPIRED CHIP

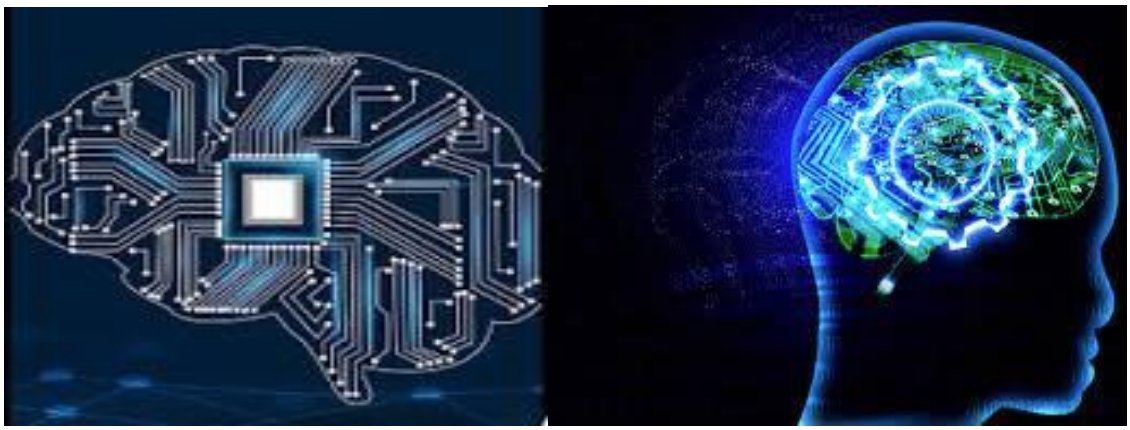
Artificial synapses and neurons are bound to neuromorphic chips, which simulate the action spikes within the human brain. These chips handle all of this processing on their own. This appears in smarter, far more energy-efficient computing systems. An energy-efficient approach to AI computing workloads could be provided by neuro-inspired computing chips combine features derived from neural systems. We analyze four key metrics for benchmarking neuro-inspired computing chips, including computing density, energy efficiency, computing accuracy, and on-chip learning capability, and discuss co-design principles, from the device to the algorithm level, for neuro-based computing chips using non-volatile memory.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:03/Issue:11/November-2021 Impact Factor- 6.752 www.irjmets.com



VII. NEUROMORPHIC SENSORS

The concept of neuromorphic systems can be extended to sensors. A neuromorphic camera is one example. Researchers are rapidly creating different variations of neuromorphic designs, as they did in the early days of neural networks. As a result, the boundary between research tools for neuromorphic computing and tools for industrial adoption is still blurry from a practical perspective. Generally speaking, neuromorphic hardware can be divided into digital, analog, and mixed-signal circuits. Several architectures have been proposed for performing neurons and synapses in hardware. Digital circuitry is most commonly used to implement neuromorphic architectures. Among the main advantages of this type of circuitry is the ease of development, low power dissipation, and reusability. The main reason for the popularity of digital neuromorphic architectures over their analog counterparts is the low development cost. An example is the Digital Neural Array (DNA), which is an array of digital neurons on a large scale. DNAs target both FPGAs and ASICs. FPGAs, for example, allow reprogramming, while ASICs offer higher density and more reliable performance despite a lower level of flexibility. Analog circuitry is less expressed than digital architectures, in spite of its greater suitability for designing neuromorphic systems. The physical properties of analog implementations are similar to those of neuromorphic architectures and, as with SNNs, these architectures are robust to noise, making them an ideal hardware implementation. Such physical characteristics include reliability and asynchronous operation. The Field-Programmable Analog Array (FPAA) is the best device for analog circuits. The Field-Programmable Neural Array, a custom design aimed at neuromorphic applications, uses programmable components to mimic neuron and synapse functions.



VIII. NEURO-INSPIRED COMPUTING USING PCRAM TECHNOLOGY

A dissection of BioNN reveals that the basic tasks of neuro-inspired computing are to replicate fundamental synapses, neurons, and their synaptic behaviors using hardware technology. Throughout the decades, there have been many implementations of neuro-inspired computing using PCRAM technology. Here, we present comprehensive discussions about how PCRAM is utilized for neural-inspired computing, including the fundamental electric-induced conductance mechanism and advanced techniques to simulate biological components and behavior, as well as current state-of-the-art in intelligent applications built on PCRAM.

ELECTRIC-INDUCED CONDUCTANCE CONTROLLABILITY

Based on biological synaptic behavior, a key precondition for implementing BioNN using hardware is the continuous regulation of conductance. The feasibility of PCRAM originates from the electric-induced controllable phase change that occurred at the active region. By applying appropriate electrical stimulation to

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:03/Issue:11/November-2021 Impact Factor- 6.752 www.irjmets.com

the active area, crystallinity and amorphous states can be controlled bilaterally and continuously. The process of crystalline-to-amorphous change accompanied by prompt quenching is triggered by high and narrow pulses, which is akin to synaptic depression. Additionally, the reverse process of mimicking synaptic potentiation involves long and moderate heating, which typically involves applying a lower and wider pulse. PCRAM can therefore be reliably controlled with special programming pulses in a bidirectional manner, i.e., multilevel reduction in conductance and cumulative enhancement in conductance.

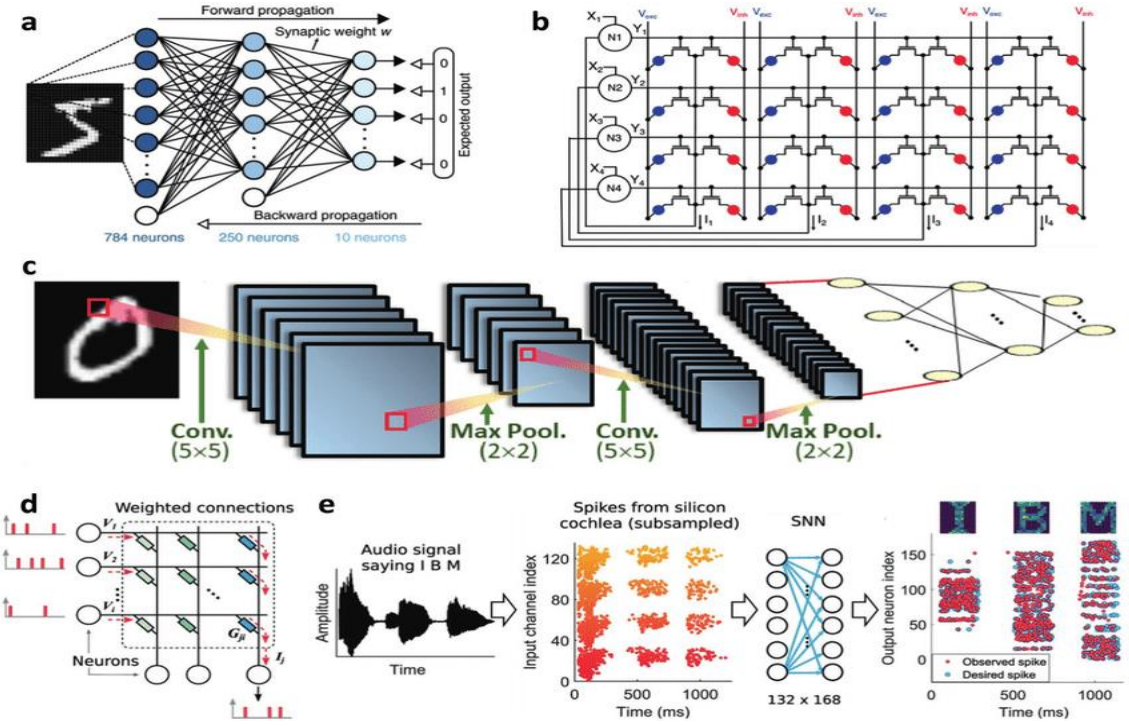
In order to holp up a wide range of workloads, SNN connectivity needs to be flexible and well-provisioned. Some networks may call for dense, all-to-all connectivity while others may call for sparse connectivity; some may have uniform graph degree patterns, others power-law distributions, some may require precision synaptic weights, for example, to support learning, while others can be done with binary connections. As a network grows, algorithm performance scales along with it, as revealed by neuron counts as well as neuron-to-neuron fanout degrees. Biologically, this rule holds true. Based on the O(N2) scaling connectivity state in many fan outs, it becomes an enormous challenge to support networks with high levels of connectivity today using integrated circuit technology.

To address the challenge, it supports a range of features to relax the sometimes severe constraints that other neuromorphic designs had inflicted on the programmer:

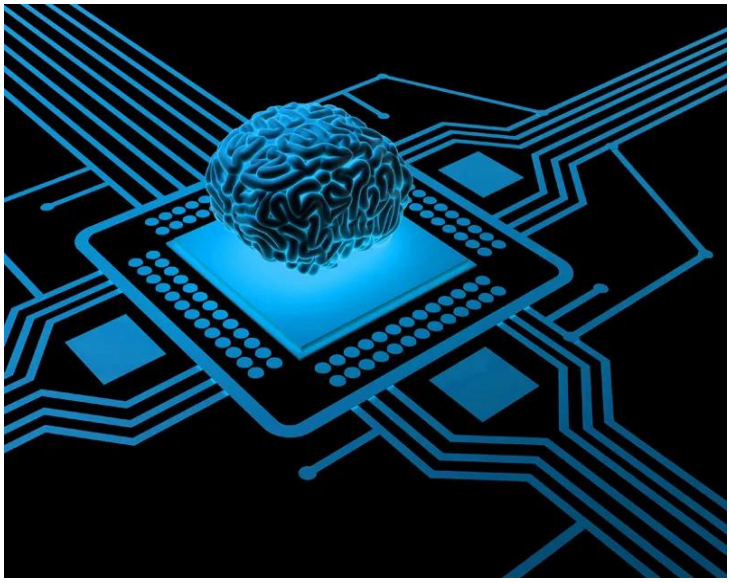
a. Sparse network concentration. In addition to thick matrix connectivity, it also supports three sparse matrix compression models in which neuron indices are calculated using index data stored with synapse state variables.
b. Core to core multicast. As the network connectivity requires, any neuron may send a unique spike to any number of objective cores.
c. Variable synaptic formats. Loihi supports any weight precision between one and nine bits, signed and unsigned, and there is no restriction on how the weight accuracy is set.

a. PCRAM Neuron

The LIF mechanism in BioNN dominates information transmission through the neuron. According to the Hodgkin–Huxley model and various threshold-based neuronal models, the LIF mechanism is driven by complex electrochemical mechanisms combined with the bilayer structure of the membrane. Postsynaptic potentials from dendrites modulate membrane potential stochastically. The sum of potentials above the firing threshold causes neurons to export new potential via axons.



e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:03/Issue:11/November-2021 Impact Factor- 6.752 www.irjmets.com



IX. CONCLUSION

Artificial intelligence is advancing rapidly throughout the world. We will have exponentially more devices connected to the cloud by the end of 2030 than we do now. In addition to more devices on the cloud, there are more data traffic and more hardware requirements. In spite of the current devices, data traffic remains high, and a considerable amount of power and resources are being consumed. Further, it is concerning to know that support structures of information technology are not growing as fast as the technology itself, and we may find ourselves in a situation sooner rather than later when no amount of hardware or software can handle the amounts of data that smart devices will generate. These problems can be solved by neuromorphic computing, which handles much larger volumes of data with drastically reduced energy consumption for the same. Artificial synapse is the groundbreaking technology that makes neuromorphic computing a reality. The artificial synapse is still being researched to fully understand and utilize its benefits. Memristors are the best examples and proof of the level of advancement that we have achieved in the recent past to take us further into the world of artificial intelligence. Although neuromorphic computing has physical and technical limitations and challenges, it is growing at a significant rate and is expected to change computing for good within the next few years.

X. REFERENCES

[1] Zhang, W., Gao, B., Tang, J. et al. Neuro-inspired computing chips. Nat Electron 3, 371–382 (2020). https://doi.org/10.1038/s41928-020-0435-7

[2] Wang, Q.; Niu, G.; Ren, W.; Wang, R.; Chen, X.; Li, X.; Ye, Z.; Xie, Y.; Song, S.; Song, Z. Phase change random access memory forneuro-inspired computing.Adv. Electron. Mater.2021, 2001241.

[3] Wang, P., Yu, S. Ferroelectric devices and circuits for neuro-inspired computing. MRS Communications 10, 538–548 (2020). https://doi.org/10.1557/mrc.2020.71

[4] ellamraju, S., Kumari, S., Girolkar, S., Chourasia, S., Tete, A.D.:Design of various logic gates in neural networks. In: Annual IEEEIndia Conference (INDICON), pp. 1–5 (2013)

[5] Maas, W.: Networks of spiking neurons: the third generation ofneural network models. Neural Netw.10, 1659–1671 (1997)

[6] Danijela Markovic,Alice Mizrahi,Damien Querlioz,Julie Grollier:

[7] https://www.researchgate.net/publication/339840879_Physics_for_Neuromorphic_Computing

[8] Buesing et al.(2011) Neural Dynamics_as_Sampling_A_Model_for_Stochastic Computation_in_Recurrent Networks_of Spiking Neurons

[9] Yellamraju, S., Kumari, S., Girolkar, S., Chourasia, S., Tete, A.D.: Design of various logic gates in neural networks. In: Annual IEEE India Conference (INDICON), pp. 1–5 (2013)

[10] Zhang, Y.; Qu, P.; Ji, Y.; Zhang, W.; Gao, G.; Wang, G.; Song, S.; Li, G.; Chen, W.; Zheng, W.; et al. A system hierarchy forbrain-inspired computing.Nature2020,586, 378–384

[11] S. Yu, "Neuro-inspired computing with emerging nonvolatile memorys," in Proceedings of the IEEE, vol. 106, no. 2, pp. 260-285, Feb. 2018, doi: 10.1109/JPROC.2018.2790840.

[12] Z. Yu, A. M. Abdulghani, A. Zahid, H. Heidari, M. A. Imran and Q. H. Abbasi, "An Overview of Neuromorphic Computing for Artificial Intelligence Enabled Hardware-Based Hopfield Neural Network," in IEEE Access, vol. 8, pp. 67085-67099, 2020, doi: 10.1109/ACCESS.2020.2985839.

[13] X. Si, J. Chen, Y. Tu, W. Huang, J. Wang, Y. Chiu, W. Wei, S. Wu, X. Sun, R. Liu, S. Yu, R. Liu, C. Hsieh, K. Tang, Q. Li, and M. Chang: A twin-8T SRAM computation-in-memory macro for multiple-bit CNN-based machine learning. In 2019 IEEE International Solid- State Circuits Conference — (ISSCC), San Francisco, CA, USA, 2019, pp. 396–398.

[14] Ao P, Wu H, Gao B, Tang J, Zhang Q, Zhang W, Yang J J and Qian H 2020 Fully hardware-implemented memristor convolutional neural network Nature

[15] S. Schmitt, J. Klaehn, G. Bellec, A. Gruebl, M. Guettler, A. Hartel, S. Hartmann, D. Husmann, K. Husmann, V. Karasenko, in 2017 Int. Joint Conf. on Neural Networks (IJCNN), IEEE, Piscataway, NJ 2017.

arXiv:2201.05808v1 [math.CV] 15 Jan 2022

SHARP BOUNDS OF THIRD HANKEL DETERMINANT FOR A CLASS OF STARLIKE FUNCTIONS AND A SUBCLASS OF q-STARLIKE FUNCTIONS

SHAGUN BANGA AND S. SIVAPRASAD KUMAR*

Abstract. Following the trend of coefficient bound problems in Geometric Function Theory, in the present paper, we obtain the sharp bound of $|H_3(1)|$ for the class $\mathcal{S}^*$, of starlike functions and $\mathcal{SL}_q^*$, of q- starlike functions related with lemniscate of Bernoulli. Bound on the initial class is also an improvement over the existing known bound and the bound on the latter class generalizes the prior known outcome. Further, we determine the extremal functions to prove the sharpness of our results.

1. Introduction

Denote the class of analytic functions $f(z) = z + \sum_{n=2}^{\infty} a_n z^n$, defined on the open unit disk $\mathbb{D}$ by $\mathcal{A}$. Let $\mathcal{S}$ be the subclass of $\mathcal{A}$ consisting of the univalent functions. For two analytic functions f and g, we say f is subordinate to g if there exists a Schwarz function $\omega(z)$ with $\omega(0) = 0$ and $|\omega(z)| < 1$ such that $f(z) = g(\omega(z))$. The normalized function f in $\mathcal{S}$ satisfying the following inequality

$$\operatorname{Re} \frac{zf'(z)}{f(z)} > 0, \quad z \in \mathbb{D},$$

belongs to the class of starlike functions, denoted by $\mathcal{S}^*$. Furthermore, various sub-classes of $\mathcal{S}^*$ have been introduced and studied by many authors in the past (see [9, 19, 24, 25]). Likewise, Sokół and Stankiewicz [27] introduced the class $\mathcal{SL}^*$, defined as

$$\mathcal{SL}^* := \left\{ f \in \mathcal{A} : \frac{zf'(z)}{f(z)} \prec \sqrt{1+z}, \quad z \in \mathbb{D} \right\}.$$

Since then, enormous work is done for the class $\mathcal{SL}^*$, for ready reference (see [1, 2, 13, 22, 26]).

Let $\tilde{q}$, $n \in \mathbb{N}$. For a function $f \in \mathcal{A}$, the $\tilde{q}^{th}$ Hankel determinant, is defined as follows:

$$H_{\tilde{q}}(n) := \begin{vmatrix} a_n & a_{n+1} & \dots & a_{n+\tilde{q}-1} \\ a_{n+1} & a_{n+2} & \dots & a_{n+\tilde{q}} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n+\tilde{q}-1} & a_{n+q} & \dots & a_{n+2\tilde{q}-2} \end{vmatrix},$$

introduced in [21] and has been studied by several authors. It also plays an important role in the study of singularities (see [6]). Noor [20] studied the rate of growth of $H_{\tilde{q}}(n)$ as $n \to \infty$ for functions in $\mathcal{S}$ with bounded boundary. Different choices of $\tilde{q}$ and n yields various types of Hankel determinants, such as for $\tilde{q} = 2$ and $n = 1$, the famous Fekete-Szegö functional is given by $H_2(1) := a_3 - a_2^2$. Furthermore, the generalized Fekete-Szegö functional is given by $a_3 - \mu a_2^2$, where μ is either real or complex. For $\tilde{q} = n = 2$, we have second order Hankel determinant $H_2(2) := a_2 a_4 - a_3^2$. Also, another type of second order Hankel determinant is obtained by taking $\tilde{q} = 2$ and $n = 3$, mathematically written as $H_2(3) := a_3 a_5 - a_4^2$. The estimations of the sharp

2010 *Mathematics Subject Classification.* 30C45, 30C50.

Key words and phrases. Sharp bounds, Third Hankel determinant, q-starlike functions, Carathéodory coefficients, lemniscate of Bernoulli.

* corresponding author.

bounds for these $H_{\tilde{q}}(n)$ is obtained by many authors for various sub-classes of $\mathcal{A}$ (see [5, 23, 29]). Third order Hankel determinant, given by

$$H_3(1) = a_3(a_2a_4 - a_3^2) - a_4(a_4 - a_2a_3) + a_5(a_3 - a_2^2), \tag{1.1}$$

is obtained when $\tilde{q} = 3$ and $n = 1$. Sharp bound of $|H_3(1)|$ was not obtained for any class of analytic functions before 2018. It is achieved by Kowalczyk *et al.* [12], for functions in $\mathcal{A}$ satisfying $\mathrm{Re}(f(z)/z) > \alpha$, $\alpha \in [0, 1)$ and in [11] for convex functions. Following which, Banga and Kumar [4] recently derived sharp bound of third Hankel determinant as $|H_3(1)| \leq 1/36$ for functions in $\mathcal{SL}^*$, which earlier was calculated to be $43/576$ in [23]. Lecko *et al.* [16] found that $1/9$ is the sharp bound of the third Hankel determinant for starlike functions of order $1/2$. The credit of initiation of sharp bound of $|H_3(1)|$ goes to Kwon *et al.* [14] who deduced p_4 in terms of p_1, where $p_i's$ are the coefficients of the functions in the Carathéodory class $\mathcal{P}$, defined by:

$$p(z) = 1 + p_1z + p_2z^2 + p_3z^3 + p_4z^4 + \cdots \ (z \in \mathbb{D}).$$

Let us recall the $q-$derivative of a complex valued function defined on a subset of $\mathbb{C}$, defined as below

$$(D_qf)(z) = \begin{cases} \dfrac{f(z) - f(qz)}{(1-q)z}, & z \neq 0 \\ f'(0), & z = 0, \end{cases}$$

where $q \in (0, 1)$. Whenever f is differentiable on a given subset of $\mathbb{C}$, the above definition of $q-$derivative implies

$$\lim_{q\to 1^-} (D_qf)(z) = \lim_{q\to 1^-} \frac{f(z) - f(qz)}{(1-q)z} = f'(z).$$

Furthermore, Taylor series expansion of f yields that

$$(D_qf)(z) = 1 + \sum_{n=2}^{\infty} [n]_q a_n z^{n-1},$$

where

$$[n]_q = \sum_{k=0}^{n-1} q^k = 1 + q + q^2 + \cdots + q^{n-1}, \ n \in \mathbb{N}.$$

The initiation of the above defined $q-$calculus is done by Jackson [8]. In Geometric Function Theory, sub-classes of normalized analytic functions have been studied with different view points. Ismail *et al.* [7] generalized the class $\mathcal{S}^*$ of starlike functions by introducing a new class with the usage of $q-$calculus. This marked the beginning of introduction of $q-$version of various classes in Geometric Function Theory. For instance, Srivastava and Bansal [28] study a certain family of $q-$Mittag-Leffler functions and Mahmood *et al.* [18] dealt $q-$starlike functions associated with conic domains. Recently, Khan *et al.* [10] used $q-$derivative operator to define a new subclass of starlike functions related with the lemniscate of Bernoulli, given as

$$\mathcal{SL}_q^* := \left\{ f \in \mathcal{A} : \frac{z(D_qf)(z)}{f(z)} \prec \sqrt{\frac{2(1+z)}{2+(1-q)z}}, \ \ z \in \mathbb{D} \right\},$$

or equivalently, a function $f \in \mathcal{A}$ is in $\mathcal{SL}_q^*$ if it satisfies the following

$$\left| \left(\frac{z(D_qf)(z)}{f(z)} \right)^2 - \frac{1}{1-q} \right| < \frac{1}{1-q}.$$

This implies on choosing $\omega = z(Dqf)(z)/f(z)$, the analytic characterization of the class $\mathcal{SL}_q^*$ can be expressed as $|\omega^2 - 1/(1-q)| < 1/(1-q)$, which is the interior of the right loop of the lemniscate of Bernoulli. The specialty of this class lies in the fact that it reduces to a well-known class $\mathcal{SL}^*$,

when $q \to 1^-$. The authors in [10] obtained the sharp bounds of Fekete-Szegö functional, $|H_2(2)|$, initial coefficients a_2, a_3, a_4 and a_5 and upper bound of third Hankel determinant for functions in $\mathcal{SL}_q^*$.

Our study focuses on the estimation of sharp bound of $|H_3(1)|$ for functions in $\mathcal{SL}_q^*$ and $\mathcal{S}^*$. It was found in [3] that $|H_3(1)| \leq 16$ for functions in $\mathcal{S}^*$, which is improved by Zaprawa [30], wherein he proved $|H_3(1)| \leq 1$. Later in [15], it was further improved to $8/9$. Again, in 2021, Zaprawa *et al.* [31] calculated the same to be $5/9$, to which we eventually improve in the present paper to a sharp estimate of $4/9$. In addition, we obtain $|H_3(1)| \leq \frac{(1+q)^2}{16q^2(1+q+q^2)^2}$ for functions in ${\mathcal{SL}^*}_q$. This bound apart from being sharp is an improvement over the bound obtained in [10]. Moreover, for $q \to 1^-$, this bound reduces to earlier known sharp bound for $\mathcal{SL}^*$ [4]. We also give extremal functions to justify our claims.

We state below a lemma for the formulae of p_2, p_3 [17] and p_4 [14] in order to prove our results.

Lemma 1.1. *Let $p \in \mathcal{P}$ and of the form $1 + \sum_{n=1}^{\infty} p_n z^n$. Then*

$$2p_2 = p_1^2 + \lambda(4 - p_1^2),$$

$$4p_3 = p_1^3 + 2p_1(4 - p_1^2)\lambda - p_1(4 - p_1^2)\lambda^2 + 2(4 - p_1^2)(1 - |\lambda|^2)\mu$$

and

$$\begin{aligned} 8p_4 =& p_1^4 + (4 - p_1^2)\lambda(p_1^2(\lambda^2 - 3\lambda + 3) + 4\lambda) \\ & - 4(4 - p_1^2)(1 - |\lambda|^2)(p_1(\lambda - 1)\mu + \overline{\lambda}\mu^2 - (1 - |\mu|^2)\delta), \end{aligned}$$

for some δ, λ and μ such that $|\delta| \leq 1$, $|\lambda| \leq 1$ and $|\mu| \leq 1$.

2. Main Results

This section begins with the following result.

Theorem 2.1. *Let $q \in (0,1)$ and $f \in \mathcal{SL}_q^*$ of the form $f(z) = z + \sum_{n=2}^{\infty} a_n z^n$. Then we have*

$$|H_3(1)| \leq \frac{(1+q)^2}{16q^2(1+q+q^2)^2}.$$

Proof. For $f \in \mathcal{SL}_q^*$, we refer [10] for the expressions of a_2, a_3 and a_4. On the similar lines, we compute

$$\begin{aligned} a_5 = \frac{1}{32768q^4(1+q^2)(1+q+q^2)}\Big(& 512p_1p_3q^2(2 - 10q - 8q^2 - 9q^3 + 3q^4) + p_1^4(8 - 140q + 802q^2 \\ & - 1435q^3 - 340q^4 - 1193q^5 + 1015q^6 - 320q^7 + 35q^8) + 32p_1^2p_2q(6 - 68q + 175q^2 + 89q^3 \\ & + 148q^4 - 93q^5 + 15q^6) + 256q^2(1+q+q^2)(16p_4q + p2^2(2 - 13q + 3q^2))\Big). \end{aligned}$$

Now substituting the values of above $a_i's$ in (1.1) with $p := p_1 \in [0,2]$, we obtain

$$\begin{aligned}H_3(1) := \frac{1}{4194304\ q^2(1+q^2)(1+q+q^2)^2}\Big(&-8192pp_2p_3(-14-28q-13q^2-28q^3-12q^4+q^5)\\ &-512p_1^3p_3(-14-10q-217q^2+23q^3-8q^4+5q^5+q^6)+16p_1^4p_2(-31+1111q-10148q^2\\ &+3026q^3-594q^4-84q^5-19q^6+3q^7)+p_1^6(239-4972q+35429q^2-13002q^3+3964q^4\\ &+370q^5+63q^6-44q^7+q^8)+4096(-16p_3^2(1+q)^2(1+q^2)+16p_2p_4(1+q+q^2)^2\\ &+p_2^3(-13+q)(1+q+q^2)^2)+256p_1^2(16p_4(-15+q)(1+q+q^2)^2+p_2^2(-27-102q\\ &+670q^2-188q^3+6q^4+14q^5+3q^6))\Big).\end{aligned}$$

Applying Lemma 1.1 in the above equation for the values of p_2, p_3 and p_4 and further reducing it to the simpler form, we arrive at

$$H_3(1) = \frac{\tau_1(p,\lambda)+\tau_2(p,\lambda)\mu+\tau_3(p,\lambda)\mu^2+\zeta(p,\lambda,\mu)\delta}{4194304\ q^2(1+q^2)(1+q+q^2)^2}, \tag{2.1}$$

whenever δ, μ, $\lambda \in \overline{\mathbb{D}}$ and

$$\begin{aligned}\tau_1(p,\lambda) := Ap^6+p^2(4-p^2)\lambda\Big(&8(-15+183q-804q^2+434q^3-242q^4-20q^5-3q^6+3q^7)p^2\\ &+64(45+66q+262q^2+28q^3+62q^4+6q^5+3q^6)(4-p^2)\lambda-512(7+15q-3q^2\\ &+17q^3+5q^4-q^5)(4-p^2)\lambda^2-2048(7-q)(1+q+q^2)^2\lambda+4096q^2(4-p^2)\lambda^3\\ &-512(7-q)(1+q+q^2)^2p^2\lambda^2+128(22+50q+35q^2+59q^3+20q^4+q^5+q^6)p^2\lambda\Big)\\ &-2048(5-q)(1+q+q^2)^2(4-p^2)^2\lambda^3,\\ \tau_2(p,\lambda) := (4-p^2)(1-|\lambda|^2)\Big(&256(6+2q+41q^2-15q^3-5q^5-q^6)p^3+2048(7-q)(1+q+q^2)^2p^3\lambda\\ &+p(4-p^2)\lambda(2048(6+12q+5q^2+12q^3+4q^4-q^5)-16384q^2\lambda)\Big),\\ \tau_3(p,\lambda) := (4-p^2)(1-|\lambda|^2)(&2048(7-q)(1+q+q^2)^2p^2\overline{\lambda}-(4-p^2)(16384q^2|\lambda|^2\\ &+16384(1+q^2)(1+q)^2)),\\ \zeta(p,\lambda,\mu) := (4-p^2)(1-|\lambda|^2)(1-|\mu|^2)&(1+q+q^2)^2(-(14336-2048q)p^2+16384(4-p^2)\lambda),\end{aligned}$$

where $A := 55-308q+1349q^2-698q^3+620q^4-46q^5-25q^6-20q^7+q^8$. Taking modulus over equation (2.1) and applying triangle inequality, we get

$$|H_3(1)| \le \frac{|\tau_1(p,\lambda)|+|\tau_2(p,\lambda)|y+|\tau_3(p,\lambda)|y^2+|\zeta(p,\lambda,\mu)|}{4194304\ q^2(1+q^2)(1+q+q^2)^2} \le \tilde{T}(p,x,y),$$

where $x := |\lambda|$, $y := |\mu|$ and the fact $|\delta| \le 1$ and

$$\begin{aligned}\tilde{T}(p,x,y) &:= \frac{t_1(p,x)+t_2(p,x)y+t_3(p,x)y^2+t_4(p,x)(1-y^2)}{4194304\ q^2(1+q^2)(1+q+q^2)^2}\\ &=: \frac{T(p,x,y)}{4194304\ q^2(1+q^2)(1+q+q^2)^2}.\end{aligned}$$

with

$$\begin{aligned}
t_1(p,x) &:= Ap^6 + p^2(4-p^2)x\Big(8(15-183q+804q^2-434q^3+242q^4+20q^5+3q^6-3q^7)p^2\\
&\quad + 64(45+66q+262q^2+28q^3+62q^4+6q^5+3q^6)(4-p^2)x + 512(7+15q-3q^2\\
&\quad + 17q^3+5q^4-q^5)(4-p^2)x^2 + 2048(7-q)(1+q+q^2)^2x + 4096(4-p^2)x^3\\
&\quad + 512(7-q)(1+q+q^2)^2p^2x^2 + 128(22+50q+35q^2+59q^3+20q^4+q^5+q^6)p^2x\Big)\\
&\quad + 2048(5-q)(1+q+q^2)^2(4-p^2)^2x^3,\\
t_2(p,x) &:= (4-p^2)(1-x^2)\Big(256(6+2q+41q^2-15q^3-5q^5-q^6)p^3 + 2048(7-q)(1+q+q^2)^2p^3x\\
&\quad + p(4-p^2)x(2048(6+12q+5q^2+12q^3+4q^4-q^5)+16384q^2x)\Big),\\
t_3(p,x) &:= (4-p^2)(1-x^2)(2048(7-q)(1+q+q^2)^2p^2x + (4-p^2)(16384q^2x^2\\
&\quad + 16384(1+q^2)(1+q)^2)),\\
t_4(p,x) &:= (4-p^2)(1-x^2)(1+q+q^2)^2((14336-2048q)p^2+16384(4-p^2)x).
\end{aligned}$$

In order to achieve the desired bound, we need to maximize $T(p,x,y)$ in the closed cuboid $\mathfrak{C}$: $[0,2]\times[0,1]\times[0,1]$. We accomplish this by estimating maximum values in the interior of $\mathfrak{C}$, interior of the six faces and finally on the twelve edges.

I. We begin with interior points of $\mathfrak{C}$, which means taking $(p,x,y)\in(0,2)\times(0,1)\times(0,1)$. For this, we calculate

$$\begin{aligned}
\frac{\partial T}{\partial y} &= (4-p^2)(1-x^2)(2y(16384(4-p^2)((1+q^2)(1+q)^2+q^2x^2)+2048p^2x(7-q)(1+q+q^2)^2\\
&\quad - ((14336-2048q)p^2+16384(4-p^2)x)(1+q+q^2)^2) + 256(6+2q+41q^2-15q^3-5q^5\\
&\quad - q^6)p^3 + 2048(7-q)(1+q+q^2)^2p^3x + 2048px(4-p^2)(6+12q+5q^2+12q^3+4q^4-q^5)\\
&\quad + 16384pq^2(4-p^2)x^2).
\end{aligned}$$

On solving $\partial T/\partial y = 0$, we obtain $y = y_0$, given as

$$y_0 := \frac{\tilde{A}}{2048(1-x)\left((7-q)(1+q+q^2)^2p^2 - 8(4-p^2)\left(\frac{1+2q+2q^2+2q^3+q^4}{q^2}-x\right)\right)},$$

where $\tilde{A} := p^3(128(6+2q+41q^2-15q^3-5q^5-q^6)+1024(7-q)(1+q+q^2)^2x)+1024px(4-p^2)(6+12q+5q^2+12q^3+4q^4-q^5+8q^2x)$. For $y_0\in(0,1)$, we must have

$$(7-q)(1+q+q^2)^2p^2 > 8(4-p^2)\left(\frac{1+2q+2q^2+2q^3+q^4}{q^2}-x\right)$$

and

$$\tilde{A} + 16384(4-p^2)(1-x)\left(\frac{1+2q+2q^2+2q^3+q^4}{q^2}-x\right) < 2048(1-x)(7-q)(1+q+q^2)^2p^2. \tag{2.2}$$

Let us assume $p\to 2$, then there exists $x\in(0,\frac{101}{216})$ for every $q\in(0,1)$ such that (2.2) holds. Moreover when we consider $x\in[\frac{101}{216},1)$, then there does not exist any $p\in(0,2)$ for all $q\in(0,1)$ such that (2.2) holds. Assuming $x\to 0$, we compute (2.2) holds for $p\geq 1.48855$ for every $q\in(0,1)$. In fact whenever $p\in(0,1.48855)$, there does not exist any $x\in(0,1)$ for all $q\in(0,1)$ such that

(2.2) holds. Thus we conclude possible solution exists in $[1.48855, 1) \times (0, \frac{101}{216})$ for inequality (2.2). A computation shows

$$\left.\frac{\partial T}{\partial p}\right|_{y=y_0} \neq 0,$$

in this interval. Therefore, there does not exist any critical point in the interior of $\mathfrak{C}$.

II. Now we compute the maximum value of T in the interior of all the six faces of $\mathfrak{C}$. On the face $p = 0$, $T(p, x, y)$ reduces to

$$\begin{aligned} T(0, x, y) = {} & 262144(1 - x^2)((1 + q^2)(1 + q)^2 + q^2x^2 - (1 + q + q^2)^2x)y^2 \\ & + 32768x(1 + q + q^2)^2(x^2(5 - q) + 8(1 - x^2)), \end{aligned} \tag{2.3}$$

which in turn differentiating with respect to y becomes

$$\frac{\partial T}{\partial y} = 524288y(1 - x^2)(x - 1)\left(x - \frac{1 + 2q + 2q^2 + 2q^3 + q^4}{q^2}\right) \neq 0 \quad x, y \in (0, 1).$$

This clearly shows there does not exist any critical point for $T(0, x, y)$ in $(0, 1) \times (0, 1)$.

On the face $p = 2$,

$$\tilde{T}(p, x, y) = \tilde{T}(2, x, y) = \frac{A}{65536q^2(1 + q^2)(1 + q + q^2)^2} \leq \frac{(1 + q)^2}{16q^2(1 + q + q^2)^2},$$

$x, y \in (0, 1)$, as we have $-4041 - 8500q - 6843q^2 - 8890q^3 - 3476q^4 - 46q^5 - 25q^6 - 20q^7 + q^8 \leq 0$ for $q \in (0, 1)$.

On the face $x = 0$, $T(p, x, y)$ becomes

$$\begin{aligned} T(p, 0, y) = {} & p^6(55 - 308q + 1349q^2 - 698q^3 + 620q^4 - 46q^5 - 25q^6 - 20q^7 + q^8) \\ & + 256(4 - p^2)(-p^3(-6 - 2q - 41q^2 + 15q^3 + 5q^5 + q^6)y + 64(4 - p^2)(1 \\ & + q)^2(1 + q^2)y^2 + 8p^2(-7 + q)(1 + q + q^2)^2(-1 + y^2)) := h_1(p, y). \end{aligned} \tag{2.4}$$

On solving $\frac{\partial h_1}{\partial y} = 0$, we get

$$y =: y_1 = \frac{p^3(-6 - 2q - 41q^2 + 15q^3 + 5q^5 + q^6)}{16(32(1 + q)^2(1 + q^2) + p^2(-15 - 29q - 35q^2 - 27q^3 - 13q^4 + q^5))}. \tag{2.5}$$

For $0 < p \leq 1.46$, we have $y_1 \leq 0$ for every $q \in (0, 1)$. There exists some $q \in (0, 1)$ whenever $p \in (1.46, 2)$ such that $y_1 > 0$. On substituting (2.5) in $\frac{\partial h_1}{\partial p}$ and simplifying further, we get $\frac{\partial h_1}{\partial p} \neq 0$, where $p \in (1.46, 2)$, $q \in (0, 1)$.Thus $h_1(p, y)$ has no critical point in $(0, 2) \times (0, 1)$.

On the face $x = 1$, $T(p, x, y)$ reduces to

$$\begin{aligned} T(p, 1, y) =: {} & -32768(-5 + q)(1 + q + q^2)^2 + p^6(1 + q)^3(-1 + 7q + 19q^2 + 9q^3 \\ & + q^4 + q^5) + 1024p^2(77 + 146q + 246q^2 + 140q^3 + 94q^4 + 6q^5 + 3q^6) \\ & - 32p^4(929 + 1783q + 2636q^2 + 1666q^3 + 878q^4 - 4q^5 + 29q^6 + 3q^7) =: h_2(p). \end{aligned} \tag{2.6}$$

On differentiating h_2 with respect to p, we obtain

$$\begin{aligned} \frac{\partial h_2}{\partial p} = {} & 6p^5(1 + q)^3(-1 + 7q + 19q^2 + 9q^3 + q^4 + q^5) + 2048p(77 + 146q + 246q^2 + 140q^3 + 94q^4 \\ & + 6q^5 + 3q^6) - 128p^3(929 + 1783q + 2636q^2 + 1666q^3 + 878q^4 - 4q^5 + 29q^6 + 3q^7), \end{aligned}$$

further which becomes 0 at $p = 0$ and $p = p_0$, given by

$$p_0 := \sqrt{\frac{32(929 + 1783q + 2636q^2 + 1666q^3 + 878q^4 - 4q^5 + 29q^6 + 3q^7)}{3(1 + q)^3(-1 + 7q + 19q^2 + 9q^3 + q^4 + q^5)} - \tilde{A}},$$

where $\tilde{A} = \dfrac{64\sqrt{2}A_0}{3(-1+4q+37q^2+86q^3+92q^4+50q^5+15q^6+4q^7+q^8)}$ and

$$A_0 = (107909+414041q+1008402q^2+1557100q^3+1804144q^4+1471838q^5+913014q^6 + 363176q^7+107408q^8+9900q^9+6570q^{10}+346q^{11}+41q^{12}+15q^{13})^{(1/2)}.$$

A calculation yields $p=0$ is a point of minima and p_0 is a point of maxima and maximum value is given by a huge mathematical expression in q which is computed to be less than $262144(1+q)^2(1+q^2)$.

On the face $y=0$, we have $T(p,x,0) =: h_3(p,x)$, given by

$$\begin{aligned} h_3(p,x) :=& p^6(55-308q+1349q^2-698q^3+620q^4-46q^5-25q^6-20q^7+q^8) \\ &+ 2048(4-p^2)(1+q+q^2)^2(-1+x^2)(-32x+p^2(-7+q+8x)) \\ &+ 8(4-p^2)x(-1024(-5+q)(1+q+q^2)^2x^2+32p^2x(101-2q^5 \\ &+ 3q^6+16x+2q^4(51+8x)+4q^3(29+20x)+2q(85+24x) \\ &+ 2q^2(207-64x+32x^2))-p^4(-15+3q^7+8x+22q^4(-11 \\ &+ 8x)+4q^5(-5+8x)+q^6(-3+8x)+q(183-272x+128x^2) \\ &+ q^3(434-720x+384x^2)+4q^2(-201+384x-352x^2+128x^3))). \end{aligned}$$

A calculation yields that there is no common solution to the system of equations $\frac{\partial h_3}{\partial x}=0$ and $\frac{\partial h_3}{\partial p}=0$ in $(0,2)\times(0,1)$. Similarly we can show that there does not exist any critical point for $T(p,x,1)$.

III. Finally, we estimate the maximum value on the edges of the cuboid $\mathfrak{C}$.
Starting with the $T(p,0,0) =: h_4(p)$, given by

$$\begin{aligned} h_4(p) =& (1+q+q^2)^2(4p^2(14336-2048q)-p^4(14336-2048q))+(55 \\ &- 308q+1349q^2-698q^3+620q^4-46q^5-25q^6-20q^7+q^8)p^6, \end{aligned}$$

obtained from (2.4). On solving $\frac{\partial h_4}{\partial p}=0$, we get either $p=0$ or $p =: p_0$, given by

$$\begin{aligned} p_0 :=& \frac{1}{\sqrt{3A}}\Bigg(2048(7-q)(1+q+q^2)^2-64\sqrt{2}\Bigg(23933+97507q+203268q^2+309564q^3+313752q^4 \\ &+250248q^5+114774q^6+34938q^7-6156q^8-644q^9+1352q^{10}+192q^{11}-75q^{12}+3q^{13}\Bigg)^{1/2}\Bigg)^{1/2}. \end{aligned}$$

We compute that the function $h_4(0)=0$ is a minimum value of $h_4(p)$ and $h_4(p_0)$ is a huge mathematical expression in q which is also a maximum value of $h_4(p)$. Further we have $\tilde{T}(p_0,0,0) \le (1+q)^2/16q^2(1+q+q^2)^2$. Substituting $y=1$ in equation (2.4), we obtain

$$\begin{aligned} T(p,0,1) = h_5(p) =& (4-p^2)(-256(-6-2q-41q^2+15q^3+5q^5+q^6)p^3 \\ &+ (4-p^2)(16384(1+q^2)(1+q)^2))+Ap^6. \end{aligned}$$

The function $h_5(p)$ is a decreasing function of p for all q. Thus

$$\max_{p\in[0,2]} \tilde{T}(p,0,1) = \tilde{T}(0,0,1) = \frac{(1+q)^2}{16q^2(1+q+q^2)^2}.$$

Form equation (2.6), which is independent of y, we get $\tilde{T}(p,1,0) = \tilde{T}(p,1,1) = \tilde{T}(p,1,y)$. Thus $\tilde{T}(p,1,0) = \tilde{T}(p,1,1) \le \frac{(1+q)^2}{16q^2(1+q+q^2)^2}$. Substituting $x=0$ in (2.3), we obtain $\tilde{T}(0,0,y) = y^2(1+$

$q)^2/16q^2(1+q+q^2)^2$, which is clearly an increasing function of y for all q and we have

$$\tilde{T}(0,0,y) \leq \tilde{T}(0,0,1) = \frac{(1+q)^2}{16q^2(1+q+q^2)^2}.$$

Evaluating (2.6) at $p=0$, we get

$$\tilde{T}(0,1,y) = \frac{5-q}{128q^2(1+q^2)}.$$

The value of $\tilde{T}(p,x,y)$ on the edges $p=2$, $x=1$; $p=2$, $x=0$; $p=2$, $y=0$; $p=2$, $y=1$ are all respectively equal to $\tilde{T}(2,1,y) = \tilde{T}(2,0,y) = \tilde{T}(2,x,0) = \tilde{T}(2,x,1) = \tilde{T}(2,x,y)$ as $\tilde{T}(2,x,y)$ is independent of both x and y, which further equals to

$$\frac{A}{65536q^2(1+q^2)(1+q+q^2)^2} \leq \frac{(1+q)^2}{16q^2(1+q+q^2)^2}.$$

Evaluating equation (2.3) at $y=0$, we deduce

$$T(0,x,0) = h_6(x) = 32768(1+q+q^2)^2x(8-(3+q)x^2).$$

On solving $h_6'(x)=0$, we get

$$x = x_0 := \frac{512(1+q+q^2)}{\sqrt{294912+688128q+1081344q^2+884736q^3+491520q^4+98304q^5}}.$$

A computation shows that x_0 is a point of maxima and maximum value is given by

$$\max_{x\in[0,1]} h_6(x) = h_6(x_0) = \frac{\sqrt{2}(1+q+q^2)}{12\sqrt{3(3+q)}} \ (0<q<1).$$

Also, we have

$$\max_{0\leq x\leq 1} \tilde{T}(0,x,0) \leq \frac{(1+q)^2}{16q^2(1+q+q^2)^2}.$$

Now evaluating (2.3) at $y=1$, we obtain

$$T(0,x,1) = 262144(1-x^2)((1+q^2)(1+q)^2+q^2x^2) + 32768x^3(5-q)(1+q+q^2)^2,$$

which is clearly a decreasing function of x and attains maximum value at $x=0$, given by $\frac{(1+q)^2}{16q^2(1+q+q^2)^2}$.

Altogether I-III, yield $|H_3(1)| \leq \frac{(1+q)^2}{16q^2(1+q+q^2)^2}$. The result is sharp as equality occurs for the function $\tilde{f}:\mathbb{D}\to\mathbb{C}$, satisfying the following

$$\frac{z(D_q\tilde{f})(z)}{\tilde{f}(z)} = \sqrt{\frac{2(1+z^3)}{2+(1-q)z^3}}.$$

∎

Let $q\to 1^-$ in the above Theorem, then it reduces to the following result obtained by Banga and Kumar [4].

Corollary 2.2. *Let $f\in\mathcal{SL}^*$. Then $|H_3(1)|\leq 1/36$.*

Moreover, extremal functions also coincide in the case of $q\to 1^-$.

Theorem 2.3. *Let $f\in\mathcal{S}^*$ of the form $f(z)=z+\sum_{n=2}^{\infty}a_nz^n$. Then the sharp bound for third order Hankel determinant for such functions is given by*

$$|H_3(1)|\leq 4/9. \tag{2.7}$$

Proof. For $f \in \mathcal{S}^*$, we have

$$\frac{zf'(z)}{f(z)} = \frac{1+\omega(z)}{1-\omega(z)}, \tag{2.8}$$

for some Schwarz function $\omega(z)$. Define a function, $p(z) = \frac{1+\omega(z)}{1-\omega(z)}$, then evidently $p \in \mathcal{P}$. The equation (2.8) now reduces to

$$\frac{zf'(z)}{f(z)} = p(z) = 1 + p_1 z + p_2 z^2 + p_3 z^3 + \cdots .$$

The Taylor series of which yield

$$a_2 = p_1,\ a_3 = \frac{p_2 + p_1^2}{2},\ a_4 = \frac{p_1^3 + 3p_1p_2 + 2p_3}{6}$$

and

$$a_5 = \frac{p_1^4 + 6p_1^2p_2 + 3p_2^2 + 8p_1p_3 + 6p_4}{24}.$$

Here, we assume $p_1 =: p$ lies in the interval $[0, 2]$ due to the invariant property of class $\mathcal{P}$ under rotation. Equation (1.1), together with the above expressions of $a_i's$, yield

$$H_3(1) = \frac{-p^6 + 3p^4p_2 + 8p^3p_3 + 24pp_2p_3 - 9p^2p_2^2 - 18p^2p_4 - 9p_2^3 - 16p_3^2 + 18p_2p_4}{144}.$$

Applying Lemma 1.1 in the above equation for the values of p_2, p_3 and p_4 and further reducing it to the simpler form, we arrive at

$$H_3(1) = \frac{1}{1152}\bigg(\tau_1(p,\lambda) + \tau_2(p,\lambda)\mu + \tau_3(p,\lambda)\mu^2 + \varsigma(p,\lambda,\mu)\delta\bigg).$$

Where δ, μ, $\lambda \in \overline{\mathbb{D}}$,

$$\begin{aligned}
\tau_1(p,\lambda) &:= -2p^2\lambda^2(4-p^2)^2 - 10p^2\lambda^3(4-p^2)^2 + p^2\lambda^4(4-p^2)^2 \\
&\quad + 3p^4\lambda(4-p^2) + 3p^4\lambda^2(4-p^2) - 36p^2\lambda^2(4-p^2) - 9p^4\lambda^3(4-p^2), \\
\tau_2(p,\lambda) &:= (4-p^2)(1-|\lambda|^2)\left(12p^3 + 36p^3\lambda + p\lambda(4-p^2)(20-4\lambda)\right), \\
\tau_3(p,\lambda) &:= (4-p^2)(1-|\lambda|^2)\left(36p^2\overline{\lambda} - 4(4-p^2)(|\lambda|^2+8)\right), \\
\varsigma(p,\lambda,\mu) &:= (4-p^2)(1-|\lambda|^2)(1-|\mu|^2)\left(-36p^2 + 36\lambda(4-p^2)\right).
\end{aligned}$$

Assuming $x := |\lambda|$, $y := |\mu|$ and using the fact $|\delta| \leq 1$, we have

$$|H_3(1)| \leq \frac{|\tau_1(p,\lambda)| + |\tau_2(p,\lambda)|y + |\tau_3(p,\lambda)|y^2 + |\varsigma(p,\lambda,\mu)|}{1152} \leq S(p,x,y),$$

where

$$S(p,x,y) := \frac{1}{1152}\bigg(s_1(p,x) + s_2(p,x)y + s_3(p,x)y^2 + s_4(p,x)(1-y^2)\bigg) \tag{2.9}$$

with

$$\begin{aligned}
s_1(p,x) &:= 2p^2x^2(4-p^2)^2 + 10p^2x^3(4-p^2)^2 + p^2x^4(4-p^2)^2 + 3p^4x(4-p^2) \\
&\quad + 3p^4x^2(4-p^2) + 36p^2x^2(4-p^2) + 9p^4x^3(4-p^2), \\
s_2(p,x) &:= (4-p^2)(1-x^2)(12p^3 + px(4-p^2)(20+4x) + 36p^3x), \\
s_3(p,x) &:= (4-p^2)(1-x^2)(32(4-p^2) + 4x^2(4-p^2) + 36p^2x), \\
s_4(p,x) &:= (4-p^2)(1-x^2)(36p^2 + 36x(4-p^2)).
\end{aligned}$$

Our aim is to maximize $S(p,x,y)$ in the closed cuboid $\mathfrak{C} : [0,2]\times[0,1]\times[0,1]$. We accomplish this by obtaining the maximum values in the interior of $\mathfrak{C}$, in the interior of the six faces and on the twelve edges.

I. First we consider the interior points of $\mathfrak{C}$. Let $(p,x,y) \in (0,2)\times(0,1)\times(0,1)$. In order to achieve the maximum value in the interior of $\mathfrak{C}$, we partially differentiate equation (2.9) with respect to y and further reduce it to a simpler expression as

$$\frac{\partial S}{\partial y} = \frac{1}{1152}(4-p^2)(1-x^2)(8y(x-1)((4-p^2)(x-8)+9p^2)+4p(x(4-p^2)(5+x)+p^2(3+9x))).$$

Now $\frac{\partial S}{\partial y} = 0$ yields

$$y =: y_0 = \frac{2p(x(4-p^2)(5+x)+p^2(3+9x))}{(1-x)((4-p^2)(x-8)+9p^2)}.$$

In order to find the critical points, we first ensure y_0 should lie in the interval $(0,1)$, which is possible only when

$$p^3(6+18x)+2px(4-p^2)(5+x)+(1-x)(8-x)(4-p^2) < 9p^2(1-x) \tag{2.10}$$

and

$$9p^2 > (4-p^2)(8-x). \tag{2.11}$$

We determine the common solutions for the above inequalities. A computation shows that inequality (2.11) holds for all $x \in (0,1)$ whenever $p > 1.37199$, but inequality (2.10) does not hold in $(0,2)\times(0,1)$. Therefore the function S has no critical point in the given domain of values.

II. Here below we calculate the maximum value on the six faces of the cuboid $\mathfrak{C}$.
On the face $p = 0$, $S(p,x,y)$ becomes

$$h_1(x,y) := S(0,x,y) = \frac{(1-x^2)(y^2(x-1)(x-8)+9x)}{18}, \tag{2.12}$$

where $x, y \in (0,1)$. We calculate

$$\frac{\partial h_1}{\partial y} = \frac{(1-x^2)y}{9}\left((x-1)(x-8)\right) \neq 0, \quad x,y \in (0,1).$$

Clearly, we can infer from above that h_1 has no critical point in $(0,1)\times(0,1)$.

On the face $p = 2$, $S(p,x,y)$ becomes

$$S(2,x,y) = 0, \quad x,y \in (0,1). \tag{2.13}$$

On the face $x = 0$, $S(p,x,y)$ becomes

$$S(p,0,y) =: h_2(p,y) = \frac{(4-p^2)}{288}\left(3p^3y + y^2(8(4-p^2)-9p^2)+9p^2\right), \tag{2.14}$$

$y \in (0,1)$ and $p \in (0,2)$. Now we differentiate $h_2(p,y)$ partially with respect to y and obtain

$$\frac{\partial h_2}{\partial y} = \frac{(4-p^2)}{288}\left(3p^3 + 2y(8(4-p^2)-9p^2)\right), \quad p \in (0,2) \text{ and } y \in (0,1).$$

On solving $\partial h_2/\partial y = 0$, we get

$$y = \frac{3p^3}{2(17p^2-32)}, \tag{2.15}$$

which belongs to $(0,1)$ only when $p > p_0 \approx 1.47073$. Upon substituting the value of y from equation (2.15) in $\partial h_2/\partial p = 0$, we arrive at

$$\frac{(p(16384 - 25600p^2 + 12944p^4 - 2048p^6 - 51p^8)}{64(32 - 17p^2)^2} = 0,$$

for $p = 1.20671$ in $(0,2)$. Thus there does not exist any critical point of h_2 in $(0,2) \times (0,1)$.

On the face $x = 1$, $S(p,x,y)$ becomes

$$S(p,1,y) =: h_3(p) = \frac{p^2}{576}\left(176 - 40p^2 - p^4\right). \tag{2.16}$$

To find the maximum value of h_3, we solve $\partial h_3/\partial p = 0$, which implies $p =: p_0 \approx 1.42948$ in $(0,2)$. A further calculation reveals $h_3''(p_0) < 0$, indicating p_0 is the point of maxima and

$$S(p,1,y) \leq S(p_0,1,y) \approx 0.319595,\ p \in (0,2) \text{ and } y \in (0,1).$$

On the face $y = 0$, $S(p,x,y)$ becomes

$$S(p,x,0) =: h_4(p,x) = (4-p^2)(144x(1-x^2)+p^4x(3+x-x^2-x^3)+4p^2(9-9x+2x^2+19x^3+x^4)).$$

A computation yields

$$\begin{aligned}\frac{\partial h_4}{\partial p} =& 2p(3p^4x(-3-x+x^2+x^3)+16(9-18x+2x^2+28x^3+x^4)\\ &-8p^2(9-12x+x^2+20x^3+2x^4))\end{aligned}$$

and

$$\frac{\partial h_4}{\partial x} = (4-p^2)(144(1-3x^2)+p^4(3+2x-3x^2-4x^3)+4p^2(-9+4x+57x^2+4x^3)).$$

We observe that there is no common solution for the equations $\frac{\partial h_4}{\partial p} = 0$ and $\frac{\partial h_4}{\partial x} = 0$, which indicates there does not exist any critical point of $h_4(p,x)$ in $(0,2) \times (0,1)$.

On the face $y = 1$, $S(p,x,y)$ becomes $S(p,x,1)$, given as

$$\begin{aligned}h_5(p,x) :=& \frac{1}{1152}\Big((1-x^2)(512+64x^2+16px(20+4x)+p^2(176x^2+160x^3+16x^4-256+144x\\ &-32x^2)+p^4(12x-40x^2-44x^3-8x^4+32-36x+4x^2)+p^3(48+144x-8x(20+4x))\\ &+p^5(-12-36x+x(20+4x))+p^6(-3x-x^2+x^3+x^4))\Big).\end{aligned}$$

On solving $\frac{\partial h_5(p,x)}{\partial x} = 0$ and $\frac{\partial h_5(p,x)}{\partial p} = 0$, we observe that there is no common solution to these equations, hence there does not exist any critical point of h_5 in $(0,2) \times (0,1)$.

III. Finally, we find the maximum values attained by $S(p,x,y)$ on the edges of the cuboid $\mathfrak{C}$. The equations (2.12),(2.13),(2.14) and (2.16) are appropriately used to evaluate $S(p,x,y)$ below for particular values of p, x and y.

(i) $S(p,0,0) = p^2(4-p^2)/32 =: l_1(p)$. Now, $l_1'(p) = 0$ for $p = 0$ and $p =: \gamma_0 = \sqrt{2}$. Simply by second derivative test, we obtain $p = 0$ is the point of minima and maximum value $1/8$ is attained at γ_0. So, we have

$$S(p,0,0) \leq \frac{1}{8}, \quad p \in [0,2].$$

(ii) $S(p,0,1) = (4-p^2)(32-8p^2+3p^3)/288$, which is a decreasing function of p in the given range of p. Thus maximum value is obtained at $p = 0$ and

$$S(p,0,1) \leq S(0,0,1) = \frac{4}{9},\ p \in [0,2].$$

(iii) Since $S(p,1,y)$ is independent of y, we obtain $S(p,1,0) = S(p,1,1) = p^2(176 - 40p^2 - p^4)/576 = h_3(p)$, given in (2.16). Thus

$$S(p,1,0) = S(p,1,1) \leq 0.319595,\ p \in [0,2].$$

(iv) $S(0,0,y) = 4y^2/9$, clearly which attains maximum value $4/9$ at $y = 1$. So

$$S(0,0,y) \leq \frac{4}{9},\ y \in [0,1].$$

(v) $S(0,1,y) = S(2,0,y) = S(2,1,y) = 0,\ y \in [0,1]$.

(vi) $S(0,x,0) = x(1-x^2)/2 =: l_3(x)$. Now $l_3'(x) = (1-3x^2)/2 = 0$ gives $x = \gamma_1 := 1/\sqrt{3}$ in the interval $[0,1]$. Further second derivative of $l_3(x)$ is negative at γ_1. Thus γ_1 is the point of maxima and

$$S(0,x,0) \leq \frac{1}{3\sqrt{3}} = 0.19245,\ x \in [0,1].$$

(vii) $S(0,x,1) = (1-x^2)(x^2+8)/18$, which is a decreasing function of x in $[0,1]$. So clearly maximum value is attained at $x = 0$ and we have

$$S(0,x,1) \leq \frac{4}{9},\ x \in [0,1].$$

(viii) $S(2,x,0) = S(2,x,1) = 0,\ x \in [0,1]$.

Taking into account these cases I-III, the inequality (2.7) is proved. Consider the function $\tilde{f} : \mathbb{D} \to \mathbb{C}$ as follows

$$\tilde{f}(z) = z \exp\left(\int_0^z \frac{\left(\frac{1+t^3}{1-t^3}\right) - 1}{t} dt\right) = z + \frac{2z^4}{3} + \cdots,$$

clearly belongs to $\mathcal{S}^*$ and for which, we have $a_2 = a_3 = a_5 = 0$ and $a_4 = 2/3$. This shows the bound $|H_3(1)|$ is sharp as equation (1.1) yields $|H_3(1)| = 4/9$ for this function. ∎

Acknowledgment. The first author is supported by a Research Fellowship from the Department of Science and Technology, New Delhi (Ref No. IF170272).

References

[1] Ali, R.M., Cho, N.E., Ravichandran, V., Kumar, S.S., *Differential subordination for functions associated with the lemniscate of Bernoulli*, Taiwanese J. Math. **16**(2012), no. 3, 1017–1026.

[2] Ali, R.M., Jain, N.K., Ravichandran, V., *Radii of starlikeness associated with the lemniscate of Bernoulli and the left-half plane*, Appl. Math. Comput. **218**(2012), no. 11, 6557–6565.

[3] Babalola, K.O., *On H3(1) Hankel determinants for some classes of univalent functions*. In: Dragomir, S.S., Cho, J.Y. (eds.) Inequality theory and applications, **6**(2010), 1–7. Nova Science Publishers, New York.

[4] Banga, S, Kumar S.S., *The sharp bounds of the second and third Hankel determinants for the class $\mathcal{SL}^*$*, Mathematica Slovaca **70**(2020), no. 4, 849-862.

[5] Cho, N. E., Kumar, V., *Initial coefficients and fourth Hankel determinant for certain analytic functions*, Miskolc Mathematical Notes. **21**(2020), no. 2, 763–779.

[6] Dienes, P.: *The Taylor series: an introduction to the theory of functions of a complex variable*, Dover Publications, Inc., New York, 1957.

[7] Ismail, M.E.H., Merkes, E., Styer, D., *A generalization of starlike functions*, Complex Variables, Theory and Application: An International Journal, **14**(1990), no. 1–4, 77–84.

[8] Jackson, D.O., Fukuda, T., Dunn, O., Majors, E., *On q-definite integrals*, In Quart. J. Pure Appl. Math.(1910)

[9] Janowski, W., *Extremal problems for a family of functions with positive real part and for some related families*, Ann. Polon. Math. **23**(1970/1971), no. 2, 159–177.

[10] Khan, N., Shafiq, M., Darus, M., Khan, B., Ahmad, Q., *Upper bound of the third Hankel determinant for a subclass of q-starlike functions associated with Lemniscate of Bernoulli*, J. Math. Inequal, **14**(2020), no. 1.

[11] Kowalczyk, B., Lecko,A., Lecko, M., Sim, Y. J., *The sharp bound of the third Hankel determinant for some classes of analytic functions*, Bull. Korean Math. Soc. **55**(2018), no. 6, 1859–1868.

[12] Kowalczyk, B., Lecko,A., Lecko, M., Sim, Y. J., *The sharp bound for the Hankel determinant of the third kind for convex functions*, Bull. Aust. Math. Soc. **97**(2018), no. 3, 435–445.
[13] Kumar, S. S., Kumar, V., Ravichandran, V., Cho, N. E., *Sufficient conditions for starlike functions associated with the lemniscate of Bernoulli*, J. Inequal. Appl. **2013**, 2013:176, 13 pp.
[14] Kwon, O. S., Lecko, A., Sim, Y. J., *On the fourth coefficient of functions in the Carathéodory class*, Comput. Methods Funct. Theory **18**(2018), no. 2, 307–314.
[15] Kwon, O. S., Lecko, A., Sim, Y. J., *The bound of the Hankel determinant of the third kind for starlike functions*, Bull. Malays. Math. Sci. Soc. **42**(2019), no. 2, 767–780.
[16] Lecko, A., Sim, Y.J., ´Smiarowska, B., *The sharp bound of the Hankel determinant of the third kind for starlike functions of order 1/2*, Complex Anal. Oper. Theory **13**(2019), no. 5, 2231–2238.
[17] Libera, R. J., Złotkiwicz, E. J., *Early coefficients of the inverse of a regular convex function*, Proc. Amer. Math. Soc. **85**(1982), no. 2, 225–230.
[18] Mahmood, S., Jabeen, M., Malik, S.N., Srivastava, H.M., Manzoor, R., Riaz, S.M. *Some coefficient inequalities of q-starlike functions associated with conic domain defined by q-derivative*, Journal of Function Spaces, 2018.
[19] Mendiratta, R., Nagpal, S., Ravichandran, V., *On a subclass of strongly starlike functions associated with exponential function*, Bull. Malays. Math. Sci. Soc. **38**(2015), no. 1, 365–386.
[20] Noor, K. I., *Hankel determinant problem for the class of functions with bounded boundary rotation*, Rev. Roumaine Math. Pures Appl. **28**(1983), no. 8, 731–739.
[21] Pommerenke, C.: *On the coefficients and Hankel determinants of univalent functions*, J. London Math. Soc. **1**(1966), no. 1, 111–122.
[22] Ravichandran, V., Verma, S., *Bound for the fifth coefficient of certain starlike functions*, C. R. Math. Acad. Sci. Paris **353**(2015), no. 6, 505–510.
[23] Raza, M., Malik, S. N., *Upper bound of the third Hankel determinant for a class of analytic functions related with lemniscate of Bernoulli*, J. Inequal. Appl. **2013**(2013), no. 412, 8 pp.
[24] Robertson, M.S., *Certain classes of starlike functions*, Michigan Math. J. **32**(1985), no. 2, 135–140.
[25] Rønning, F., *Uniformly convex functions and a corresponding class of starlike functions*, Proc. Amer. Math. Soc. **118**(1993), no. 1, 189–196.
[26] Sokół, J., *Radius problems in the class $\mathcal{SL}^*$*, Appl. Math. Comput. **214**(2009), no. 2, 569–573.
[27] Sokół, J., Stankiewicz, J., *Radius of convexity of some subclasses of strongly starlike functions*, Zeszyty Nauk. Politech. Rzeszowskiej Mat. No. **19**(1996), 101–105.
[28] Srivastava, H.M., Deepak, B., *Close-to-convexity of a certain family of q-Mittag-Leffler functions*, J. Nonlinear Var. Anal **1**(2017), no. 1, 61-69.
[29] Zaprawa, P., *Second Hankel determinants for the class of typically real functions*, Abstr. Appl. Anal. **2016**, Art. ID 3792367, 7 pp.
[30] Zaprawa, P., *Third Hankel determinants for subclasses of univalent functions*, Mediterr. J. Math. **14**(2017), no. 1, Art. 19, 10 pp.
[31] Zaprawa, P., Milutin, O., Tuneski, N., *Third Hankel determinant for univalent starlike functions*, Revista de la Real Academia de Ciencias Exactas, Físicas y Naturales. Serie A. Matemáticas **115**(2021), no. 2, 1-6.

Department of Applied Mathematics, Delhi Technological University, Delhi–110042, India
Email address: shagun05banga@gmail.com

Department of Applied Mathematics, Delhi Technological University, Delhi–110042, India
Email address: spkumar@dce.ac.in

SHORT TERM EFFECT OF PHYSICAL EXERCISE ON SELECTIVE ATTENTION USING EEG AND STROOP TASK

1st Divyashikha Sethia
Department of Computer Engineering
Delhi Technological University
New Delhi, India
divyashikha@dtu.ac.in

2nd Taksh Kamlesh
Department of Computer Engineering
Delhi Technological University
New Delhi, India
taksh.pathak@gmail.com

3rd Mrigank Singh
Department of Computer Engineering
Delhi Technological University
New Delhi, India
mriganksingh1@gmail.com

4th Sonia Baloni Ray
Department of Social Sciences and Humanities
IIIT Delhi
New Delhi, India
sonia@iiitd.ac.in

2022 14th International Conference on COMmunication Systems & NETworkS (COMSNETS) | 978-1-6654-2104-1/22/$31.00 ©2022 IEEE | DOI: 10.1109/COMSNETS53615.2022.9668373

***Abstract*—This paper aims to investigate the short-term effect of physical exercise on Selective Attention for university students. Selective attention allows the participant to focus on essential details ignoring the other trivial information. Stroop task is a test widely used to study Selective Attention. EEG waves are very closely related to attention and can help investigate the effect of physical exercise on attention. We propose a methodology to investigate High-Intensity interval (HIIT) Exercise (Tabata routine) on Selective Attention on university students. We prove that Response time statistically after the Post-HIIT Stroop Task is significantly less than that of the pre-HIIT Stroop Task. We analyze the pre-HIIT Stroop test EEG data and post-HIIT Stroop test EEG data to get insights using EEG time-domain features: mean, variance, kurtosis, skewness, Shannon entropy, and total power. We observe that the post-HIIT variance for all channels and total power for the Fpz channel is higher, while pre-HIIT Shannon Entropy for all channels is higher for all 10 participants. Results indicate that exercise has a positive impact on stimulating the brain. Hence it impacts selective attention, as also indicated by the statistical analysis from the Stroop test results.**

***Index Terms*—Electroenchaphologram, HIIT (High-Intensity Interval), Stroop, Selective Attention**

I. Introduction

The behavioural and cognitive process of focusing on a discrete part of information while ignoring the rest is known as attention. It varies from person to person and from event to event. Sohlberg and Mateer's model (1987, 1989) is the most accepted model worldwide, classifying attention as a sub-process group. This model [1] divides attention into the following categories:- Arousal, Focused, Sustained, Selective, Divided, Sustained, and Alternating.

Selective Attention is the attention process that focuses on a single piece of information in the surrounding atmosphere. It allows a person to focus on essential details ignoring the other trivial information. Selective attention involves three stages [2]:-

1) Pre-attentively analyze the properties of a stimulus.
2) Analyze if the given stimulus has a pattern.
3) Focus attention on the stimuli.

This paper aims to analyze the short-term effect of physical exercise on selective attention, which plays a vital role in our day-to-day life. Whether physical activity has an impact or no impact is an essential question of this study, especially in this COVID-19 era. We use the "neuroelectric enobia-8" headset, which measures the Electroencephalogram (EEG) data. In addition, the study uses the Stroop task to measure the selective attention of the participants before and after the HIIT (High-Intensity Interval) Exercise. Stroop Color and Word Test (SCWT) is a psychological test designed by John Ridley Stroop (1935) [3] in which the participant has to tell the colour of the word rather than the word itself. The study uses the Stroop Task provided by the PsytoolKit; an online web-based test creator [4] [5]. This online task contains a consent form for performing the study and then 5 minutes Stroop Color and Word Test.

We statistically prove a significant difference in response time for the pre and post-Stroop tests response time. Hence, results indicate that High-Intensity physical exercise positively affects short-term Selective Attention. Very few researchers have investigated the effect of exercise on short-term Selective attention either through a psychological test such as the Stroop test or measuring the physiological parameters such as EEG.

This study collects the Stroop Task data of 41 participants, and out of these, only 10 participants EEG data due to COVID-19 restrictions as a collection of EEG data requires a visit to

Authorized licensed use limited to: DELHI TECHNOLOGICAL UNIV. Downloaded on January 25,2022 at 08:41:33 UTC from IEEE Xplore. Restrictions apply.

the lab. Nevertheless, the statistic comparison of pre-HIIT and post-HIIT Stroop data shows that the difference between them is statistically significant and indicates a positive impact of exercise on Selective attention.

Data cleaning is necessary to remove various artefacts arising from eye blinks and muscular movements in the EEG data. To remove these artefacts, the model applies the Butterworth filter at a frequency of 0.5 Hz. It then extracts the features of EEG signals in the time-domain, namely: mean, variance, kurtosis, skewness, Shannon entropy, and total power. The comparison shows that post-HIIT variance for all channels and total power for the Fpz channel is greater than the pre-HIIT variance. In contrast, pre-HIIT Shannon Entropy for all channels is higher for all 10 participants. These results show that High-Intensity Interval Exercise stimulates the brain, thus affecting attention.

The rest of the paper comprises section II which describes the related work. Section III discusses the methodology comprising of data collection and artefact removal. Section IV describes the results obtained, followed by conclusion and future in section V.

II. Related Work

The domain of establishing physical exercise concerning cognitive functions, especially selective attention, remains a new field of study for researchers.

Alves et al. [6] showed that acute high-intensity interval exercise leads to a increase in response in time for the Stroop task, which is a measure of Selective Attention. It has better results as compared to the Digi Span Test (a short-term memory test). The exercise routine for experimentation begins with 10 minutes of cycling as physical exercise followed by stretching exercise as a control.

Lebeau et al. [7] analyzed the effects of acute exercise on older adults. The Control Performance Task (CPT) designed for this is a car driving simulator and other tasks, including Stroop Task and Trail Making Test. The Trail Making Test is a neuropsychological trial of visual consideration and task switching. It comprises of two sections in which the subject is told to associate a bunch of 25 specks as fast as could be expected while as yet keeping up with precision. The authors find no significant improvement in the simulated driving task before and after the acute exercise routine. However, there is a substantial improvement in the Stroop task reaction time and Trial Making Test after the exercise routine.

Aly et al. [8] contributed to this field by proving that regular physical activity contributes to better neural efficiency in resource allocation to tasks. The study divides the participants into two groups. The first group consists of active participants exercising regularly. The second group consists of less active participants who either do not perform exercise or do exercise irregularly. The results show that the active group's Response Time (RT) is better than that of the less active group.

Yanagisawa et al. [9] showed that acute moderate exercise improves the Stroop Task response time with increased oxygen flow in the brain's dorsolateral prefrontal lobe. The authors studied the cortical activation of the brain using multichannel functional near-infrared spectroscopy (fNIRS).

Endo et al. [10] studied whether the dynamic exercises affect cognitive functions and improved blood flow in the brain's prefrontal cortex. The authors use the Stroop task before and after ergometer exercises performed for 15 minutes at 20, 40, and 60% of maximum voluntary activity to measure cognitive functions and near-infrared Spectroscopy to determine regional cerebral blood flow changes. The results indicate that the Ergometer exercise only at a moderate level of 40% has a better Stroop response. At the 40% level, there is also an increase in oxy-haemoglobin in the prefrontal cortex during the Stroop task. However, the deoxyhemoglobin in those areas remains unchanged at all levels.

None of the previous research papers investigate the impact of short-term exercise using both Stroop test and EEG together as done in this paper.

III. Methodology

This section deals with a detailed explanation of the methodology used in the study. Fig. 1. shows the proposed methodology.

A. Data Collection

The study comprises EEG data of 10 participants and Stroop Task data of 41 participants aged 18 to 25. The participants have neither taken caffeine intake in any form in the past 24 hours nor have done physical exercise, which might interfere with data readings. The participants filled a consent form before participating in the experimental procedure. They gave their consent that they do not have any illness or diseases that might affect their health after performing High-Intensity Interval Training. They are also reminded from time to time to stop the exercise if they feel too much fatigue and cannot continue.

Each participant wore a "Neuroelectric Enobia-8" [11] headset to measure the EEG data. The channels used in this study in the "neuroelectric enobia 8" headset consist mainly of channels Fp1, Fpz, Fp2, F3, Fz, F4, F6, F7, which are a part of the brain's prefrontal lobe. The main reason for selecting these channels is the vivid research done in the field asserting that physical exercise leads to maximum effect in the prefrontal region [9] [10].

The experiment routine is approximately around 40 to 45 minutes. The test starts once the headset configures correctly and all eight channels are green. Fp1, Fpz, Fp2, F3, Fz, F4, F6, F7 are the eight channels recognized with increased cognitive functions during the acute moderate and dynamic exercise.

The participants, after the Stroop test, are allowed to rest for 5 minutes. After the resting period, a High-Intensity Interval

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:41:33 UTC from IEEE Xplore. Restrictions apply.

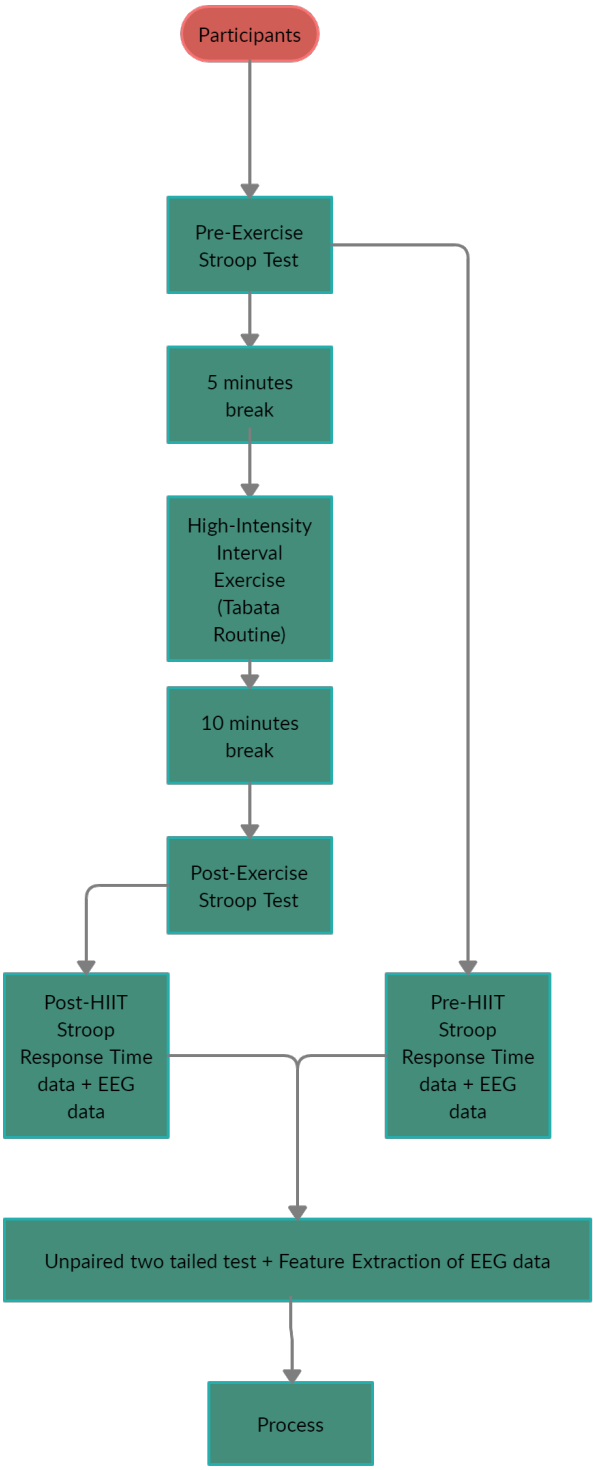


Fig. 1. Overall Workflow

Exercise routine (HIIT) is there. The routine has four high-intensity exercises: burpees, mountain climbers, jumping jacks, and squat jumps. There are four sets of exercises of 20 seconds, with a break of 10 seconds after every set and 30 seconds of rest between exercises. The total time for HIIT is between 8 to 10 minutes. Continuous monitoring of EEG cap connectivity and everything is maintained throughout the process to ensure fewer data collection errors.

After the workout routine, the participant gets 10 minutes of break and can come to a normal state. A post-HIIT Stroop Test is there of 5 minutes to measure the selective attention after HIIT. After the data collection process completes successfully, the neuroelectric headset and watch are removed and appropriately sanitized. This whole process is represented graphically in Fig. 1.

B. Artefact Removal

Artifact removal is a necessary pre-processing step while analyzing signal data. A large number of artefacts imply noisy data, which is not suitable for any form of analysis. In this work, we utilize a first-order high-pass Butterworth filter for the process of artefact removal. A Butterworth filter [12] functions as a bandpass filter where the frequency response is as flat as possible. The cutoff frequency for this high-pass filter is 0.5 Hz.

C. Feature Extraction

We measure statistical features such as mean, standard deviation, standard error, response time, kurtosis, skewness, non-linear energy, and Shannon entropy on the time-domain of the signal, separately for pre-HIIT and post-HIIT data and compare them. We then apply a bandpass filter on the signal to obtain the frequency bands corresponding to human EEG waves: alpha, beta, theta, delta, and gamma waves. We calculate the total power on each wave, again separately for pre-HIIT and post-HIIT data. We perform all these calculations individually for all 8 channels FP1, FP2, FPZ, F3, FZ, F4, F7, and F8.

IV. Results

The study compares the time-domain feature values on EEG as well as the Response Time values observed during the post-HIIT Stroop test and pre-HIIT Stroop test for all participants, for each participant channels separately.

It is observed from the pre-HIIT and post-HIIT Stroop task Response time that the Post-HIIT Stroop Task Average Response Time is shorter than the Pre-Stroop Average Response Time for all participants indicating that there is an increase in the Reaction Time of the participants.

To perform the paired two-tailed test, we take the null hypothesis that the mean of the two groups is equal, i.e., there is no impact on Selective attention. The t-value is 7.2660, and the paired two-tailed value is less than 0.0001, which is statistically significant. The result of all the above calculations are mentioned in TABLE I.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:41:33 UTC from IEEE Xplore. Restrictions apply.

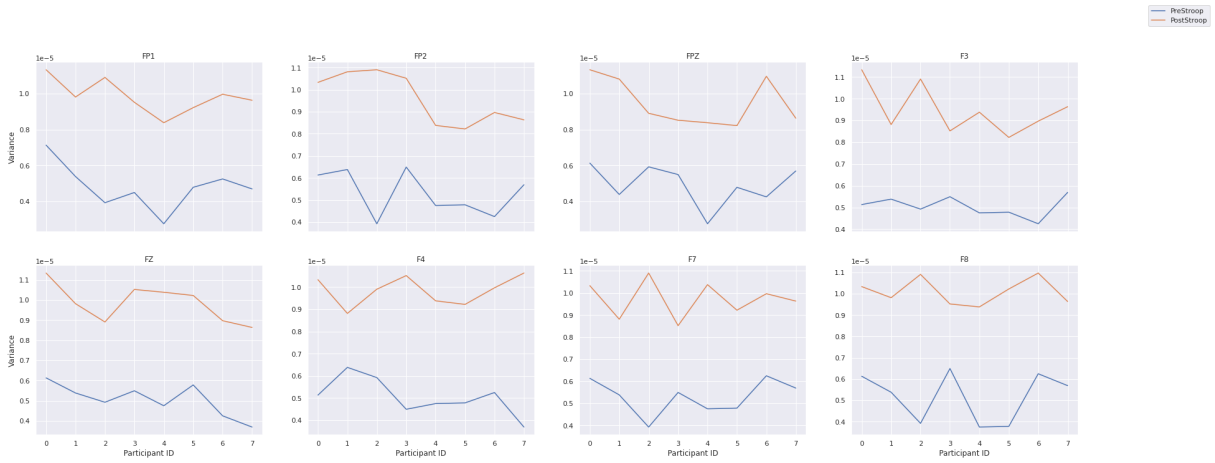


Fig. 2. Variance across 8 participants for all 8 channels

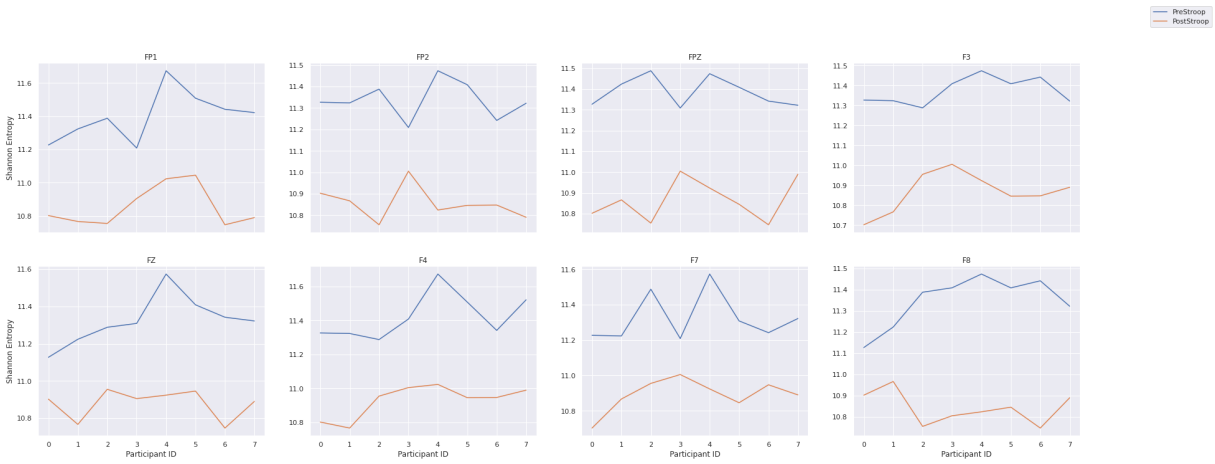


Fig. 3. Shannon Entropy across 8 participants for all eight channels

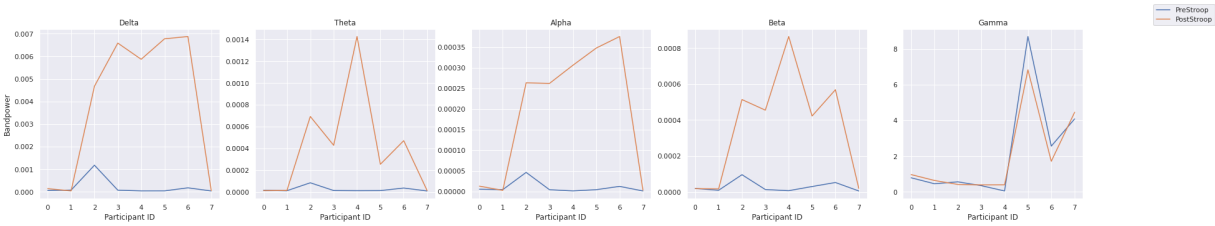


Fig. 4. Bandpower of all 5 EEG bandwidths for channel FpZ

TABLE I
STATISTICAL DATA OF PRE AND POST HIIT STROOP TASK

Statistical Measure	Pre-HIIT Stroop RT in ms	Post-HIIT Stroop RT in ms
Mean	862.63700	803.85528
Standard-Deviation	74.62694	63.50054
Standard Error of Mean	11.79956	10.04032
Number of Samples	40	40

We monitor trends for each feature value to indicate a change in EEG signal output for features variance, Shannon Entropy, and Bandpower. None of the previous research papers investigate the impact of short-term exercise using both the Stroop test and EEG together, as done in this paper. The first important observation is that the post-HIIT Stroop EEG data variance for eight channels is greater than that for the pre-HIIT Stroop data. This trend is consistent for all 8 participants. It implies that the amplitude of the EEG signals deviate more from their expected values during the post-HIIT Stroop period

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:41:33 UTC from IEEE Xplore. Restrictions apply.

than the pre-HIIT Stroop period. Fig 2 shows the comparison between the pre-HIIT and post HIIT Stroop data variance.

The second consistent trend observed is that the Shannon Entropy of pre-HIIT Stroop EEG data is greater than that of post-HIIT Stroop EEG data. As Shannon Entropy indicates how easy it is to "guess the value of a variable," this feature implies the opposite result of the variance. It further reinforces the fact that variance is higher for post-HIIT Stroop EEG data. There is more randomness in EEG signals recorded during post-HIIT Stroop duration than in pre-HIIT Stroop duration. Fig 3 shows the comparison between the pre-HIIT and post HIIT Stroop data Shannon Entropy.

The third observation is that the total signal power of post-HIIT Stroop data in the FpZ channel is higher than that of pre-HIIT Stroop data. This trend is consistent with Alpha, Beta, Theta, and Delta bands but not the Gamma band. Fig 4 shows the analysis done on the pre-HIIT and post HIIT Stroop data total power. According to Kim, [13], frontal lobes of the brain are primarily responsible for attention. Since Fpz corresponds to the center of the frontal lobe, we can conclude that attention is greater after HIIT.

The statistical comparison of Stroop data and EEG data analysis shows a significant increase in short-term selective attention for the participants after the High-Intensity Interval Exercise. The EEG data shows that post-HIIT, there is stimulation in the brain, thus affecting attention. The analysis shows that the effect is positive.

V. Conclusion and Future Work

This paper investigates the short-term impact of physical activity on Selective attention, using the Stroop task and EEG neuroelectric enobia-8 headset for ten university participants. The High-Intensity Interval Training routine involves the famous Tabata routine of burpees, jumping jacks, mountain climbers, and squat jumps. We recorded EEG data pre-HIIT and post-HIIT and cleaned that data using the Butterworth bandpass filter.

Results indicate that the post-HIIT Stroop Task response time is less than that of pre-HIIT Stroop Task Response Time, and statistically using the paired two-tailed test, we proved that the difference in response time is very significant. In addition, the post-HIIT Stroop EEG data variance is higher than the variance for pre-HIIT Stroop data for each participant. However, the reverse is true when comparing the Shannon entropy. These two observations signify that there is more randomness in post-HIIT Stroop signal data.

While contrasting different EEG bands (Alpha, Beta, Theta, Delta, and Gamma), we determine that for all EEG bands except the gamma band, the total power of the signal during post-HIIT duration is higher than the total power during the pre-HIIT duration. This trend is consistent for all participants for channel FpZ. Since the Fpz channel corresponds most to attention, we can say that attention is stronger after HIIT. All results prove a significant increase in short-term selective attention for the participants after the High-Intensity Interval Exercise.

Our work until now has given promising results, but on some of the parameters, either we have obtained zero results or very vague results. There can be two possible causes for these results. The first reason can be that our control time taken for rest after the exercise and before the post-HIIT Stroop task is too high (10 minutes). The second reason can be that our exercise intensity is too high, and as Endo et al. [10] proved that too high-intensity exercise gave no results in Stroop test response time. Therefore, we aim to modify our experiment procedure in the future by decreasing the rest time and increasing the number of participants in the study.

References

[1] "Attention one of our cognitive domains," last Accessed March, 2021. [Online]. Available: t.ly/MbKz

[2] L. Madrigal and P. Romero, "Selective attention," 2014, last Accessed March, 2021. [Online]. Available: https://www.slideshare.net/gsjus/attention-34276525

[3] J. R. Stroop, "Studies of interference in serial verbal reactions." *Journal of Experimental Psychology: General*, vol. 121, no. 1, p. 15, 1992.

[4] G. Stoet, "Psytoolkit: A software package for programming psychological experiments using linux," *Behavior research methods*, vol. 42, no. 4, pp. 1096–1104, 2010.

[5] ——, "Psytoolkit: A novel web-based method for running online questionnaires and reaction-time experiments," *Teaching of Psychology*, vol. 44, no. 1, pp. 24–31, 2017.

[6] Alves *et al.*, "Influence of acute high-intensity aerobic interval exercise bout on selective attention and short-term memory tasks," *Perceptual and motor skills*, vol. 118, no. 1, pp. 63–72, 2014.

[7] Lebeau *et al.*, "The effects of acute exercise on driving and executive functions in healthy older adults," *International Journal of Sport and Exercise Psychology*, pp. 1–19, 2020.

[8] M. Aly and H. Kojima, "Relationship of regular physical activity with neuroelectric indices of interference processing in young adults," *Psychophysiology*, vol. 57, no. 12, p. e13674, 2020.

[9] Yanagisawa *et al.*, "Acute moderate exercise elicits increased dorsolateral prefrontal activation and improves cognitive performance with stroop test," *Neuroimage*, vol. 50, no. 4, pp. 1702–1710, 2010.

[10] Endo *et al.*, "Dynamic exercise improves cognitive function in association with increased prefrontal oxygenation," *The Journal of Physiological Sciences*, vol. 63, no. 4, pp. 287–298, 2013.

[11] "Neurelectric enobia-8 headset." [Online]. Available: https://www.neuroelectrics.com/solutions/enobio/8

[12] S. Daud and R. Sudirman, "Butterworth bandpass and stationary wavelet transform filter comparison for electroencephalography signal," pp. 123–126, 2015.

[13] Kim *et al.*, "Biosignal based attention monitoring to support nuclear operator safety-relevant tasks," *Frontiers in computational neuroscience*, vol. 14, p. 111, 2020.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:41:33 UTC from IEEE Xplore. Restrictions apply.

Single OTRA-Based Multifunctional Filter Configuration

Garima Mann
Electrical Engineering Department
Delhi Technological University
Delhi, India
garimasolanki@dtu.ac.in

Pragati Kumar
Electrical Engineering Department
Delhi Technological University
Delhi, India
pragatikumar@dtu.ac.in

D. R. Bhaskar
Electrical Engineering Department
Delhi Technological University
Delhi, India
drbhaskar@dtu.ac.in

2021 7th International Conference on Signal Processing and Communication (ICSC) | 978-1-6654-2739-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICSC53193.2021.9673176

***Abstract*— This paper presents a new single-input single-output (SISO) type voltage mode filter configuration that employ single operational transresistance amplifier (OTRA) and four admittances. The proposed configuration is capable of realizing second order low pass filter (LPF), high pass filter (HPF) and band pass filter (BPF) using different choices of branch admittances. The proposed filter responses have been verified through PSPICE simulations using CMOS OTRA implemented in 0.18µm TSMC technology and also validated using MATLAB evaluations.**

Keywords—Voltage Mode Filters, Operational Transresistance Amplifier, Voltage Mode

I. INTRODUCTION

Over the years, analog active filters have played a vital role in the implementation of continuous time signal processing circuits. These filters can be broadly classified as SISO type [1], multiple-input multiple-output (MIMO) type [1], single-input multiple-output (SIMO) type [2], and multiple-input single-output (MISO) type [3]. In addition to the classical operational amplifiers, various active building blocks (ABBs) viz. current conveyors and its variants (CC), current feedback operational amplifiers (CFOA), current differencing buffered amplifiers (CDBA), operational transconductance amplifiers (OTA), current differencing transconductance amplifiers (CDTA), OTRA and many more have been used in the literature for designing such active filters [1-6].

Because of the presence of virtual ground at the input terminals, the effect of parasitics is removed at the input terminals of OTRA due to which, OTRA has received prominent attention in the design of filters, oscillators and non-linear analog applications. Also, its low output impedance facilitates easy cascadability. Therefore, various filters using OTRA as an active block, have been reported in the literature [7-30]. These filters can further be classified as single OTRA-based [7-20] and multiple OTRA-based [21-30] filter configurations. Since, the proposed work deals with single OTRA-based biquads, it is worthwhile to present a detailed overview of single OTRA-based biquad filters presented by various authors so that the proposed work is put in the proper perspective.

In [7], Kilinc and Cam presented a generalized single OTRA-based structure employing three admittances. This circuit provides a first order and second order all pass filters (APF), alongwith a notch filter by proper choice(s) of admittances. Cakir et. al in [8], reported a APF configuration using single OTRA and four admittances. It may be mentioned that in this circuit, the intrinsic property of OTRA has not been used. Yet, another inverting and non-inverting APF structure was proposed by Cam et. al in [9], using single OTRA, two resistors and a capacitor. Three, current mode first order APF topologies, were proposed by Kacar in [10]. Two of these circuits use one resistor and one capacitor while the third circuit uses two resistors and one capacitor to realize first order APF functions. In [11], Ranjan et. al realized a band pass filter of order three employing four resistors and four capacitors. Since the circuit uses four capacitors which is non-canonic for the realization of a third order filter. A MISO type universal filter configuration of order one using single OTRA was proposed by Gahalawat et. al in [12]. In [13], Kilinc et. al reported a MISO type single OTRA-based second order universal filter structure employing as many as eight passive components (four resistors and four capacitors). In [14], Bhatt et. al presented a MISO type single OTRA-based second order universal filter based on RC-CR decomposition technique. In [15], a single OTRA-based biquad configuration employing three resistors and three capacitors was reported by Gokcen and Cam to realize second order filter functions which required component matching constraints. Pandey et. al [16] reported a second order filter topology employing a single OTRA (utilizing only one terminal of OTRA while the other terminal was left open) and five admittances. A generalized n^{th} order filter configuration was proposed by Kilinc and Cam in [17] using five resistors and three capacitors to realize third APF and third order LPF responses. A third order MISO type universal filter topology employing five resistors and six capacitors was presented by Ghosh et. al in [18] and Ghosh et. al in [19]. Tewari and Arya in [20] presented a fractional order LPF of order 1.5 using a single OTRA.

A careful perusal of the different single OTRA-based filters presented above reveals that:

(i) The number of passive components required for a particular filter function are non-canonic [7,11,13,14,17,18,19]
(ii) Requirement of component matching constraints for the realization of different filter functions [8,13,14,15,17,18]
(iii) Utilization of all input terminals of OTRA was not used since one of the input terminals is either left open or grounded [8,10,16]

Thus, the main aim of the proposed work is to present a new single OTRA-based second order filter configuration employing canonic number of capacitors and relatively lower number of resistors as compared to previously reported filters utilising the intrinsic property of the OTRA without

978-1-6654-2739-5/21/$31.00 ©2021 IEEE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:37:37 UTC from IEEE Xplore. Restrictions apply.

imposing any component matching constraints (The term canonic is used for the proposed filter structure with minimum number of components, active and passive). Workability of the presented filter structure has been verified through PSPICE simulations and MATLAB evaluations.

II. Proposed Configuration

In Fig. 1, the circuit notation of an OTRA is displayed.

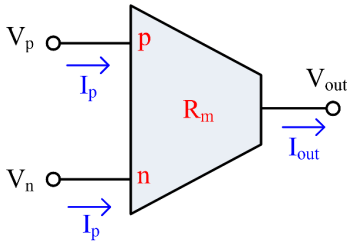


Fig. 1 Symbol of an OTRA

The port relations of OTRA are described by the matrix equation as:

$$\begin{bmatrix} V_p \\ V_n \\ V_{out} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ R_m & -R_m & 0 \end{bmatrix} \begin{bmatrix} I_p \\ I_n \\ I_0 \end{bmatrix} \quad (1)$$

From equation (1) the output voltage is given by:

$$V_o = R_m(I_p - I_n)$$

where R_m = transresistance gain. For an ideal OTRA, $R_m = \infty$, which makes both the input currents equal.

The proposed filter topology is shown in Fig. 2.

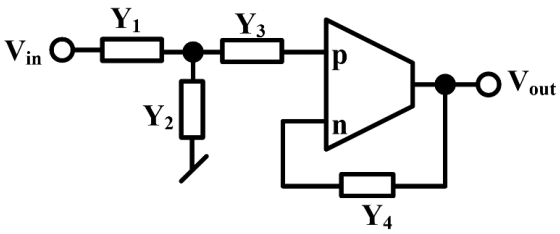


Fig. 2 Proposed filter configuration

The transfer function for the proposed structure is given by equation (2) (assuming ideal OTRA):

$$\frac{V_{out}}{V_{in}} = \frac{Y_1 Y_3}{Y_4(Y_1+Y_2+Y_3)} \quad (2)$$

From equation (2), various 2nd order filter responses can be obtained by appropriate selection of branch admittances as given below:

Case 1: If we select $Y_1=G_1$, $Y_2=sC_2$, $Y_3=G_3$ and $Y_4=sC_4+G_4$, a 2nd order LPF can be obtained whose transfer function is given by:

$$\frac{V_{out}(s)}{V_{in}(s)} = \frac{\frac{G_1 G_3}{C_2 C_4}}{s^2 + s\left(\frac{G_1+G_3}{C_2} + \frac{G_4}{C_4}\right) + \frac{G_4}{C_2 C_4}(G_1+G_3)} \quad (3)$$

The pole frequency (ω_o), pole quality factor (Q) and gain (H_o) of 2nd order LPF are given by:

$$\omega_o = \sqrt{\frac{G_4(G_1+G_3)}{C_2 C_4}} \qquad Q = \frac{\sqrt{C_2 C_4 G_4 (G_1+G_3)}}{C_4(G_1+G_3)+G_4 C_2}$$
$$H_o = \frac{G_1 G_3}{G_4(G_1+G_3)} \quad (4)$$

Case 2: For $Y_1=G_1$, $Y_2=G_2$, $Y_3=sC_3$ and $Y_4=sC_4+G_4$, equation (2) yields to:

$$\frac{V_{out}(s)}{V_{in}(s)} = \frac{\left(\frac{G_1}{C_4}\right)s}{s^2 + s\left(\frac{G_1+G_2}{C_3} + \frac{G_4}{C_4}\right) + \frac{G_4}{C_3 C_4}(G_1+G_2)} \quad (5)$$

Hence, for these values of admittances, a BPF can be realized for which ω_o, BW and H_o may be expressed as:

$$\omega_o = \sqrt{\frac{G_4(G_1+G_2)}{C_3 C_4}} \qquad BW = \frac{G_1+G_2}{C_3} + \frac{G_4}{C_4}$$
$$H_O = \frac{G_1 C_3}{C_4(G_2+G_1)+C_3 G_4} \quad (6)$$

Case 3: For HPF function realization, the admittances were selected as: $Y_1=sC_1G_1/(G_1+sC_1)$, $Y_2=G_2$, $Y_3=sC_3$ and $Y_4=G_4$, and the transfer function can be expressed as:

$$\frac{V_{out}(s)}{V_{in}(s)} = \frac{\left(\frac{G_1}{G_4}\right)s^2}{s^2 + s\left(\frac{G_1+G_2}{C_3} + \frac{G_1}{C_1}\right) + \frac{G_1 G_2}{C_1 C_3}} \quad (7)$$

The filter parameters of HPF can be obtained from equation (7) and may be expressed as:

$$\omega_O = \sqrt{\frac{G_1 G_2}{C_1 C_3}} \qquad Q = \frac{\sqrt{C_1 C_3 G_1 G_2}}{G_1(C_1+C_3)+G_2 C_1}$$
$$H_O = \frac{G_1}{G_4} \quad (8)$$

where $G_i = 1/R_i$

III. Non-Ideal Analysis

The consequences of major non-idealities present in the OTRA on the realized filter are considered in this section. We have used the single pole model of trans-resistance gain R_m [31] as given in equation (9) to evaluate the non-ideal filter parameters:

$$R_m(s) = \frac{R_0}{1+\frac{s}{\omega_0}} \quad (9)$$

For high frequency applications, $R_m(s)$ reduces to:

$$R_m(s) = \frac{1}{sC_p} \text{ where } C_p = \frac{1}{R_0 \omega_0} \quad (10)$$

Using the above non-ideal model of the OTRA, the transfer function of the proposed structure of Fig. 2 may be expressed as:

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:37:37 UTC from IEEE Xplore. Restrictions apply.

$$\frac{V_{out}^{/}(s)}{V_{in}^{/}(s)} = \frac{Y_1 Y_3}{Y_4(Y_1+Y_2+Y_3)+sC_p(Y_1+Y_2+Y_3)} \quad (11)$$

From equation (11), the non-ideal transfer functions and parameters of each filter functions can be derived. To conserve space, we present the transfer function of LPF (corresponding to case 1 presented in section II). The non-ideal transfer function of LPF is given by:

$$\frac{V_{out}^{/}(s)}{V_{in}^{/}(s)} = \frac{\frac{G_1 G_3}{C_2(C_4+C_p)}}{s^2 + \frac{As}{C_2(C_4+C_p)} + \frac{G_4(G_1+G_3)}{C_2(C_4+C_p)}} \quad (12)$$

where $A = \left[(G_1+G_3)(C_4+C_p)+C_2G_4\right]$

The non-ideal filter parameters of the LPF may be evaluated from equation (12) and are given by:

$$\omega_{NI} = \omega_o \sqrt{\frac{C_4}{C_4+C_p}}, \; H_{NI}=H_O$$

$$Q_{NI} = \frac{\sqrt{G_4(G_1+G_3)C_2(C_4+C_p)}}{(C_4+C_p)(G_1+G_3)+C_2C_4}$$

It can be seen that the parasitic affects the filter parameters of the proposed filter functions and it may be minimized by selecting the value of passive components such that $C_4 >> C_p$.

IV. Simulation Results and Matlab Evaluations

Using CMOS OTRA implemented in 0.18µm CMOS technology [32] as shown in Fig. 3, the functionality of the presented filter structure has been validated through MATLAB evaluations and PSPICE program. The power supply voltages used were ±1.85V. The bias current (I_B) and bias voltage (V_{B1}) of CMOS OTRA were taken as 18µA and 0.36V respectively. The MOSFETs aspect ratios shown in Fig. 3 have been provided in Table I.

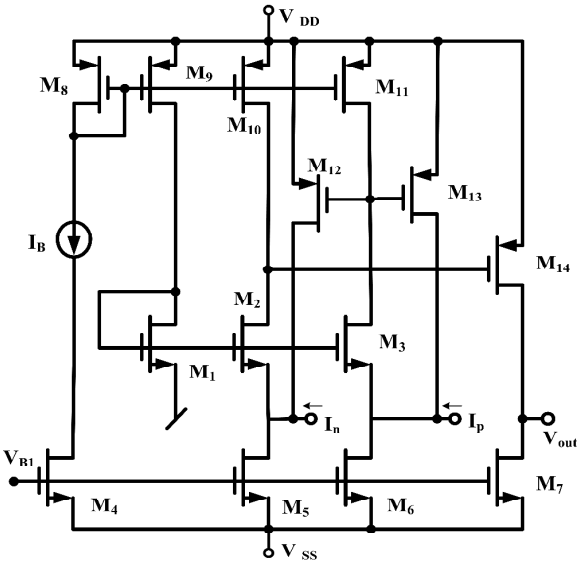
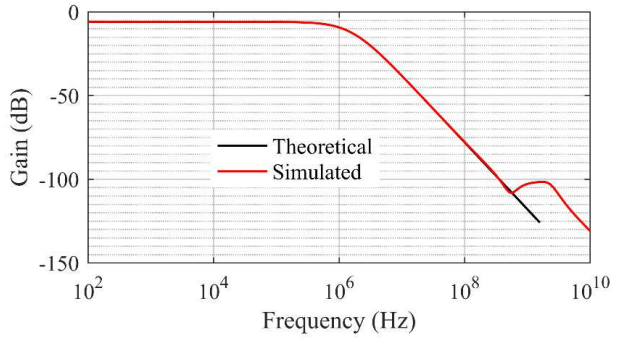


Fig. 3 CMOS Realization of an OTRA [32]

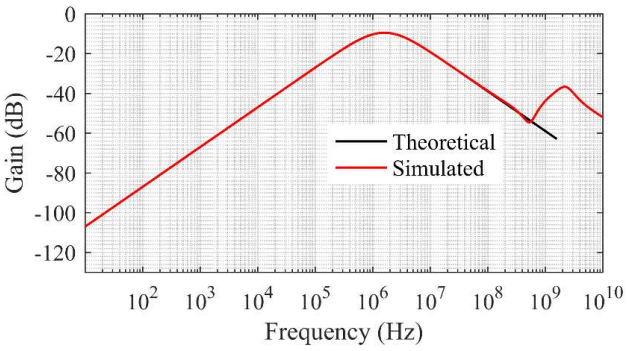
TABLE I MOSFETS ASPECT RATIOS FOR FIG. 3

MOSFETs	Aspect Ratio (W/L) in µM
M_1-M_3, M_{12}, M_{13}	36/0.9
M_4, M_7	3.6/0.9
M_5,M_6	10.8/0.9
M_8-M_9	18/0.9
M_{14}	18/0.18

The proposed filter was designed for a nominal value of the pole frequency/centre frequency of 1.59MHz for LPF, HPF and BPF. The passive component values were taken as: $R_1 = R_3 = R_4 = 14.1$ kΩ and $C_2 = C_4 = 10$ pF (for LPF), $R_1 = R_2 = R_4 = 10$ kΩ and $C_1 = C_3 = 10$ pF (for HPF) and $R_1 = R_2 = R_4 = 14.1$ kΩ and $C_3 = C_4 = 10$ pF (for BPF). For these values of passive components, the DC gain and the Q of LPF, HPF and BPF were obtained as (-6dB and 0.4713), (0dB and 0.33) and (-9dB and 0.4713) respectively. The corresponding values of the -3dB frequency of LPF and HPF are 937.46 kHz and 4.25 MHz respectively while the corresponding value for the bandwidth for the BPF is 3.36 MHz. The simulated values of the 3-dB frequencies for LPF, HPF and BPF were found to be 950KHz, 4.008MHz and 3.24MHz respectively. It may be noted that the 3-dB frequencies of the realized filter are different from the pole frequencies, as the design values of the pole quality factor 'Q' is less than 0.5 (real poles). The simulated PSPICE magnitude responses of different filter functions have been shown in Fig. 4 alongwith MATLAB evaluations. Fig. 5 displays the transient responses of the proposed filter functions for the input signal of 50 mV at different frequencies (100 kHz for LPF, 30.59 MHz for HPF and 1.59 MHz for BPF).

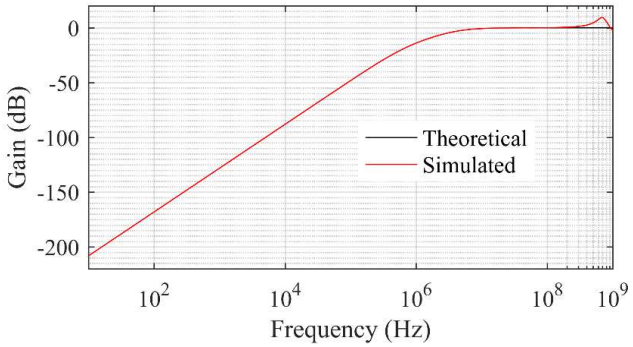


(a)



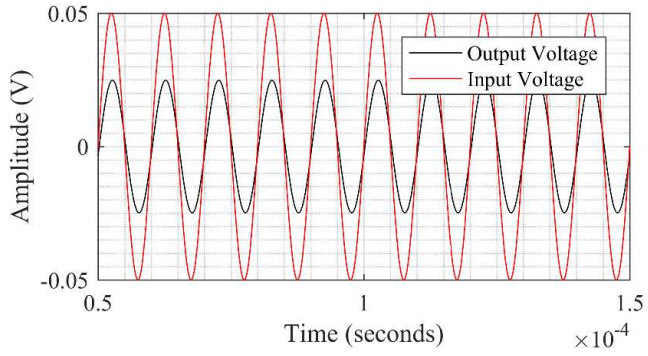
(b)

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:37:37 UTC from IEEE Xplore. Restrictions apply.

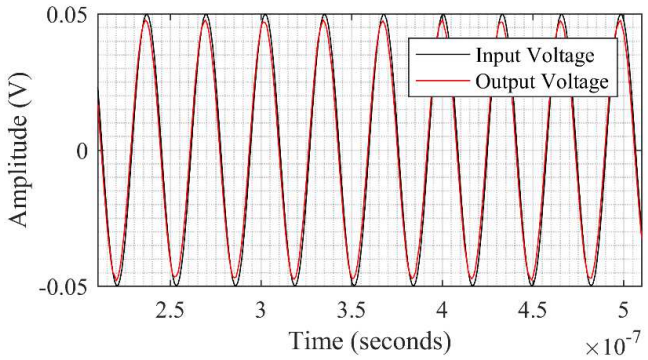


(c)

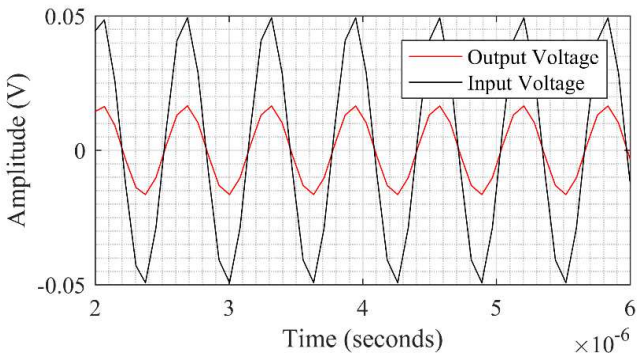
Fig. 4 Gain versus frequency plots (a)LPF (b)BPF and (c)HPF



(a)



(b)



(c)

Fig. 5 Transient responses of (a) LPF (b) HPF and (c) BPF

In Table II, a comparative analysis of the proposed filter configuration with earlier reported filter structures has been demonstrated.

TABLE II COMPARISON OF VARIOUS SINGLE OTRA-BASED SECOND ORDER MULTIFUNCTIONAL FILTERS

REF. NO.	(i)	(ii)	(iii)	(iv)	(v)
[13]	4	8(4R+4C)	No	No	Yes
[14]	4	8(4R+4C)	No	No	Yes
[15]	3	6(3R+3C)	No	Yes	Yes
[16]	5	5(2/3R+3/2C)	Yes	Yes	No
Proposed	4	5(3R+2C)	Yes	Yes	Yes

(i) Number of Admittances (ii) Number of passive components (iii) Realization without component matching (iv) Single Input (v) Use of intrinsic property of OTRA

V. CONCLUSIONS

In this communication, we have presented a new multifunctional biquad filter structure using single OTRA and four admittances. When compared with previously reported single OTRA-based second order filters, the proposed filter structure is the only circuit utilizing intrinsic property of the OTRA to realize second order low pass, high pass and band pass filter responses with canonic number of capacitors. Workability of all the proposed active filter circuit has been established by PSPICE simulations using a CMOS OTRA implemented in 0.18µm CMOS technology and MATLAB evaluations. It is believed that the presented filter configuration will enrich the existing repertoire of single OTRA-based biquad filters.

ACKNOWLEDGEMENT

This work has been performed at Analog Signal Processing Laboratory of Electrical Engineering Department, Delhi Technological University, New Delhi- 110042, India

REFERENCES

[1] R. L. Geiger, E. Sanchez-Sinencio, "Active filter design using operational transconductance amplifiers: A tutorial," *IEEE circuits and devices magazine*, vol. 1, no. 2, pp. 20-32, Mar 1985.

[2] W.Chunhua, Z. Ling and L. Tao, "A new OTA-C current-mode biquad filter with single input and multiple outputs," *AEU-International Journal of Electronics and Communications*, vol. 62, no. 3, pp.232-234, 2008.

[3] W. Tangsrirat, T. Dumawipata and W. Surakampontorn,"Multiple-input single-output current-mode multifunction filter using current differencing transconductance amplifiers," AEU- *International Journal of Electronics and Communications*, vol. 61, no. 4, pp.209-214, 2007.

[4] A. K. Singh and R. Senani, "Active-R design using CFOA-poles: new resonators, filters, and oscillators," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 48, no. 5, pp.504- 511, 2001.

[5] C. Toumazou and F.J. Ledgey, "Universal active filter using current conveyors," *Electronics Letters*, vol. 22, no.12, pp.662-664, 1986.

[6] S. A. Bashir and N. A. Shah, "Voltage mode universal filter using current differencing buffered amplifier as an active device," *Circuits and Systems*, vol. 3, no. 03, 2012.

[7] S. Kılınç and U. Çam, "Cascadable allpass and notch filters employing single operational transresistance amplifier," *Computers & Electrical Engineering*, vol. 31, no. 6, pp.391-401, 2005.

[8] C. Cakir, U. Cam and O. Cicekoglu, "Novel allpass filter configuration employing single OTRA," *IEEETransactions on circuits and systems II: Express Briefs*, vol. 52, no. 3, pp.122-125, 2005.

[9] U. Cam, C. Cem and O. Cicekoglu, "Novel transimpedance type first-order all-pass filter using single OTRA," *AEU-International Journal of Electronics and Communications*, vol. 58, no. 4, pp.296-298, 2004.

[10] F. Kaçar, "Operational transresistance amplifier based current-mode all-pass filter topologies." *In 2009 Applied Electronics* IEEE, pp. 149-152, Sep 2009.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:37:37 UTC from IEEE Xplore. Restrictions apply.

[11] A. Ranjan, M. Ghosh and S.K. Paul, "Voltage-mode third order band pass filter employing operational transresistance amplifier," *In 2012 5th International Conference on Computers and Devices for Communication (CODEC)* IEEE, pp. 1-3, Dec 2012.

[12] H. Gahlawat, H. Kumar, J. Kamnani, G. Dagar and R. Pandey, "Voltage Mode Universal First Order Filter Employing Single OTRA," *i-Manager's Journal on Electronics Engineering*, vol. 5, no. 4, pp.24, 2015.

[13] S. Kılınç, A.U. Keskin. and U. Çam, "Cascadable voltage-mode multifunction biquad employing single OTRA," *Frequenz*, vol. 61, no. 3-4, pp.84-86, 2007.

[14] V. Bhatt, P. Benjwal and M. Joshi, "OTRA Based Second Order Universal Filter and its optimization like Butterworth, Chebyshev and Bessel," *International Conference on Research in Intelligent and Computing in Engineering*, pp. 143-150, 2017.

[15] A. Gokcen and U. Cam, "MOS-C single amplifier biquads using the operational transresistance amplifier" *AEU-International Journal of Electronics and Communications*, vol.63, no. 8, pp.660-664, 2009.

[16] R. Pandey, N. Pandey, M. Singh, M. Jain and S.K. Paul, "Voltage mode biquadratic filter using single OTRA," *In 2012 IEEE 5th India International Conference on Power Electronics (IICPE)*, IEEE, pp.1-4, Dec 2012.

[17] S. Kilinç and U. Çam, "Realization of n-th Order Voltage Transfer Function Using a Single Operational Transresistance Amplifier," *Etri Journal*, vol. 27, no. 5, pp.647-650, 2005.

[18] M. Ghosh, S.K. Paul and R.K. Ranjan, "Universal filter using operational transresistance amplifier," *In 2012 IEEE 5th India International Conference on Computers and Devices for Communication(CODEC)*, IEEE, pp.1- 4, 2012.

[19] M. Ghosh, S.K. Paul, R.K. Ranjan and A. Ranjan, "Third order universal filter using single operational transresistance amplifier," *Journal of Engineering*, 2013.

[20] P.K. Tewari and R. Arya, "Design and Realization of Fractional Low-Pass Filter of 1.5 Order Using a Single Operational Transresistance Amplifier," *International Journal of Signal Processing, Image Processing and Pattern Recognition*, vol. 9, no. 9, pp.69-76, 2016.

[21] A. Gökçen, S. Kilinc, S. and U. Çam, "Fully integrated universal biquads using operational transresistance amplifiers with MOS-C realization," *Turkish Journal of Electrical Engineering & Computer Sciences*, vol. 19, no.3, pp.363-372, 2011.

[22] R. Anurag, R. Pandey, N. Pandey, M. Singh and M. Jain, "OTRA based shadow filters," *In 2015 Annual IEEE India Conference (INDICON)* , IEEE, pp.1-4, Dec 2015.

[23] N.I. Khachab and S.M. Naeim, "Fourth order filter structures using Operational Trans-resistance Amplifiers," *In ICM 2011 Proceeding*, IEEE, pp.1-4, Dec 2011.

[24] R. Pandey, N. Pandey, S.K.. Paul, A. Singh, B. Sriram and K. Trivedi, "Voltage mode OTRA MOS-C single input multi output biquadratic universal filter," *Advances in Electrical and Electronic Engineering*, vol. 10, no. 5, pp.337-344, 2012.

[25] A. Gökçen and U. Çam, "A 5th order video band elliptic filter topology using OTRA based Fleischer Tow Biquad with MOS-C Realization," *Natural and Engineering Sciences*, vol. 1, no. 2, pp.44-52, 2016.

[26] N. Pandey, R. Pandey, R. Anurag and R. Vijay, "A Class of Differentiator-Based Multifunction Biquad Filters Using OTRAs," *Advances in Electrical and Electronic Engineering*, vol. 18, no. 1, pp.31-40, 2020.

[27] Y.S. Hwang, D.S. Wu, J.J. Chen, C.C. Shih, and W.S Chou, "Design of current-mode MOSFET-C filters using OTRAs," *International Journal of Circuit Theory and Applications*, vol. 37, no. 3, pp.397-411, 2009.

[28] P. Tripathi, P. Pahalwan, P. Gola, N. Pandey and R. Pandey, "Design of digitally controlled OTRA based filter for hearing aid application," *In 2016 IEEE 1st International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES)*, IEEE, pp. 1-5, July 2016.

[29] N. Pandey, V. Kumar, A. Goel and A. Gupta, "Electronically tunable LC high pass ladder filter using OTRA," *ICTACT J. Microelectronics*, vol.3, no.3, pp.446-451, 2017.

[30] J. J. Chen, H. W. Tsao, S. I. Liu, and W. Chiu, "Parasitic-capacitance-insensitive current-mode filters using operational transresistance amplifiers," *IEE Proceedings-Circuits, Devices and Systems*, vol. 142, no. 3, pp.186-192, 1995.

[31] C. Sánchez-López, E. Martínez-Romero and E. Tlelo- Cuautle. "Symbolic analysis of OTRAs-based circuits." *Journal of applied research and technology*, vol.9, no.1, pp: 69-80, 2011

[32] H. Mostafa, and A. M. Soliman, "A modified CMOS realization of the operational transresistance amplifier (OTRA)," *Frequenz*, vol. 60, no. 3-4, pp.70-77, 2006.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:37:37 UTC from IEEE Xplore. Restrictions apply.

Software-Defined Networking: A Traffic Engineering Approach

Shanu Bhardwaj
Department of Computer Science and Engineering
Delhi Technological University
Delhi, India
shanubhardwaj1@gmail.com

Ashish Girdhar
Department of Computer Science and Engineering
Delhi Technological University
Delhi, India
ashishgirdhar@dtu.ac.in

***Abstract*—Software-Defined Networking (SDN) is the new networking approach that overcomes the obstacles that are faced by the conventional networking paradigm. The core idea of SDN is to separate the control plane from the data plane. This idea improves the network in many ways such as efficient utilization of resources, management of the network, innovation with new evolution, reduced cost, and many others. To manage all these changes, there is a great need for an efficient traffic engineering tool to improve the utilization of resources for the better performance of the network. Traffic engineering is also responsible for the analysis and monitoring of real-time data traffic. This paper mainly focuses on the structure of traffic engineering in SDN. In addition, the scope of various parameters of traffic engineering in the SDN environment and setup experimentations are also demonstrated. Hence, this work can leverage traffic engineering in the environment of SDN to enhance the network for better use in the future.**

Keywords—SDN, traffic engineering, traffic network, traffic analysis, network parameters.

I. INTRODUCTION

With the growth and development of many new applications of the Internet of Things [1], cloud computing, and many more in the network, the conventional architecture is not sufficient to meet the needs of the current environment [2]. Therefore, a new paradigm is designed by some researchers to prevail over the conventional architecture, named as Software-Defined Networking [3]. The problem in the conventional network is that both the planes of SDN are integrated into the same appliance [4]. As an outcome, the conventional architecture cannot provide the global perspective of the network and even, each device requires manual configuration. Hence, the new approach increases flexibility builds the network to configure easily and more programmable by distinctive the control/network plane from the data/physical plane with a global perspective of the centralized network [5].

Traffic Engineering is the study in which the measurement and analysis of data traffic take place to upgrade the performance of the network in an efficient manner [7]-[8]. It is the mechanism to enhance the performance of the network by providing dynamic behaviors of predicting the data traffic, analyzing, design the data routing schema, and transmitting the data [9]. To generate these dynamics behaviors, network observing plays a sprightly role. In conventional architecture, the technologies used for traffic engineering include Internet-Protocols and Multi-Protocols Switching based on Traffic Engineering.

Even though the SDN [10] furnishes hold up with traffic engineering but still there is not any research that shows the structure of SDN with traffic engineering which is of substantial significance for the future of SDN [11]. Hence, this paper provides the structure of new emerging technology SDN with traffic engineering. Also, provides the reach of traffic engineering in SDN to enhance the architecture of the network for better use.

The remainder of the paper is organized in such a manner: Section II provides the literature work done in the SDN with traffic engineering followed by past to future scenarios of traffic engineering. Section III describes the structure of SDN in traffic engineering with different parameters and measures. Section IV discusses the reach of traffic engineering in the SDN. Section V provides a conclusion.

II. PAST WORK

As shown in fig.1, the evolution of traffic engineering from the past to the future of SDN. In the early, Asynchronous Transfer Mode (ATM) [12] switching was used as a traffic engineering appliance. ATM traffic engineering can transmit different services that work simultaneously on the network. In this transfer mode, connection-oriented communication is taking place, which means the connection can be established even before forwarding the data to the destination.

After some development and re-growth of new terminologies, there was an evolution of the IP routing [13] scheme to pass on the data packets from source host to destination host [14]. As with the growth of emerging automation such as the IoT [15], Cloud computing [16], Sensor network [17], and many more the data traffic is increasing day by day. So to overcome that limitation multi-protocol routing was used.

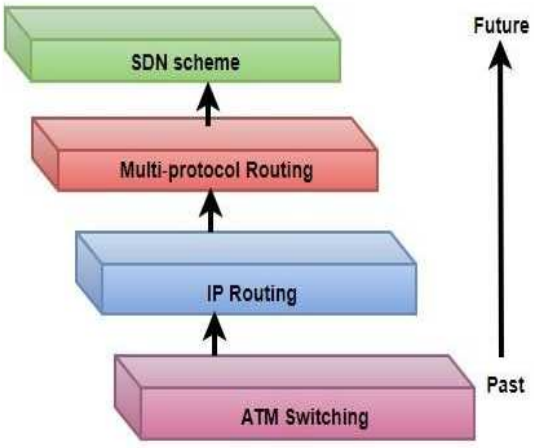


Fig.1. Past to Future Traffic Engineering in SDN

2021 IEEE 8th Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON) | 978-1-6654-0962-9/21/$31.00 ©2021 IEEE | DOI: 10.1109/UPCON52273.2021.9667584

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:49:54 UTC from IEEE Xplore. Restrictions apply.

III. Structure Of Traffic Engineering In SDN

As shown in Fig.2, the structure of traffic engineering in SDN includes two sections: measurement of data traffic and management of data traffic. Both the approaches of traffic engineering include the proper connection of topologies, different packet counters, end-to-end network analysis, traffic matrices, and many more. With the analysis of traffic, validation of the SDN network can be done and can also predict the statistics of packet encountered from receiver to sender.

A. SDN Traffic Measurement

Traffic measurement mostly deals with the status of the network in the environment of SDN. It also measures the performance of data packets in the form of monitoring [18], measuring [19], and gathering the status of the network of the SDN environment. The traffic measurement is categorized into three directions: Network parameters, Collective parameters, and Traffic analysis.

1) Network Parameter: The network parameter shows the current scenario of the entire network. Therefore, it is necessary to know the status of every node in the network, to pass on the data from the source host to the target host in the SDN environment [20]. The network parameters mainly consist of the topology parameter of the network, the data traffic parameter of the network, and the performance parameter.

2) Collective Parameter: To Collective parameters are the general parameters of the network traffic which consists of two modes of measurement: Active mode and Passive mode. Both modes provide the measurement mechanism for the monitoring of data traffic [21]. The active mode actively produces the files whereas the passive mode just monitors the transit of data from source host to destination host. For security purposes, detecting the data traffic in the passive mode is much more difficult than that of the active mode [22].

3) Traffic Analysis: Analysis is the most crucial part of any aspect; similarly, traffic analysis also plays a main role in the network measurement and provides different parameters such as congestion in the network, delay of the data from source host to destination host, and packet loss during transit. Better prediction and analysis of the data traffic result in good scheduling of the traffic network [23].

B. SDN Traffic Management

Figures Traffic management mostly deals with the management of the network via measures provided by the traffic measurement in the SDN environment [24]. It manages the network traffic with the information provided in the network. Traffic management is categorized into further three categories: Load balancing management, Energy-saving management, and Hybrid SDN management.

1) Load Balancing Management: Load balancing in SDN means the calculation decision is not distributed in each device rather it is centralized and has multiple links to improve the utilization of the path for the better strategy of the load balancing [25][26].

2) Energy Saving Management: For environmental protection, there is a great need for the reduction in the consumption of energy [27]. Therefore, SDN is not only considering the resource utilization, management of the network but also for the energy-saving mechanism [28]. The energy-saving mechanism mainly depends on various network parameters such as switches, links, and many others.

3) Hybrid SDN Management: Hybrid SDN management comes into existence when some of the limitations occur in the SDN network [29]. So, to overcome these obstacles there is a need for hybrid SDN management. In hybrid SDN, a global aspect of the whole network can be obtained accurately. But still, there are a lot of issues to manage the data traffic in the hybrid network of the SDN environment [30].

IV. Scope Of Traffic Engineering In SDN

The existence of SDN has applied to various environments of the network such as large-scale networks, cellular networks, and many more [31]. Traffic engineering is the technology that is very important for the SDN evolution. Fig. shows the scope of traffic engineering in SDN which includes four major thrusts that are flow management, updating topologies, traffic characterization, and fault tolerance [32] as demonstrated in fig.3.

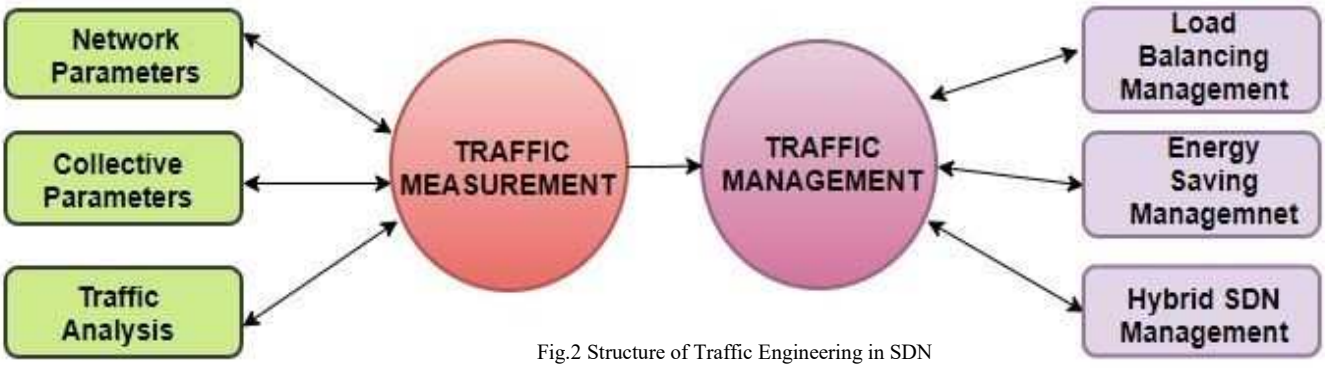


Fig.2 Structure of Traffic Engineering in SDN

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:49:54 UTC from IEEE Xplore. Restrictions apply.

A. *Flow Management*

When the flow of data traffic does not match with any of the rules present inflow table. Then, either the flow of the packet is sent by the switch to the centralized controller of the SDN, or the flow path is computed in the SDN controller. Therefore, to resolve the issues, the flow management of traffic engineering is designed to address the relationship between load balancing and latency [33].

B. *Topology Update*

The topology update consists of the new network policies that are updated in the flow table whenever there is a change in the network [34]. SDN focuses on changes in the policy, not on the link failure or the change of the element in the network. This policy manages flow control at each level of the network.

C. *Fault Tolerance*

Fault tolerance works on the reliability of the network. By this means, SDN can perform well in the failure of the network. Failure in fault tolerance is in the infrastructure of the network such as controller, paths, and switches.

D. *Traffic Characterization*

Traffic characterization is the analysis of the data traffic which includes traffic monitoring tools, collection of data traffic, data mining, pattern classification, flow data checking, programming errors, software for debugging, and many more [35]. Traffic monitoring tool plays a significant role in the traffic characterization for detecting the failures in the network. Therefore, advanced and efficient traffic monitoring tools must be developed to achieve accuracy in the measurement of the traffic, less complexity in the network, reachability of data from source host to destination host, and less overhead.

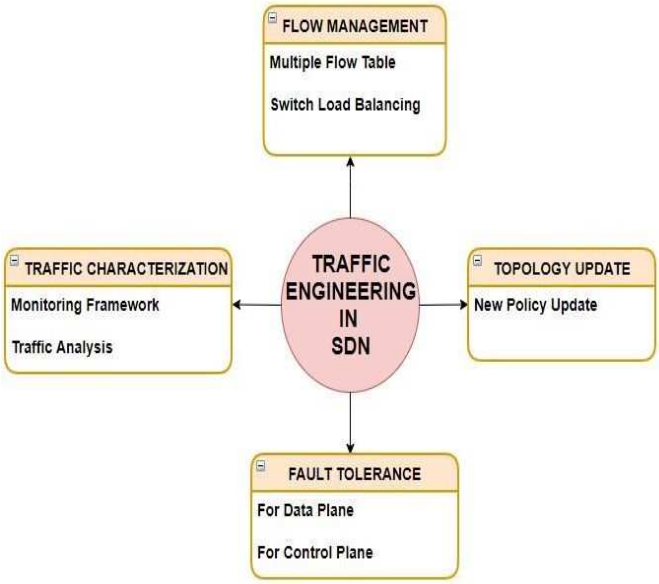


Fig.3 Scope of Traffic Engineering in SDN

V. Discussion

For the traffic analysis in Software-Defined Networking, the creation of topology in the Mininet emulator is the first step. Fig. 4 is the screenshot of the creation of topology in the emulator, having four nodes, one switch, and one default controller.

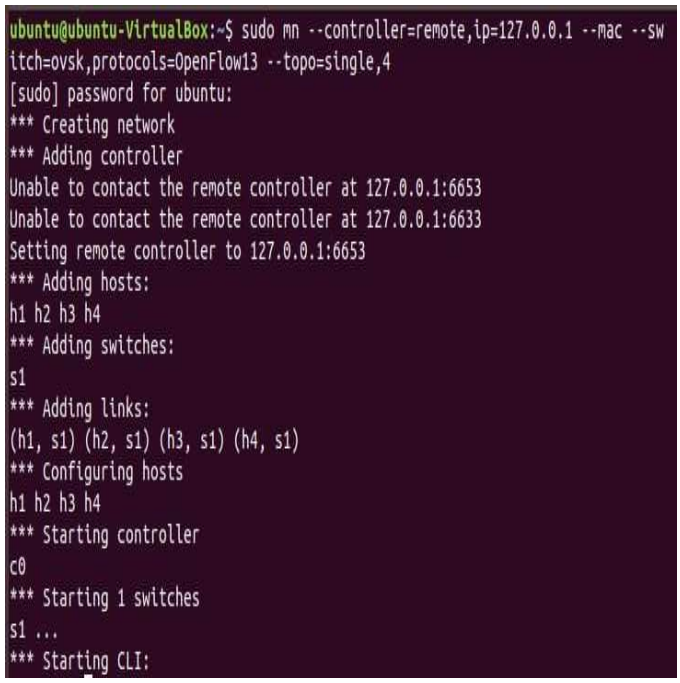


Fig. 4 Topology creation in SDN

After the creation of the topology, the traffic is the main entity to showcase the analysis of the performance metrics in the network. With the help of controller C0, the ping command has been executed to check the connection is established or not. Fig. 5 is the screenshot of the hping tool in the VM test machine. The resultant shows 0% packet loss with acceptance of 5 packets.

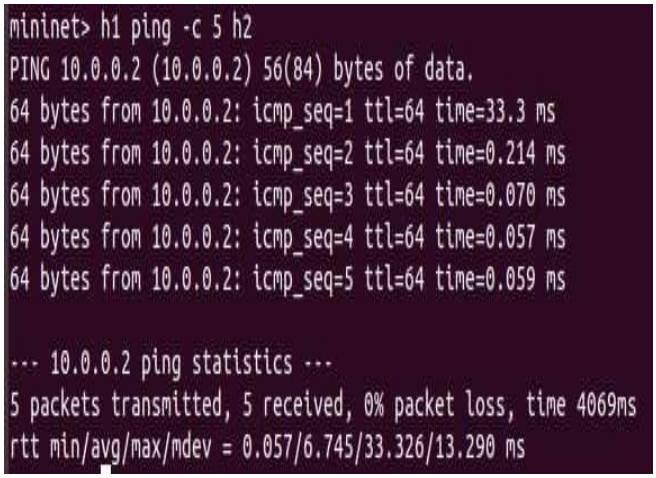


Fig. 5 Hping tool on Virtual Machine.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:49:54 UTC from IEEE Xplore. Restrictions apply.

As per the results, the work on traffic engineering in the field of Software-Defined Networking is increasing day by day. With the help of a dimension research analysis tool, the paper concluded that the research topic traffic engineering using SDN has a great demand as the number of publications and amount of work done has increased over the past 10 years.

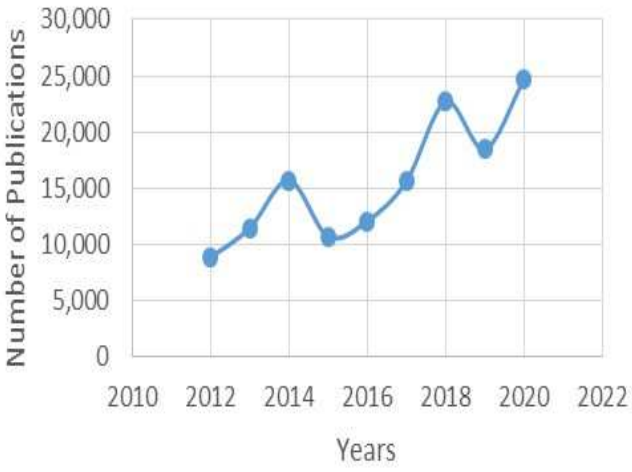


Fig. 6 Work done in Traffic Engineering in the field of SDN

CONCLUSION

With the growth of technologies, traffic engineering is one of the crucial aspects of the application of SDN. Traffic engineering is also responsible for the analysis and monitoring of real-time data traffic. In any networking approach, monitoring and analyzing real-time data traffic is a significant mechanism to observe the transmission of data packets from the source host to a destination host. Therefore, this paper demonstrated the structure of SDN's traffic engineering. Also, it presented the scope of different parameters of traffic engineering in the environment of SDN. Hence, this work holds the traffic engineering in the environment of SDN to improve the network in various ways such as to manage all the alterations done in the network, improve the utilization of resources for the better performance of the network, management data traffic in the network, reduction of the cost in the SDN network, for the innovations in the existing solutions, and many more.

REFERENCES

[1] Wan, J., Zou, C., Zhou, K., Lu, R., & Li, D. (2014). IoT sensing the framework with inter-cloud computing capability in vehicular networking. Electronic Commerce Research, 14(3), 389-416.

[2] Wan, J., Zhang, D., Sun, Y., Lin, K., Zou, C., & Cai, H. (2014). VCMIA: A novel architecture for integrating vehicular cyber-physical systems and mobile cloud computing. Mobile Networks and Applications, 19(2), 153160.

[3] Farhady, H., Lee, H., & Nakao, A. (2015). Software-defined networking: A survey. Computer Networks, 81, 79-95.

[4] Queiroz, W., Capretz, M. A., & Dantas, M. (2019). An approach for SDN traffic monitoring based on big data techniques. Journal of Network and Computer Applications, 131, 28-39.

[5] Akyildiz, I. F., Lee, A., Wang, P., Luo, M., & Chou, W. (2014). A roadmap for traffic engineering in SDN-OpenFlow networks. Computer Networks, 71, 1-30.

[6] Agarwal, S., Kodialam, M., & Lakshman, T. V. (2013, April). Traffic engineering in software-defined networks. In 2013 Proceedings IEEE INFOCOM (pp. 2211-2219). IEEE.

[7] Bhardwaj, S., Panda, S. N., & Datta, P. (2020, March). Comparison and Performance Evaluation of Software-Defined Networking Controllers. In 2020 International Conference on Emerging Smart Computing and Informatics (ESCI) (pp. 276281). IEEE.

[8] Suárez-Varela, J., & Barlet-Ros, P. (2018). Flow monitoring in Software-Defined Networks: Finding the accuracy/performance tradeoffs. Computer Networks, 135, 289-301.

[9] Awduche, D., Chiu, A., Elwalid, A., Widjaja, I., & Xiao, X. (2002). Overview and principles of Internet traffic engineering (pp. 2070-1721). RFC 3272, may.

[10] Badotra, S., & Panda, S. N. (2019). A review on SoftwareDefined Networking enabled IoT cloud computing. IIUM Engineering Journal, 20(2), 105-126.

[11] Bhardwaj, S., & Panda, S. N. (2019). SDWSN: SoftwareDefined Wireless Sensor Networking. Internal Journal of Innovative Technology and Exploring Engineering, 8(12), 1064-1071.

[12] J.J. Bae, T. Suda, Survey of traffic control schemes and protocols in ATM networks, Proc. IEEE 79 (2) (1991) 170189.

[13] N. Wang, K. Ho, G. Pavlou, M. Howarth, An overview of routing optimization for internet traffic engineering, Commun, Surv. Tut. IEEE 10(1) (2008) 36-56 (First Quarter).

[14] G. lannaccone, C,-N. Chuah, R. Mortier, S. Bhattacharyya, C. Diot, Analysis oflink failure in an IP backbone, in; Proceeding of the 2nd ACM SIGCOMM Workshop on Internet Measurement, November 2002, PP, 237-242.

[15] Hameed, S., Khan, F. I., & Hameed, B. (2019). Understanding security requirements and challenges in Internet of Things (IoT): A review. Journal of Computer Networks and Communications, 2019.

[16] Sharma, P. K., Park, J. H., Jeong, Y. S., & Park, J. H. (2019). Shsec: sdn based secure smart home network architecture for internet of things. Mobile Networks and Applications, 24(3), 913-924.

[17] Bhardwaj, S., & Panda, S. N. (2020). A Study on Noninvasive Body Wearable Sensors. In Intelligent Communication, Control and Devices (pp. 345-351). Springer, Singapore.

[18] Van Adrichem, N. L., Doerr, C., & Kuipers, F. A. (2014, May). Opennetmon: Network monitoring in openflow software-defined networks. In 2014 IEEE Network Operations and Management Symposium (NOMS) (pp. 1-8). IEEE.

[19] Yuan, L., Chuah, C. N., & Mohapatra, P. (2010). ProgME: towards programmable network measurement. IEEE/ACM Transactions on Networking, 19(1), 115-128.

[20] Z. Hu and J. Luo, ''Cracking network monitoring in DCNs with SDN,'' in Proc. IEEE Comput. Commun., Apr./May 2015, pp. 199–207

[21] M. Yu, L. Jose, and R. Miao, ''Software defined traffic measurement with OpenSketch,'' in Proc. 10th USENIX Symp. Netw. Syst. Design Implement. (NSDI), 2013, pp. 29–42

[22] M. Moshref, M. Yu, R. Govindan, and A. Vahdat, ''DREAM: Dynamic resource allocation for software-defined measurement,'' ACM SIGCOMM Comput. Commun. Rev., vol. 44, no. 4, pp. 419–430, 2015.

[23] S. A. Mehdi, J. Khalid, and S. A. Khayam, ''Revisiting traffic anomaly detection using software defined networking,'' in Proc. Int. Symp. Recent Adv. Intrusion Detection, Menlo Park, CA, USA, Sep. 2011, pp. 161–180.

[24] M. Al-Fares, S. Radhakrishnan, B. Raghavan, N. Huang, and A. Vahdat, ''Hedera: Dynamic flow scheduling for data center networks,'' in Proc. Netw. Syst. Design Implement. Symp. (NSDI), vol. 10. 2010, p. 19.

[25] A. R. Curtis, W. Kim, and P. Yalagandula, ''Mahout: Low-overhead datacenter traffic management using end-hostbased elephant detection,'' in Proc. IEEE INFOCOMM, Apr. 2011, pp. 1629–1637.

[26] M. Chen, H. Jin, Y. Wen, and V. C. M. Leung, ''Enabling technologies for future data center networking: A primer,'' IEEE Netw., vol. 27, no. 4, pp. 8–15, Jul. 2013.

[27] Bhardwaj, S., Panda, S. N., & Datta, P. (2020, December). Layer-Based Attacks in the Ternary Planes of SoftwareDefined Networking. In 2020 IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-ECE) (pp. 292-295). IEEE.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:49:54 UTC from IEEE Xplore. Restrictions apply.

[28] S. Agarwal, M. Kodialam, and T. V. Lakshman, ‘‘Traffic engineering in software defined networks,’’ in Proc. IEEE INFOCOM, Apr. 2013, pp. 2211–2219.

[29] Y. Guo, Z. Wang, X. Yin, X. Shi, and J. Wu, ‘‘Traffic engineering in SDN/OSPF hybrid network,’’ in Proc. IEEE 22nd Int. Conf. Netw. Protocols (ICNP), Oct. 2014, pp. 563– 568.

[30] Chowdhury, S. R., Bari, M. F., Ahmed, R., & Boutaba, R. (2014, May). Payless: A low cost network monitoring framework for software defined networks. In 2014 IEEE Network Operations and Management Symposium (NOMS) (pp. 1-9). IEEE.

[31] Tahaei, H., Salleh, R. B., Ab Razak, M. F., Ko, K., & Anuar, N. B. (2018). Cost effective network flow measurement for software defined networks: A distributed controller scenario. IEEE Access, 6(2), 5182-5198.

[32] Srivastava, V., & Pandey, R. S. (2020). A Reward Based Formal Model for Distributed Software Defined Networks. Wireless Personal Communications, 23(1), 1-17.

[33] Freris, N. M. (2019). A software-defined architecture for control of IoT cyberphysical systems. Cluster Computing, 22(4), 1107-1122.

[34] Silva, J. B., Silva, F. S. D., Neto, E. P., Lemos, M., & Neto, A. (2020). Benchmarking of mainstream SDN controllers over open off-the-shelf software-switches. Internet Technology Letters, 3(3), 152-160.

[35] Meena, R. C., Bundele, M., & Nawal, M. (2020, February). RYU SDN Controller Testbed for Performance Testing of Source Address Validation Techniques. In 2020 3rd International Conference on Emerging Technologies in Computer Engineering: Machine Learning and Internet of Things (ICETCE), February 07-08, 2020, Jaipur, India, pp. 16. IEEE.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:49:54 UTC from IEEE Xplore. Restrictions apply.

2021 IEEE 6th International Conference on Computing, Communication and Automation (ICCCA)
Aurel Vlaicu University of Arad, Romania. Dec 17-19, 2021

Solution of Transportation Problem Under Spherical Fuzzy Set

1st Vineet Kumar
Department of Applied Mathematics
Delhi Technological University
Delhi, India
vineetsrivastava21@gmail.com

2nd Anjana Gupta
Department of Applied Mathematics
Delhi Technological University
Delhi, India
anjanagupta@dce.ac.in

3rd H.C. Taneja
Department of Applied Mathematics
Delhi Technological University
Delhi, India
hctaneja@dce.ac.in

2021 IEEE 6th International Conference on Computing, Communication and Automation (ICCCA) | 978-1-6654-1473-9/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICCCA52192.2021.9666372

***Abstract*—Transportation problem (TP) is used to solve real-life problems as agriculture, management and engineering science etc. But in real-life transportation problems, the cost depends on uncontrollable factors like fuel price, weather conditions and consumers behaviour etc. To deal with this type of imprecision, fuzzy numbers and their extension have been introduced by the authors. In literature, many methods are available for finding the solution of TP under fuzzy set (FS), intuitionistic fuzzy set (IFS) and Pythagorean fuzzy set (Pyfs) etc. In the proposed work, we have investigated the solution of the spherical fuzzy transportation problem (SFTP) and presented three different models of the spherical fuzzy transportation problem. Spherical fuzzy set (SFS) arithmetic operations and algorithms are discussed here. To illustrate the proposed work we have taken the numerical problem for these sets. Finally, we have discussed the result and conclusion of the paper.**

***Index Terms*—Transportation problem, Fuzzy set, Spherical fuzzy set, Score function.**

I. Introduction

TP is a linear programming problem and originally formulated by Hitchcock [18]. In a traditional transportation dilemma, the decision-maker knows exactly how much transportation costs, how much supply there is, and how much demand there is, but in real-life transportation problems all these values depend on various factors like fuel costs, condition of whether and customer loyalty, etc. which are not fixed (imprecise).
Zadeh(1965) [4] defined the concept of the fuzzy set to deal with imprecision of real-world problems. Chanas et al. [1] worked on fuzzy TP. Dinagar and Palanivel [3] solved the TP under a fuzzy environment. Kumar and Amarpreet [19] have presented the idea of using a traditional transportation strategy to solve a fuzzy transportation challenges. Many authors have worked on fuzzy transportation, but they only address the element's membership function..

The intuitionistic fuzzy set was introduced by Atanassov [2]. The intuitionistic fuzzy set is defined by membership and non-membership function of the element to the set i.e (δ, λ) with the condition $\delta + \lambda \leq 1$. The IFTP was solved by Gani and Abbs [5] using a new ranking algorithm in which supply and demand are triangular intuitionistic fuzzy numbers. Singh and Yadav [8] introduced the transportation problem in which expenses are expressed as triangular intuitionistic fuzzy numbers. Dubey and Mehra [7] worked on linear programming with the triangular intuitionistic fuzzy numbers. Bharti [14] solved the trapezoidal intuitionistic fuzzy fractional transportation problem. Atanssov and Gargov [6] defined the interval-valued intuitionistic fuzzy set (IVIFS). Mishra et al. [10] worked on transportation problem under IVIF. Arora [9] given an approach for interval-valued fuzzy fractional TP. Bharati and Singh [11] solved the TP under IVIFS.

Yager [16] defined the pythagorean fuzzy set (Pyfs). Yager overcome the situation when $\delta + \lambda > 1$. A pythagorean fuzzy set is an extention of IFS. Zhang and Xu [12] defined the methodology of the TOPSIS method to MCDM with Pyfs. Geng et al. [13] presented an exteended TODIM method and proposed pythagorean fuzzy uncertain linguistic set. Karasan et al. [15] worked on landfill site selection problem using novel pythagorean fuzzy AHP method. Kumar et al. [20] presented the pythagorean fuzzy approach to the transportation problem. Zhang [28] proposed the interval-valued pythagorean fuzzy set (IVPFS). He extended membership and non-membership in interval numbers. Peng and Yang [21] defined the fundamental properties of IVPFS. Peng and Li [22] proposed an algorithm for IVPFS in emergency decisions depend on two measures multiparametric measure and WDBA. Rahaman et al. [23] gave the methodology to multi-attribute group decision-making based on induced IVPFS Einstein aggregation operator.

The condition where neutral membership degree calculated independently in real-life Problem, based on these two sets FS and IFS, the picture fuzzy set (PFS) introduced by Cuong and Kreinovich [29]. They used three index, membership degree (δ), neutral-membership degree $(\eta)$and non-membership degree (λ) in picture fuzzy set with the condition $0 \leq \delta+\eta+\lambda \leq 1$. Dutta and Ganju [24] proposed some characteristics of of PFS. Van and Thao [25] worked on a measure of PFS and their use in multi-attribute decision-making problem. Geetha and Selvakumari [30] solved the picture Fuzzy transportation problem using the score function.

The generalization of PFS is SFS which is introduced by Ashraf et al. [26] with condition $0 \leq \delta + \eta + \lambda \leq 1$.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:47:37 UTC from IEEE Xplore. Restrictions apply.

Consider the example, let δ=0.7, η=0.3 and λ=0.5, in this case 0.7+0.3+0.5 >1 but by squaring $(0.7)^2$ +$(0.3)^2$+$(0.5)^2 \leq 1$. By this, we can see SFS is more suitable than PFS. Spherical fuzzy sets and the spherical fuzzy TOPSIS method were studied by Gundogdu and Kahraman [27]. The spherical fuzzy set was utilised by Mahmood et al. [17] to solve decision-making and medical diagnosis challenges.

The findings of the preceding literature analysis on transportation issues, we found that no one worked on transportation problems under the spherical fuzzy sets. In our work, we have proposed the method to solve spherical fuzzy transportation problem using three different cases. In these cases we have used the score function to convert the data into a crisp transportation problems. After that, we have applied the vogel approximation method (VAM) for an initial basic feasible solution (IBFS) and verified the optimality with computational technique. This is the new method to handle the uncertainty in TP.

We have divided the whole paper into 4 Sections. In Section 2 some basic knowledge of FS, Pyfs, PFS, SFS, and arithmetic operations on spherical fuzzy numbers are presented. Methodology to solve spherical fuzzy transportation problem and their illustrative examples are given in Section 3. Section 4, presents the result and concluding remarks.

II. Preliminaries

Definition 2.1 [4] Let G be a universal set, then a fuzzy set (FS) H is defined as

$$H = \{\langle g, \theta_H(g) \,|g \in G\rangle\}$$

where

$$\mu_H(g) : G \to [0,1]$$

and $\theta_H(g)$ is the membership degree of the element g to the set H.

Definition 2.2 [16] Pythagorean fuzzy set '$\tilde{H}$' on G is defined as a set of ordered pair given by

$$\tilde{H} = \{< g, \theta_H(g), \lambda_H(g) > |g \in G\}$$

where $\theta_H(g), \lambda_H(g) : G \to [0,1]$ are the membership and non-membership degree of the element $g \in G$ respectively, with the condition $(\theta_H(g))^2 + (\lambda_H(g))^2 \leq 1$, the degree of indeterminacy is given by $\alpha_H(g) = \left(\sqrt{1-(\theta_H^2+\lambda_A^2)}\right)$

Definition 2.3 [25] Picture fuzzy set $\tilde{H}$ on a universe G is an object in the form of

$$\tilde{H} = \{[g, \theta_{\tilde{H}}(g), \eta_{\tilde{H}}(g), \lambda_{\tilde{H}}(g)]|g \in G\}$$

where $\theta_{\tilde{H}}(g)\in$[0, 1] is the degree of membership of g in $\tilde{H}$, $\eta_{\tilde{H}}(g)\in$[0,1] is the neutral membership of g in $\tilde{H}$ and $\lambda_{\tilde{H}}(g)\in$[0, 1] is the non-membership of g in $\tilde{H}$, where $\lambda_{\tilde{H}}(g), \eta_{\tilde{H}}(g), \lambda_{\tilde{H}}(g)$ satisfy the following condition

$$(g \in G), \theta_{\tilde{H}}(g) + \eta_{\tilde{H}}(g) + \lambda_{\tilde{H}}(g)$$

now (1-($\theta_{\tilde{H}}(g) + \eta_{\tilde{H}}(g) + \lambda_{\tilde{H}}(g)$)) called the refusal membership degree of g in $\tilde{H}$.

Definition 2.4 [27] A spherical fuzzy set (SFS) $\tilde{\text{H}}$ defined on universal set G is given by

$$\tilde{H}_s = \{< g, \theta_{H_s}(g), \eta_{H_S}(g), \lambda_{H_S}(g) >, g \in G\}$$

where

$$\theta_{H_s}(g) : G \to [0,1], \eta_{H_S}(g) : G \to [0,1], \lambda_{H_S}(g) : G \to [0,1]$$

and

$$0 \leq \theta_{H_s}(g)^2 + \eta_{H_S}(g)^2 + \lambda_{H_S}(g)^2 \leq 1$$

Definition 2.5 [27] Let $\tilde{\text{A}}_\text{s}$ & $\tilde{\text{B}}_\text{s}$ are two SFS. Then

1) Union:

$$\tilde{A}_s \cup \tilde{B}_s = \{\max\{\theta_{A_s}, \theta_{B_s}\}, \min\{\lambda_{A_s}, \lambda_{B_s}\}, \min\{(1 - ((\max\{\theta_{\tilde{A}_S}, \theta_{\tilde{B}_S}\})^2 + (\min\{\lambda_{\tilde{A}_S}, \lambda_{\tilde{B}_S}\})^2))^{\frac{1}{2}}, \max\{\eta_{\tilde{A}_S}, \eta_{\tilde{B}_S}\}\}\}$$

2) Intersection:

$$\tilde{A}_S \cap \tilde{B}_S = \{\min\{\theta_{A_s}, \theta_{B_s}\}, \max\{\lambda_{A_s}, \lambda_{B_s}\}, \max\{(1 - ((\min\{\theta_{\tilde{A}_S}, \theta_{\tilde{B}_S}\})^2 + (\max\{\lambda_{\tilde{A}_S}, \lambda_{\tilde{B}_S}\})^2))^{\frac{1}{2}}, \min\{\eta_{\tilde{A}_S}, \eta_{\tilde{B}_S}\}\}\}$$

3) Addition:

$$\tilde{A}_S \oplus \tilde{B}_S = \Big((\theta_{\tilde{A}_S}^2 + \theta_{\tilde{B}_S}^2 - \theta_{\tilde{A}_S}^2.\theta_{\tilde{B}_S}^2)^{\frac{1}{2}}, \lambda_{\tilde{A}_S}.\lambda_{\tilde{B}_S}, ((1-\theta_{B_S}^2)\eta_{A_S}^2 + (1-\theta_{A_S}^2)\eta_{B_S}^2 - \eta_{A_S}^2\eta_{B_S}^2)^{\frac{1}{2}}\Big)$$

4) Multiplication:

$$\tilde{A}_S \otimes \tilde{B}_S = \{\theta_{A_S}.\theta_{B_S}, (\lambda_{\tilde{A}_S}^2 + \lambda_{\tilde{B}_S}^2 - \lambda_{\tilde{A}_S}^2.\lambda_{\tilde{B}_S}^2)^{\frac{1}{2}}, ((1-\lambda_{B_S}^2)\eta_{A_S}^2 + (1-\lambda_{A_S}^2)\eta_{B_S}^2 - \eta_{A_S}^2\eta_{B_S}^2)^{\frac{1}{2}}\}$$

Definition 2.6 [27] Let $\tilde{\text{H}}_\text{S}$ be the SFS, The score function and accuracy function are defined by:

Score $(\tilde{H}_S) = (\theta_{\tilde{H}_S} - \eta_{\tilde{H}_S})^2 - (\lambda_{\tilde{H}_S} - \eta_{\tilde{H}_S})^2$, $S(\tilde{H}_S) \in [-1,1]$
Accuracy $(\tilde{H}_S)$=${\theta_{H_S}}^2 + {\lambda_{H_S}}^2 + {\eta_{H_S}}^2$, $Acc.(\tilde{H}_S) \in [0,1]$
Note that:$\tilde{H}_S < \tilde{Q}_S$ if and only if

1) S($\tilde{H}_S$)<S($\tilde{Q}_S$) or
2) S ($\tilde{H}_S$)=S ($\tilde{Q}_S$) and Acc.($\tilde{H}_S$)< Acc.($\tilde{Q}_S$)

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:47:37 UTC from IEEE Xplore. Restrictions apply.

III. Spherical Fuzzy Transportation Problem

A. Spherical Fuzzy Transportation Problem (SFTP)

The SFTP is

$$Min\tilde{Z} = \sum_{i}^{m}\sum_{j}^{n}\tilde{c}_{ij}^{SFS}x_{ij}$$

Such that

$$\sum_{j}^{n} x_{ij} = \tilde{a}_{i}^{SFS}, i = 1, 2, ...m$$

$$\sum_{i}^{m} x_{ij} = \tilde{b}_{j}^{SFS}, j = 1, 2, ...n$$

$$x_{ij} \geq 0$$

B. Proposed Methodology

Three distinct models of transportation problems were used to address the transportation problem under the spherical fuzzy set.

- Model 1 (In this paradigm, costs are expressed as spherical fuzzy numbers (SFNs), whereas supply and demand are expressed as crisp numbers.)

Destination	$\mathbf{D}_1$	$\mathbf{D}_2$	**…**	$\mathbf{D}_n$	**Supply**
Sources					
$\mathbf{S}_1$	${C_{11}}^{SFN}$	${C_{12}}^{SFN}$	…	${C_{1n}}^{SFN}$	a_1
$\mathbf{S}_2$	${C_{21}}^{SFN}$	${C_{22}}^{SFN}$	…	${C_{2n}}^{SFN}$	a_2
$\mathbf{S}_m$	${C_{m1}}^{SFN}$	${C_{m2}}^{SFN}$	…	${C_{mn}}^{SFN}$	a_m
$Demand$	b_1	b_2	…	b_n	

- Model 2 In this paradigm, supply and demand are expressed as spherical fuzzy numbers (SFNs), whereas costs are expressed as crisp numbers.)

Destination	$\mathbf{D}_1$	$\mathbf{D}_2$	**…**	$\mathbf{D}_n$	**Supply**
Sources					
$\mathbf{S}_1$	C_{11}	C_{12}	…	C_{1n}	${a_1}^{SFN}$
$\mathbf{S}_2$	C_{21}	C_{22}	…	C_{2n}	${a_2}^{SFN}$
$\mathbf{S}_m$	C_{m1}	C_{m2}	…	C_{mn}	${a_m}^{SFN}$
$Demand$	${b_1}^{SFN}$	${b_2}^{SFN}$	…	${b_n}^{SFN}$	

- Model 3 (The cost, supply, and demand are all represented by a spherical fuzzy number.)

Destination	$\mathbf{D}_1$	$\mathbf{D}_2$	**……**	$\mathbf{D}_n$	**Supply**
Sources					
$\mathbf{S}_1$	${C_{11}}^{SFN}$	${C_{12}}^{SFN}$	……..	${C_{1n}}^{SFN}$	${a_1}^{SFN}$
$\mathbf{S}_2$	${C_{21}}^{SFN}$	${C_{22}}^{SFN}$	……	${C_{2n}}^{SFN}$	${a_2}^{SFN}$
$\mathbf{S}_m$	${C_{m1}}^{SFN}$	${C_{m2}}^{SFN}$	……..	${C_{mn}}^{SFN}$	${a_m}^{SFN}$
Demand	${b_1}^{SFN}$	${b_2}^{SFN}$	……..	${b_n}^{SFN}$	

Algorithm

Step1. Write SFTP in tabular form.
Step 2. Use score function to change the SFTP into a crisp transportation problem.
Step 3. Check it is a balance or not.
Step 4. If it's balanced, move to step 5, otherwise make it balanced to add dummy variables.
Step 5. Find the IBFS by VAM.
Step 6. Using any Software MQ, LINGO, and MATLAB, test the optimality of the transportation problem.
Step7. To get the objective function, put x_{ij} in the objective function.

C. Illustrative Example

- (Model 1 SFTP): In this SFTP model 1, we have taken the condition for the proposed transportation problem is shown in Table 1, the cost are in spherical fuzzy numbers, supply and demand are crisp.

TABLE I
DATA FOR SPHERICAL FUZZY TP OF MODEL 1

	$\mathbf{D}_1$	$\mathbf{D}_2$	$\mathbf{D}_3$	$\mathbf{D}_4$	**Supply**
$\mathbf{O}_1$	(0.9,0.1,0.1)	(0.6,0.4,0.4)	(0.91,.03,.02)	(0.91,.03,.02)	26
$\mathbf{O}_2$	(0.89,0.08,0.03)	(0.74,0.16,0.1)	(0.5,0.5,0.5)	(0.7,0.3,0.3)	24
$\mathbf{O}_3$	(0.99,0.05,0.02)	(0.73,0.15,0.08)	(0.73,0.12,0.08)	(0.68,0.26,0.06)	30
Demand	17	23	28	12	80

We obtain the crisp transportation problem by using score function (Definiation 2.6) .

TABLE II
DEFUZZIFIED SPHERICAL FUZZY TRANSPORTATION PROBLEM OF MODEL 1

	$\mathbf{D}_1$	$\mathbf{D}_2$	$\mathbf{D}_3$	$\mathbf{D}_4$	**Supply**
$\mathbf{O}_1$	0.64	0.04	0.792	0.94	26
$\mathbf{O}_2$	0.73	0.40	0	0.16	24
$\mathbf{O}_3$	0.94	0.41	0.42	0.34	30
Demand	17	23	28	12	80

The IBFS is as follows
Minimum cost = $0.04\times23+0\times24+0.34\times12+0.42\times4+0.64\times3+0.94\times14$ =21.76
It has degeneracy solution because the m + n−1= 6. so, we have to proceed for the optimality, we use LINGO software.

Minimum cost
=$3\times0.64+0.04\times23+0\times24+0.94\times14+0.42\times4+0.34\times12$
=21.76

- (Model 2): Condition for the proposed transportation problem is shown in Table III where the costs are crisp, but the demand and supply are SFN.

We obtain the crisp transportation problem by using score function (Definiation 2.6).

The IBFS is as follows
Minimum cost

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:47:37 UTC from IEEE Xplore. Restrictions apply.

TABLE III
DATA FOR SPHERICAL FUZZY TP OF MODEL 2

	D_1	D_2	D_3	D_4	Supply
O_1	0.64	0.04	0.792	0.94	(0.9,0.1,0.1)
O_2	0.73	0.4	0	0.16	(0.89,0.08,0.03)
O_3	0.94	0.41	0.42	0.34	(0.99,0.05,0.02)
Demand	(0.9,0.1,0.1)	(0.6,0.4,0.4)	(0.91,0.03,0.02)	(0.99,0.05,0.02)	

TABLE IV
DEFUZZIFIED SPHERICAL FUZZY TP OF MODEL 2

	D_1	D_2	D_3	D_4	Supply
O_1	0.64	0.04	0.792	0.94	0.64
O_2	0.73	0.4	0	0.16	0.73
O_3	0.94	0.41	0.42	0.34	0.94
Demand	0.64	0.04	0.792	0.94	

$=0.64\times0.64+0\times0.73+0\times0.062+0\times0.04+0.94\times0.34=0.7292$

It has degeneracy solution because the m + n− 1= 6, so, we have to proceed for the optimality, we use LINGO software

The Minimum cost
$=0.64\times0.6+0.04\times0.04+0\times0.73+0\times0.42+0.34\times0.94+0\times0.04+0\times0.062=0.7052$

- (Model 3 SFN): Condition for the proposed transportation problem is shown in Table V. Here all the three paremeters costs, demands and supplies are Spherical fuzzy number.

TABLE V
DATA FOR SPHERICAL FUZZY TP OF MODEL 3

	D_1	D_2	D_3	D_4	Supply
O_1	(0.61,0.46,0.34)	(0.74,0.27,0.28)	(0.7,0.3,0.3)	(0.62,0.39,0.39)	(0.9,0.1,0.1)
O_2	(0.81,0.2,0.23)	(0.55,0.47,0.43)	(0.5,0.5,0.5)	(0.7,0.3,0.3)	(0.89,0.08,0.03)
O_3	(0.99,.05,.02)	(0.73,0.15,0.08)	(0.73,0.12,0.08)	(0.68,0.26,0.06)	(0.99,0.05,0.02)
Demand	(0.9,0.1,0.1)	(0.6,0.4,0.4)	(0.91,0.03,0.02)	(0.99,0.05,0.02)	

We obtain the crisp transportation problem by using score function (Definiation 2.6).

TABLE VI
DEFUZZIFIED SPHERICAL FUZZY TP OF MODEL 3

	D_1	D_2	D_3	D_4	Supply
O_1	0.0585	0.2115	0.16	0.3021	0.64
O_2	0.3355	.0128	0	0.34	0.73
O_3	0.73	0.4	0.41	0.42	0.94
Demand	0.64	0.04	0.792	0.94	

The IBFS is obtained as follows:
Min cost
$=0\times0.102+0.0585\times0.64+0\times0.73+0.42\times0.838+0.41\times0.062+0.4\times0.04=0.43082$

To obtain the optimality, we use LINGO software. Therefore, the optimal solution is as follows:
Minimum cost = 0.4308

IV. RESULT AND CONCLUSION

We have solved TP under spherical fuzzy set using three different examples. The result for these examples are shown in Table VII and compared our work with some previous work as shown in Table VIII.

TABLE VII
COMPARISION OF INTIAL BASIC FEASIBLE SOLUTION WITH OPTIMAL SOLUTION

SFTP	IBFS	Optimum Solution	Comparison
Model 1	21.76	21.76	IBFS$\geq$ optimum solution
Model 2	0.7292	0.7052	IBFS$\geq$ optimum solution
Model3	0.43082	0.4308	IBFS$\geq$ optimum solution

In model 1 SFTP, we found the IBFS is 21.76 and optimum transportation cost of 21.76. Also in model 2 SFTP, we found the IBFS is 0.7292 and optimum transportation cost of 0.7052. In mode 3 SFTP, we found the IBFS is 0.43082 and optimum transportation cost of 0.4308. So, it is clear from the above observation the optimal solution in all three models is less than or equal to IBFS.

TABLE VIII
ANALYSIS OF PROPOSED WORK WITH OTHERS AUTHORS WORK

S.No.	Author name	Work
1	Bharti(2018) [14]	TP under interval valued intuitionistic fuzzy environment.
2	Kumar et al.(2019) [20]	A pythagorean fuzzy approach to the TP
3	Geetha and Selvakumari (2020) [30]	A picture fuzzy approach to solving TP
4	Proposed work	Solution of TP under spherical fuzzy sets.

Finally conclude that we have developed the algorithm to solve the SFTP using three different models. In these models, we have used the score function to convert the data into crisp transportation problems. After that, we have applied the vogel approximation method for an IBFS and verified the optimality with computational technique. This is the new approach to tackle the uncertainty in real-life transportation problem.

REFERENCES

[1] Chanas, S., Kołodziejczyk, W.and Machaj, A., "A fuzzy approach to the transportation problem". Fuzzy Sets and Systems, 13(3), 211-221, (1984)

[2] Atanassov, K., "Intuitionistic fuzzy sets". fuzzy Sets and Systems 20 (1), 87-96, (1986).

[3] Dinagar, D. S. and Palanivel, K., "The transportation problem in fuzzy environment". International Journal of Algorithms, Computing and Mathematics, 2(3), 65-71, (2009).

[4] Zadeh, L. A. "Fuzzy sets"., Information and control, 8(3), 338-353, (1965).

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:47:37 UTC from IEEE Xplore. Restrictions apply.

[5] Nagoor Gani, A. and Abbas, S. “Intuitionistic fuzzy transportation problem”. In Proceedings of the Heber International Conference on Applications of Mathematics and Statistics (HICAMS) (pp. 528-535), (2012).
[6] Atanassov, K. T. and Gargov, G. “Interval valued intuitionistic fuzzy sets”. Fuzzy Sets and Systems 31, 343-349, (1989).
[7] Dubey, D. and Mehra, A., “Linear programming with triangular intuitionistic fuzzy number”. In Proceedings of the 7th Conference of the European Society for Fuzzy Logic and Technology (pp. 563-569). Atlantis Press, (2011).
[8] Singh, S. K. and Yadav, S. P., “A new approach for solving intuitionistic fuzzy transportation problem of type-2”. Annals of Operations Research, 243(1-2), 349-363, (2016).
[9] Arora, J., “An algorithm for interval-valued fuzzy fractional transportation problem”.(2018).
[10] Mishra, A., Kumar, A. and Khan, M. A., “A note on “Transportation problem under interval-valued intuitionistic fuzzy environment”. Journal of Intelligent andFuzzy Systems, 37(1), 897-900, (2019).
[11] Bharati, S. K. and Singh, S. R., “Transportation problem under interval-valued intuitionistic fuzzy environment”. International Journal of Fuzzy Systems, 20(5), 1511-1522, (2018).
[12] Zhang, X. and Xu, Z., “Extension of TOPSIS to multiple criteria decision making with Pythagorean fuzzy sets”. International Journal of Intelligent Systems, 29(12), 1061-1078, (2014).
[13] Geng, Y., Liu, P., Teng, F. and Liu, Z., “Pythagorean fuzzy uncertain linguistic TODIM method and their application to multiple criteria group decision making”. Journal of Intelligent and Fuzzy Systems, 33(6), 3383-3395, (2017).
[14] Bharati, S. K., “Trapezoidal intuitionistic fuzzy fractional transportation problem”. In Soft Computing for Problem Solving (pp. 833-842). Springer, Singapore,(2019).
[15] Karasan, A., Ilbahar, E. and Kahraman, C., “A novel pythagorean fuzzy AHP and its application to landfill site selection problem”. Soft Computing, 23(21), 10953-10968, (2019).
[16] Yager, R. R., “Pythagorean membership grades in multicriteria decision making”. IEEE Transactions on Fuzzy Systems, 22(4), 958-965, (2013).
[17] Mahmood, T., Ullah, K., Khan, Q. and Jan, N., “An approach toward decision-making and medical diagnosis problems using the concept of spherical fuzzy sets”. Neural Computing and Applications, 31(11), 7041-7053, (2019).
[18] Hitchcock, F. L., “The distribution of a product from several sources to numerous localities”. Journal of Mathematics and Physics, 20(1-4), 224-230, (1941).
[19] Kumar, A. and Amarpreet, K. A. U. R., “Application of classical transportation methods for solving fuzzy transportation problems”. Journal of Transportation Systems Engineering and Information Technology, 11(5), 68-80, (2011).
[20] Kumar, R., Edalatpanah, S. A., Jha, S. and Singh, R., “A Pythagorean fuzzy approach to the transportation problem”. Complex Andintelligent Systems, 5(2), 255-263, (2019).
[21] Peng, X. and Yang, Y., “Fundamental properties of interval-valued Pythagorean fuzzy aggregation operators”. International Journal of Intelligent Systems, 31(5), 444-487, (2016).
[22] Peng, X. and Li, W., “Algorithms for interval-valued pythagorean fuzzy sets in emergency decision making based on multiparametric similarity measures and WDBA”. Ieee Access, 7, 7419-7441, (2019).
[23] Rahman, K., Abdullah, S., Ali, A. and Amin, F., “Approaches to Multi-Attribute Group Decision Making Based on Induced Interval-Valued Pythagorean Fuzzy Einstein Hybrid Aggregation Operators”. Bulletin of the Brazilian Mathematical Society, New Series, 50(4), 845-869, (2019).
[24] Dutta, P. and Ganju, S., “Some aspects of picture fuzzy set”. Transactions of A. Razmadze Mathematical Institute, 172(2), 164-175, (2018).
[25] Van Dinh, N. and Thao, N. X., “Some measures of picture fuzzy sets and their application in multi-attribute decision making”. Int. J. Math. Sci. Comput.(IJMSC), 4(3), 23-41, (2018).
[26] Ashraf, S., Abdullah, S., Mahmood, T., Ghani, F. and Mahmood, T., “Spherical fuzzy sets and their applications in multi-attribute decision making problems”. Journal of Intelligent and Fuzzy Systems, 36(3), 2829-2844, (2018).
[27] Kutlu Gündoğdu, F. and Kahraman, C., “Spherical fuzzy sets and spherical fuzzy TOPSIS method”. Journal of Intelligent and Fuzzy Systems, 36(1), 337-352, (2019).
[28] Zhang, X., “Multicriteria Pythagorean fuzzy decision analysis: A hierarchical QUALIFLEX approach with the closeness index-based ranking methods”. Information Sciences, 330, 104-124, (2016).
[29] Cuong, B. C. and Kreinovich, V., “Picture Fuzzy Sets-a new concept for computational intelligence problems”. In 2013 Third World Congress on Information and Communication Technologies (WICT 2013) (pp. 1-6). IEEE, 2013.
[30] Geetha, S. S. and Selvakumari, K., “A Picture Fuzzy Approach to Solving Transportation Problem”. European Journal of Molecular and Clinical Medicine, 7(2), 4982-4990, (2020).

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:47:37 UTC from IEEE Xplore. Restrictions apply.

Rendiconti Lincei. Scienze Fisiche e Naturali
https://doi.org/10.1007/s12210-021-01038-5

RESEARCH PAPER

Spatio-temporal variations of quality of rainwater and stormwater and treatment of stormwater runoff using sand–gravel filters: case study of Delhi, India

Harsh Pipil[1] · **A. K. Haritash**[1] · **Krishna R. Reddy**[2]

Received: 16 November 2021 / Accepted: 20 December 2021
© The Author(s), under exclusive licence to Accademia Nazionale dei Lincei 2022

Abstract
The present study was undertaken to characterise rainwater and stormwater runoff, along with reusability potential of treated stormwater runoff in Delhi, India. The analysis of rainwater and stormwater runoff samples revealed that the acidic ions (SO_4^{2-} and NO_3^-) in rainfall are being neutralised by alkaline species resulting in zero event of acidic rain (pH < 5.6) over Delhi. The higher level of alkaline species (Ca^{2+}, NH_4^+, HCO_3^-, etc.) in stormwater runoff may be attributed to the alluvial nature of soil and suspended dust which is chiefly being of crustal origin in Delhi. Non-sea-salt fraction (NSSF) indicated no marine contribution towards rainwater quality. Neutralization factor (NF) validated the dominance of alkaline species (NH_4^+, HCO_3^-, K^+, Ca^{2+}, and Mg^{2+}) over the acidic species (SO_4^{2-}, NO_3^-, Cl^- etc.), thereby resulting in near neutral pH of rainwater. Physico-chemical analysis of stormwater runoff from different locations in Delhi revealed relatively higher concentration of all the dissolved species (Ca^{2+}, Mg^{2+}, HCO_3^-, PO_4^{3-}, NO_3^-, Cl^-) along with higher level of suspended impurities and the spatial variations were observed to be regulated by land use landcover pattern. The treatment of stormwater runoff by sand–gravel filter resulted in complete removal of suspended impurities and partial removal of hardness. This study proposes additional use of a constructed wetland to remove the organic impurities, PO_4^{3-}, and NO_3^-. Furthermore, the study concludes that the stormwater runoff in Delhi can easily be treated to the desired level of drinking water supply with additional provision of disinfection to narrow down the demand–supply gap for drinking water.

Keywords Rainwater · Stormwater runoff · Stormwater treatment · Neutralisation factor · Sand–gravel filter

1 Introduction

India is the second most populous country supporting 1/6th of the world's population over approximately 2.4% of world's total land mass (Profile 2021). More than 60% population in India resides in cities and increasing population has led to an alarming increase in demand of the potable water for its people. The limited freshwater sources in urban areas, change in climatic conditions, uneven distribution of rainfall, and changes in land use pattern have forced to explore availability of potable water from alternate sources (Vialle et al. 2011). The rainwater is considered as the purest form in natural environment; therefore, collection and suitable use of rain water can be an effective solution for covering the demand–supply gap. The interface of rain with earth's surface results in the dissolution of minerals/pollutants, and suspension of silt, clay, and other organic impurities causing in relatively poor quality of stormwater runoff. Since most of the urban surfaces are paved (70%) and have high

This technical paper is submitted to Rendiconti Lincei. Scienze Fisiche e Naturali.

✉ A. K. Haritash
akharitash@dce.ac.in

Harsh Pipil
harsh_phd2k19@dtu.ac.in

Krishna R. Reddy
kreddy@uic.edu

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

[2] Department of Civil, Materials, and Environmental Engineering, University of Illinois, Chicago, IL 60607, USA

Published online: 10 January 2022

runoff coefficient (0.9), most of the wet precipitation turns to surface runoff which ultimately causes flooding in low lying areas, soil and nutrient loss, and limited recharge of aquifers (Goel 2011). The quality of stormwater runoff largely depends on land use, air quality, and associated activities. Generally, the natural wet precipitation has slightly acidic nature ($5.6 < pH < 7.0$) due to dissolution of atmospheric carbon dioxide, but the stormwater runoff remains alkaline due to dissolution of minerals and other alkaline species (Ca^{2+}, Mg^{2+}, HCO_3^-, etc.) from atmosphere suspended dust and soil (Khare et al. 2004). The quality of stormwater runoff varies spatially and temporally depending upon the activities and their intensity. Whereas the industrial areas are reported to have high concentrations of dissolved impurities and toxic elements (Ambade 2015) originating from anthropogenic activities, the coastal regions reported relatively higher levels of NaCl coming from natural salt spray (Das et al. 2005; Calvo et al. 2010). As the intensity and diversity of activities is more pronounced in developing countries like India along with the higher concentration of suspended dust (Balachandran and Khillare 2001), the stormwater runoff has high suspended solid content and other dissolved impurities. Moreover, the quality of rainwater and its runoff is of particular concern in cities that represent high level of air pollutants, suspended dust, and metals, as in case of Delhi, India (Deepika et al. 2016).

Delhi, the metropolitan city, is classified as one of the most polluted cities around the globe. The city houses more than 20 million people in an area of 1484 square km with a population density of 13,500 people/km^2, thus raising a concern for public health and infrastructure development (Directorate of Economics and Statistics Office of Chief Registrar 2021). It has the demand of 5183 MLD of treated water against the supply of 4251 MLD, creating a gap of 931 MLD, which projects to approximately 341,000 million litre on annual basis. On the other hand, Delhi receives annual rainfall of approximately 650 mm per year. Considering 70% paved surfaces with runoff coefficient of 0.9, the city produces surface runoff of approximately 963,000 million litre annually which is approximately 2.8 times higher than the annual demand–supply gap. If only one-third surface runoff is trapped, stored and utilised over the year, it can easily bridge the gap to meet the water demand of the city. Moreover, the utilisation of stormwater runoff would help overcome the problem of stormwater runoff mixing with sewage, additional cost of treatment, and minimisation of secondary sewage sludge generation.

The present study was undertaken in the capital city of Delhi, India, to determine the spatio-temporal variations of quality of rainwater and stormwater runoff. As the interface of precipitation with terrestrial environment results in poor quality of runoff, the study investigated the potential of sand–gravel filters towards the treatment of stormwater runoff and further explores the possibility for reuse of treated water.

2 Materials and methods

2.1 Collection of rainwater and stormwater runoff

The samples of wet precipitation ($n = 8$) were collected from east, south, southeast and northwest regions of Delhi representing residential, commercial, industrial, and institutional-cum-residential area, respectively, during the first rain (4th June, 2021—3.4 mm) and the intermediate rain (29th July, 2021—72 mm) of the Indian monsoon season (Fig. 1). Wet collector units were placed at the height of 1 m above the roof level to avoid collection of splash. Similarly, the representative samples of stormwater runoff ($n = 59$) were collected from different areas in Delhi (based on the land use and associated activities) during four events of rainfall in the year 2021 (Fig. 1).

2.2 Characterisation of rainwater and stormwater runoff

All of the collected samples were analysed immediately following the prescribed methods of APHA (2012). The pH, electrical conductivity (EC), and total dissolved solids (TDS) were analysed using the Orion Make (USA) multimeter; TSS was determined gravimetrically; cations (Na^+, K^+, Ca^{2+}, Li^+) were analysed by Flame photometer (Systronics 128μC); Aquaread (Model: AP-5000) multimeter was used for NH_4^+ and CDOM; UV–VIS Spectrophotometer (LABINDIA UV3092) was used for determining phosphates, sulphates and nitrates; Shimadzu TOC analyzer (TOC-LCPH) was used for TOC, TC, IC and TN, while, rest of the analysis was performed using the standard titrimetric methods as per the APHA standard methods (APHA 2012).

2.3 Source characterization

Based on the results obtained for characterisation of wet precipitation and runoff, the relative abundance of acidic and alkaline ionic species, enrichment factor (EF), neutralization factor (NF), and source characterisation were also calculated. Enrichment factor for any species X in rainwater can be calculated as:

$$EF_x = \left\{\left[X/Na^+\right]_{rain}\right\} / \left\{\left[X/Na^+\right]_{seawater}\right\},$$

where $[X/Na]_{rain}$ is the ratio of concentration of species X and sodium in rainwater. The value of $EF > 1.0$ indicates concentration of a particular chemical species due to anthropogenic activities.

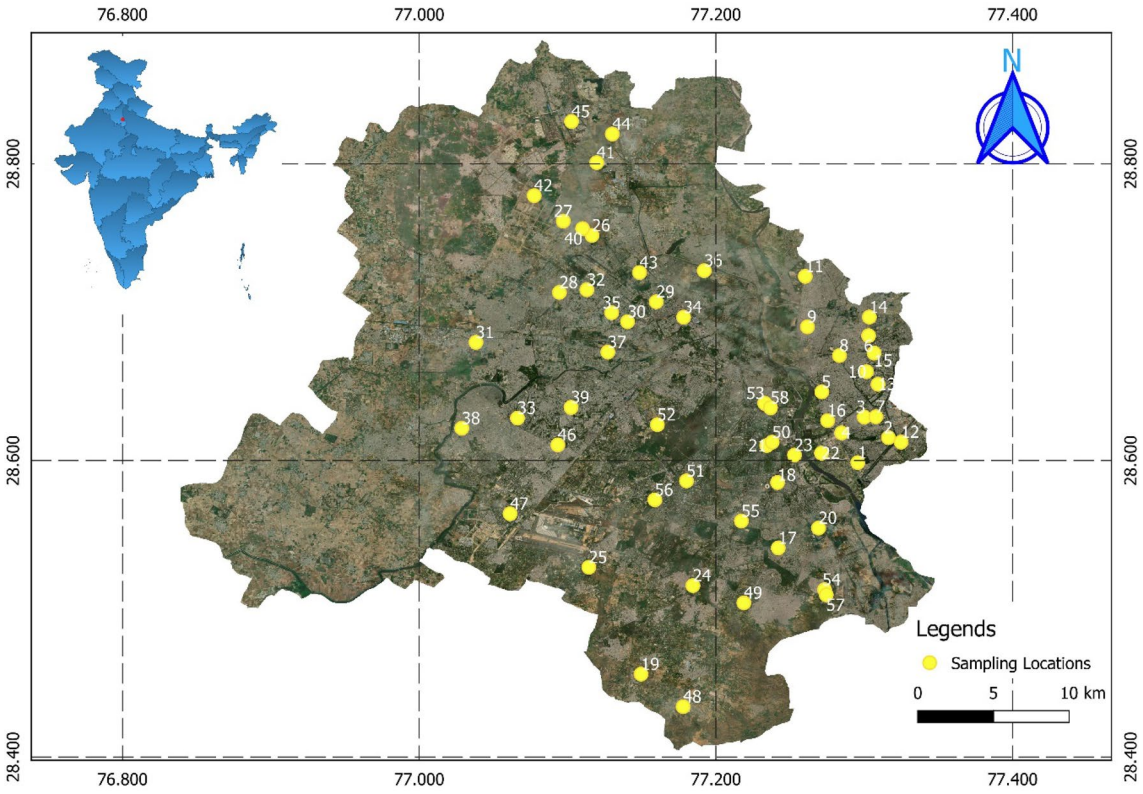


Fig. 1 Location of sampling points for stormwater runoff in Delhi, India

Neutralization factor (NF_x) is another expression for any species X, and it is calculated for validating the neutralization of acids by bases using

$$NF_x = [X] / \left[NO_3^- + SO_4^{2-}\right].$$

The $NF_x > 1.0$ denotes dominance of alkaline species (NH_4, HCO_3, K, Ca, Mg).

The non-sea-salt fraction (NSSF) of any species X in the rainwater was calculated using

$$NSSF_x = \left[X_{rain}\right] - \left\{\left[Na_{rain}\right] \times \left[X_{seawater} / Na_{seawater}\right]\right\},$$

where X_{rain} is the concentration of species X in rainwater, Na_{rain} is the concentration of sodium in rainwater. Non-sea-salt fraction (NSSF) was calculated to assess the role of local sea-salt resuspension or marine contribution towards the quality of rainwater.

2.4 Treatment of stormwater runoff

The typical stormwater runoff sample from DTU (institutional) was subjected to treatment through a sand–gravel filter system to assess the degree of treatment and the possibility towards reuse of treated water. Preliminary treatment testing was performed using the real stormwater runoff collected during the rainfall event in a separate gravel only column filter and sand–gravel column filter ($d_i = 1.2$ cm), which consisted of a 2.0 cm layer of filter material (gravel only or sand–gravel), operated at a flow rate of 20 ml/minute. The sand-gravel column filter consisted of a 1.0 cm sand layer over 1.0 cm gravel layer. The gravel used as filter media was a construction grade aggregate of size 5–7 mm available commercially. The void ratio (e) of the gravel bed was 0.46 ($\approx 46\%$) with bulk density of 1760 kg/m^3. For sand filter, construction grade river sand passing through 300 µm but retaining on 150 µm sieve, with void ratio (e) = 0.60 ($\approx 60\%$) was used.

3 Results and discussion

3.1 Quality of rainwater

The physico-chemical characterisation of rainfall revealed that most of the ionic species are present above the traceable concentrations (Table 1). Average pH of rainwater was observed to be 7.03 indicating that the alkaline species are available for neutralising the acidic ions (SO_4^{2-} and NO_3^-). The pH of 6.6 was observed during first rainfall as compared to the pH of 7.7 during the intermediate rain indicating the

Table 1 Primary characterization of rainfall and stormwater runoff in Delhi, India

Parameter	Rainfall ($n=8$)			Stormwater Runoff ($n=8$)		
	Min	Max	Mean ± SD	Min	Max	Mean ± SD
pH	5.7	7.9	7.0 ± 0.7	6.2	7.7	6.9 ± 0.6
EC (µS/cm)	36	243	110.6 ± 70.9	56	1174	336.2 ± 385.9
TDS (mg/l)	19	122	55.3 ± 35.4	28	586	179.1 ± 206.9
TSS (mg/l)	0	104	28.0 ± 36.7	28	932	205.8 ± 302.4

EC electrical conductivity, *TDS* total dissolved solids, *TSS* total suspended solids, *SD* standard deviation

Table 2 Neutralization capacity (NF) of major alkaline species present in rainwater

	NH_4^+	HCO_3^-	K^+	Ca^{2+}	Mg^{2+}
NF_1	0.25	2.42	0.39	24.46	30.46
NF_2	0.36	1.54	2.54	17.81	105.76

1 and 2 first and second rainfall spell, *NF* neutralization factor

availability of acidic species before the monsoon (owing to higher level of air pollutants), whereas the naturally originating alkaline species of crustal origin dominated during intermediate rain events during the monsoon. Relatively higher level of alkaline species (Ca^{2+}, NH_4^+, HCO_3^-, etc.) may be attributed to the alluvial nature of soil in Delhi region (Balachandran and Khillare 2001; Kulshrestha et al. 2003). Moreover, studies on chemical composition of rainfall have confirmed the alkaline nature of suspended dust in Indian sub-continent (Khare et al. 2004). No incident of acid rain was observed at any location during the study period in Delhi.

The neutralisation factor $(NH_4^+ + Ca^{2+}) / (NO_3^- + SO_4^{2-})$ (Khemani et al. 1994) values were more than 1.0 at all the locations indicating that the alkaline species dominated over the acidic species (Table 2). Further, the relative neutralising capacity of individual alkaline species calculated as per Khemani et al. (1994) revealed that calcium and magnesium were the dominant alkaline species followed by HCO_3^- that contributed towards neutralisation of acidic species. Similar observations have been made by Kulshrestha et al. (2003), Das et al. (2005) and Singh et al. (2007). Based on relative abundance of cations, it was observed that $Ca^{2+} > Mg^{2+} > Na^+ > NH_4^+ > K^+$ (Fig. 2a) indicating the order of neutralising capacity of alkaline species. Similarly, the anions $HCO_3^- > SO_4^{2-} > NO_3^- > Cl^- > PO_4^{3-}$ (Fig. 2a) also represented abundance of HCO_3^- as a neutralising species. Dominantly alkaline pH of rainwater in study area may also be attributed to precipitation of the anthropogenically produced acidic species (NO_3^- and SO_4^{2-}) during the heavy rainfall induced by Tauktae Cyclone which preceded the monsoon in year 2021. Later, dominantly dusty environment suspends the crustal dust dominated by Ca^{2+} and Mg^{2+} (Kulshrestha et al. 2003) which is reflected by the alkaline nature of precipitation. The relative contribution of acidic species towards inducing acidity in wet precipitation was calculated using NO_3^-/SO_4^{2-} ratio (Khemani et al. 1994). The value above 1.0 indicates the influence of HNO_3 towards acidity whereas values below 1.0 indicate chief contribution of H_2SO_4. It was observed that the acidity in rainwater was chiefly contributed by HNO_3 at all the locations except the dominance of H_2SO_4 in northwest, south, and southeast Delhi. The NF_x (Table 3) revealed relatively lower values during first rainfall event which may be attributed to relatively higher level of atmospheric acidic species (NO_3^- and SO_4^{2-}) which get flushed out during the first rain. Slightly higher levels in the intermediate events confirms that most of the acidic atmospheric species have already been scavenged

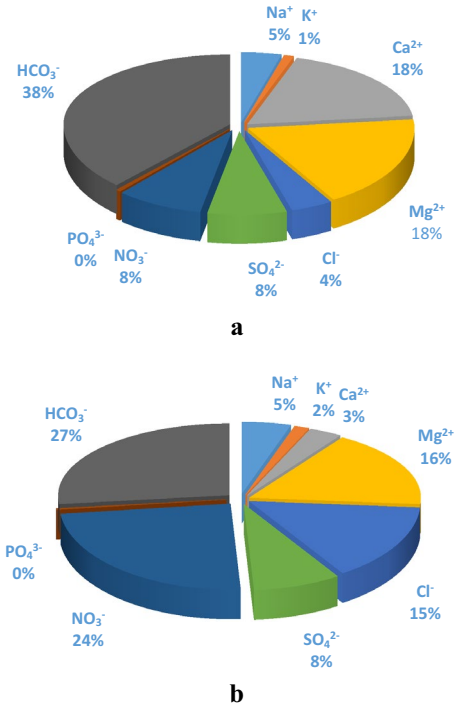


Fig. 2 **a** The percent contribution of cations and anions in rainwater in Delhi, India. **b** The percent contribution of cations and anions in stormwater runoff in Delhi, India

Table 3 Comparison of seawater ratio, enrichment factor and non-sea-salt fraction in rainwater

	Cl^-/Na^+	SO_4^{2-}/Na^+	K^+/Na^+	Ca^{2+}/Na^+	Mg^{2+}/Na^+
Seawater ratio	1.16	0.13	0.23	0.022	0.04
$Rainwater_1$	0.66	0.72	0.11	7.07	8.8
EF_1	0.57	5.76	0.5	324.19	200.52
$NSSF_1$	− 0.12	0.11	− 0.02	1.21	1.54
$Rainwater_2$	1.64	4.78	0.66	4.62	27.44
EF_2	1.14	38.23	2.9	212.04	625.14
$NSSF_2$	0.03	0.16	0.01	0.19	1.15

1 and 2 first and second rainfall spell, *EF* enrichment factor, *NSSF* non-sea-salt factor

during preceding events of rainfall, thus resulting in higher NF_x values for intermediate rain.

The negative values of NSSF (Table 3) for K^+ and Cl^- confirm no contribution of sea salt, whereas, positive value of NSSF for Ca^{2+} and Mg^{2+} confirmed the resuspension of crustal dust and its role in alkaline nature of rainwater (Table 3). The distant location of Delhi from coastal areas may be the reason for no contribution of salt spray from sea. On the other hand, proximity to Thar desert of Rajasthan and dust load from limestone and sand stone mines of Haryana results in concentration of Ca^{2+} and Mg^{2+} in atmospheric dust (Ambastha and Haritash 2021; Chitrakshi and Haritash 2018). Further, enrichment factor (EF) for different chemical species was also calculated. The results confirmed the enrichment of Mg^{2+}, Ca^{2+} and SO_4^{2-} from local source, i.e., the resuspended dust. Similar observations were also made by Khemani et al. (1994) for the Indian sub-continent indicating that the enrichment of these chemical species takes place due to resuspension of dust. The resuspension of dust may be attributed to natural dust-storms as well as fugitive emissions due to movement of vehicles (Ambastha and Haritash 2021).

3.2 Quality of stormwater runoff

The characterisation of stormwater runoff from different locations in Delhi revealed relatively higher concentration of all the dissolved species except a slight decrease in Ca^{2+} concentration which may be attributed to indirect base-exchange upon interaction with soil. Significant difference was observed in concentration of TSS, Cl^-, SO_4^{2-}, NO_3^-, and HCO_3^- (Table 4). Relatively higher level of TSS may be due to the suspension of silt and clay, whereas Cl^- and SO_4^{2-} are mainly contributed by mixing with domestic wastewater, and NO_3^- may be contributed from domestic wastewater, organic waste, compost/fertiliser applied in lawns and parks, etc. Based on relative abundance $HCO_3^- > NO_3^- > Mg^{2+} > Cl^- > SO_4^{2-} > Na^+ > Ca^{2+} > K^+$ (Fig. 2b) indicating that the contribution of crustal origin, organic waste, and mixing of domestic wastewater dominantly regulated the quality of stormwater runoff. Since the dust is a dominant feature of climate in India, its presence as TSS and Ca^{2+}, Mg^{2+}, HCO_3^-, etc. as dissolved impurities further confirms the source as natural.

The runoff collected from residential areas represented relatively higher level of TDS and EC which may be attributed to mixing of domestic wastewater (flow through open channels) with stormwater runoff. The level of TSS was observed high along the roadways and institutional areas which may be due to the suspension of deposited dust (during dry season) and accumulated street sweepings along the roads. It was observed that runoff from residential areas has relatively higher level of NH_4^+, NO_3^-, PO_4^{3-}, and TOC which may be attributed to the addition from organic matter, mulch, applied compost etc. in lawns, parks; littering of food waste, and excretory waste of stray animals and pets (Melidis et al. 2007; Gikas et al. 2012; Reddy et al. 2014). Upon comparison of runoff collected from different rainfall events, it was observed that the average stormwater runoff quality remained almost same with minimal variations. Site-specific activities associated with local environment represented variations with respect to a few parameters, e.g., the level of TSS was observed high owing to the construction activity of Central Vista Project in Central Delhi. It was observed that the TSS, Ca^{2+}, Mg^{2+}, HCO_3^-, PO_4^{3-}, Cl^-, and NO_3^- were the major parameters that require the treatment before reuse of the treated runoff.

3.3 Stormwater runoff treatment and reuse potential

The study revealed almost complete removal of total suspended solids while only 11% of TDS were removed from the stormwater runoff passed through the sand–gravel filter (Table 5). It was observed that the gravel filter can be used to remove the floating as well as suspended impurities from stormwater runoff, however, the TSS removal efficiency can be augmented by using sand layer along with gravel layer but it reduces the hydraulic conductivity of filter (Kandra et al. 2014). The increase in TDS of sand–gravel treated runoff might be attributed to exchange of 'Na' of sand with 'Ca' of water resulting in natural water-softening. Greeley and Leach (1978) observed that the particles with size larger than 60 µm carry positive charges and 'Na' remains as a part of adsorbed ions over the soil particles or as accumulated mineral in regions which have a natural source of Na^+. Given the higher adsorption potential of Ca^{2+} compared to Na^+, the replacement of Ca^{2+} by Na^+ takes place easily although the degree of ion exchange is considerably low. The results indicated 25–30% removal of hardness indicating that 'Ca'

Table 4 Measured stormwater runoff quality in different urban locations and comparison with drinking water standard (IS 10500:2012)

Parameters	DWS	Commercial ($n=9$)			Industrial ($n=6$)			Institutional ($n=9$)			Residential ($n=18$)			Road/Highway ($n=17$)		
		Min	Max	Mean ± SD	Min	Max	Mean ± SD	Min	Max	Mean ± SD	Min	Max	Mean ± SD	Min	Max	Mean ± SD
pH	6.5–8.5	6.2	7.9	7.01 ± 0.55	6.7	7.5	7.10 ± 0.30	6.0	8.2	7.23 ± 0.64	5.3	8.0	6.76 ± 0.70	6.1	8.9	7.31 ± 0.85
EC (µS/cm)	–	81.0	724.0	272.69 ± 213.42	117.9	485.0	251.78 ± 154.80	80.7	476.0	224.44 ± 133.22	56.1	1375.0	392.76 ± 427.10	21.8	661.0	225.70 ± 204.68
TDS (mg/l)	500	40.9	365.0	136.69 ± 107.41	59.3	242.0	125.92 ± 77.11	40.4	238.0	113.23 ± 65.52	28.0	686.0	201.67 ± 218.46	24.3	331.0	113.63 ± 99.55
TSS (mg/l)	–	48.0	800.0	374.22 ± 243.06	52.0	1150.0	314.67 ± 416.15	44.0	2308.0	504.89 ± 743.52	20.0	2004.0	382.33 ± 574.48	18.0	3448.0	692.18 ± 1114.45
Na^+ (mg/l)	–	0.8	48.6	14.87 ± 18.65	3.5	23.7	11.38 ± 8.77	0	45.4	11.99 ± 15.13	0	111.4	18.15 ± 27.16	0.4	24.0	7.64 ± 8.62
K^+ (mg/l)	–	2.4	31.6	10.17 ± 10.32	2.5	15.0	7.31 ± 5.35	0	26.3	6.79 ± 7.95	0	84.4	15.18 ± 24.17	1.3	34.07	9.04 ± 9.90
Ca^{2+} (mg/l)	75	3.6	28.3	13.09 ± 7.76	6.6	15.1	10.75 ± 3.57	0	45.8	17.08 ± 15.08	0	59.5	16.30 ± 16.68	0	22.8	9.32 ± 7.72
Mg^{2+} (mg/l)	30	4.9	175.0	57.21 ± 48.33	15.0	74.0	40.14 ± 24.71	7.9	326.0	72.32 ± 97.39	10.0	449.0	88.21 ± 103.76	3.9	250.0	55.33 ± 64.29
Li^+ (mg/l)	–	0.01	0.18	0.06 ± 0.05	0	0.2	0.09 ± 0.08	0	0.6	0.10 ± 0.20	0	0.3	0.05 ± 0.08	0	0.18	0.05 ± 0.05
NH_4^+ (mg/l)	–	0.6	6.8	2.45 ± 2.25	0.2	5.3	1.79 ± 1.91	0.6	3.3	1.61 ± 0.86	0.4	59.8	6.00 ± 13.88	0.15	10.2	2.67 ± 3.10
Cl^- (mg/l)	250	2.6	236.8	52.87 ± 82.69	3.3	131.8	35.07 ± 50.62	2.4	88.76	16.19 ± 27.70	1.4	343.7	72.20 ± 112.70	1.04	98.9	23.46 ± 34.06
SO_4^{2-} (mg/l)	200	5.8	50.6	14.71 ± 13.83	6.9	57.4	19.58 ± 19.03	5.8	162.3	31.89 ± 49.28	1.9	183.4	27.23 ± 42.50	3.8	16.9	8.69 ± 3.77
NO_3^- (mg/l)	45	2.1	122.8	42.30 ± 49.00	1.8	72.8	15.88 ± 27.95	2.3	62.0	22.80 ± 27.58	2.3	478.8	114.31 ± 126.97	1.6	124.5	28.97 ± 43.38
PO_4^{3-} (mg/l)	–	0.14	7.9	1.89 ± 2.42	0.2	5.5	1.62 ± 2.01	0.12	2.7	0.75 ± 0.87	0.04	17.4	2.07 ± 4.19	0.07	3.2	1.10 ± 1.08
HCO_3^- (mg/l)	–	0	311.1	120.31 ± 91.39	79.3	231.8	128.10 ± 61.97	36.6	198.3	97.60 ± 49.86	39.6	677.1	163.68 ± 183.08	24.4	408.7	113.21 ± 108.08
TN (mg/l)	–	0.9	4.9	2.60 ± 1.56	0.6	46.5	10.03 ± 18.01	0.8	5.1	2.92 ± 1.63	1.1	47.2	6.34 ± 10.84	0.9	7.6	3.07 ± 2.37
TKN (mg/l)	–	4.2	18.2	11.04 ± 4.87	4.2	7.0	6.07 ± 1.14	2.8	16.8	7.31 ± 4.52	2.8	40.6	11.90 ± 9.44	2.8	19.6	8.89 ± 4.51
CDOM (µg/l)	–	25.5	72.3	43.80 ± 15.48	24.2	74.7	43.62 ± 22.91	11.9	92.6	45.82 ± 29.39	12.0	194.6	51.20 ± 43.94	17.5	195.5	49.65 ± 46.25
TOC (mg/l)	–	3.6	44.6	17.69 ± 14.34	3.7	41.6	15.05 ± 14.39	2.1	30.9	11.70 ± 10.91	1.7	1002.5	101.75 ± 246.82	2.9	30.2	6.60 ± 6.54
TC (mg/l)	–	8.7	99.6	36.29 ± 28.91	10.9	65.0	30.78 ± 21.81	6.3	60.3	26.30 ± 17.70	6.7	1005.0	122.41 ± 248.69	6.3	98.9	31.16 ± 30.56
IC (mg/l)	–	5.2	54.9	18.73 ± 15.18	7.2	23.4	19.75 ± 7.07	4.2	29.4	13.97 ± 7.76	2.5	96.9	24.89 ± 29.00	3.4	53.9	15.74 ± 14.88

EC electrical conductivity, *TDS* total dissolved solids, *TSS* total suspended solids, *TN* total nitrogen, *TKN* total Kjeldahl nitrogen, *CDOM* coloured dissolved organic matter, *TOC* total organic carbon, *TC* total carbon, *IC* inorganic carbon, *DWS* drinking water standard, *SD* standard deviation

Table 5 Characteristics of stormwater runoff treated with gravel and sand–gravel filter

Filter type	Gravel filter			Sand–gravel filter		
Parameter	Inlet	Outlet	Efficiency (%)	Inlet	Outlet	Efficiency (%)
pH	9.4	9.0	–	6.4	8.4	–
EC (µS/cm)	220	194	11	396	704	–
TDS (mg/l)	110	97	11	198	356	–
TSS (mg/l)	82	14	83	605	0	100
Ca^{2+} (mg/l)	10	14	–	100	60	40
Mg^{2+} (mg/l)	10	1	90	150	120	20
Total Hardness (mg/l)	20	15	25	250	180	28
Cl^- (mg/l)	101	83	18	268	120	55
NO_3^- (mg/l)	62	53	15	59	13	78

EC electrical conductivity, *TDS* total dissolved solids, *TSS* total suspended solids

and 'Mg' are being removed. The sand–gravel filter was observed to remove NO_3^- as well but it saturated with time with accumulation of TSS.

The study proposes to use a hybrid configuration of constructed wetland and sand–gravel filter for removal of other pollutants and neutralise the variations of flow. The horizontal flow large-sized natural wetlands act as buffer, reduce the flow rate resulting in increased detention time, and consequently relatively higher removal efficiency. Several studies have reported the removal of phosphates and nitrates (Haritash et al. 2017; Nandakumar et al. 2019; Pipil et al. 2021) using *Canna* lily, *Phragmites* and *Brachiaria* based-constructed wetland (Haritash et al. 2015). In addition, plant-based remediation will also remove toxic metals (Karaca et al. 2018) and other organic impurities (Cameselle and Gouveia 2019).

Further, on comparison of stormwater runoff characteristics for different land use type with Indian standards for drinking water as per IS 10500:2012 (BIS 2012; Table 4), it can be said that average runoff characteristics lies near Indian drinking water standards. It can be intercepted by the combined filtration and wetland techniques to improve its quality to bridge the demand and supply gap for drinking water.

4 Conclusion

The characterisation of rainwater revealed that none of the rainfall events was acidic. The basic nature was attributed to richness of Ca^{2+}, Mg^{2+}, and HCO_3^- in suspended dust in India. The acidic ions (SO_4^{2-} and NO_3^-) in rainfall are being neutralised by alkaline species resulting in average neutral pH. The neutralisation factor greater than unity at all the locations confirmed the findings. The negative values of NSSF for K^+ and Cl^- confirmed no contribution of sea salt, while positive value of NSSF for Ca^{2+} and Mg^{2+} confirmed the resuspension of crustal dust and its role in alkaline nature of rainwater. Stormwater runoff from residential areas had relatively higher level of NH_4^+, NO_3^-, PO_4^{3-}, and TOC making it difficult to treat using a sand–gravel filter. Further, the stormwater runoff from other areas can be treated with a sand–gravel filter to remove the objectionable suspended impurities and part of hardness inducing species (Ca and Mg). It is proposed to configure constructed wetlands and sand–gravel filter to treat stormwater runoff to desired level for organic impurities, NO_3^-, PO_4^{3-} and other pollutants. The treated stormwater runoff can be augmented with conventional treatment followed by disinfection to meet the demand for drinking water supply for approximately 300 billion litres of treated stormwater runoff for public consumption in a sustainable manner.

Acknowledgements The authors are grateful to Shivani Yadav, Harshit Chawla, and Sonam Taneja, all PhD scholars, in the Department of Environmental Engineering at DTU for their assistance during this study.

Author contributions HP: conceptualization, execution, data compilation and draft writing; AKH: conceptualization, editing, supervision, reviewing and editing; KRR: conceptualization, editing, supervision, reviewing and editing.

Declarations

Conflict of interest The authors have no conflicts of interest to declare that are relevant to the content of this article.

Human and animal rights The authors declare that no involvement of human and animals in the research work.

References

Ambade B (2015) Chemical composition of runoff water in Raipur city, central India. Appl Water Sci 5:1–12. https://doi.org/10.1007/s13201-014-0163-0

Ambastha SK, Haritash AK (2021) Emission of respirable dust from stone quarrying, potential health effects, and its management. Environ Sci Pollut Res. https://doi.org/10.1007/s11356-021-16079-4

APHA (2012) Standard methods for the examination of water and wastewater, 22nd edn. American Public Health Association American Water Works Association, Water Environment Federation, Washington

Balachandran S, Khillare PS (2001) Occurrence of acid rain over Delhi. Environ Monit Assess 71:165–176. https://doi.org/10.1023/A:1017541809985

Beuro of Indian Standards for Drinking water 2012 (BIS 2012). https://doi.org/10.1016/B978-0-12-812843-5.00018-6

Calvo AI, Olmo FJ, Lyamani H, Alados-Arboledas L, Castro A, Fernández-Raga M, Fraile R (2010) Chemical composition of wet precipitation at the background EMEP station in Víznar (Granada, Spain) (2002–2006). Atmos Res 96(2–3):408–420. https://doi.org/10.1016/j.atmosres.2010.01.013

Cameselle C, Gouveia S (2019) Phytoremediation of mixed contaminated soil enhanced with electric current. J Hazard Mater 361:95–102. https://doi.org/10.1016/j.jhazmat.2018.08.062

Chitrakshi, Haritash AK (2018) Hydrogeochemical characterization and suitability appraisal of groundwater around stone quarries in Mahendragarh, India. Environ Earth Sci 77:252. https://doi.org/10.1007/s12665-018-7431-5

Das R, Das SN, Misra VN (2005) Chemical composition of rainwater and dust fall at Bhubaneswar in the east coast of India. Atmos Environ 39(32):5908–5916. https://doi.org/10.1016/j.atmosenv.2005.06.030

Deepika, Gour P, Haritash AK (2016). Air pollution tolerance of trees in an educational institute in Delhi. Int J Environ Sci 6(6):979–986. http://10.0.23.200/ijes.6092

Directorate of Economics and Statistics Office of Chief Registrar (2021) Profile of Delhi: National Capital Territory – Delhi. http://des.delhigovt.nic.in/DoIT/DOIT_DM/district%20profile.pdf. Accessed 2 Oct 2021

Gikas GD, Tsihrintzis VA (2012) Assessment of water quality of first-flush roof runoff and harvested rainwater. J Hydrol 466:115–126. https://doi.org/10.1016/j.jhydrol.2012.08.020

Goel MK (2011) Runoff Coefficient. In: Singh VP, Singh P, Haritashya UK (eds) Encyclopedia of Snow, Ice and Glaciers. Encyclopedia of Earth Sciences Series. Springer, Dordrecht. https://doi.org/10.1007/978-90-481-2642-2_456

Greeley R, Leach R (1978). A preliminary assessment of the effects of electrostatics on aeolian process. Rep Planet Geol Program, NASA TM79729, 236–237

Haritash AK, Dutta S, Sharma A (2017) Phosphate uptake and translocation in a tropical Canna-based constructed wetland. Ecol Process 6:12. https://doi.org/10.1186/s13717-017-0079-3

Haritash AK, Sharma A, Bahel K (2015) The potential of Canna lily for wastewater treatment under Indian conditions. Int J Phytoremediation 17(10):999–1004. https://doi.org/10.1080/15226514.2014.1003790

Kandra HS, Deletic A, McCarthy D (2014) Assessment of impact of filter design variables on clogging in stormwater filters. Water Resour Manage 28:1873–1885. https://doi.org/10.1007/s11269-014-0573-7

Karaca O, Cameselle C, Reddy R (2018) Mine tailing disposal sites: contamination problems, remedial options and phytocaps for sustainable remediation. Rev Environ Sci Biotechnol 17:205–228. https://doi.org/10.1007/s11157-017-9453-y

Khare P, Goel A, Patel D, Behari J (2004) Chemical characterization of rainwater at a developing urban habitat of Northern India. Atmos Res 69(3–4):135–145

Khemani LT, Momin GA, Rao PSP, Pillai AG, Safai PD, Mohan K, Rao MG (1994) Atmospheric pollutants and their influence on acidification of rain water at an industrial location on the west coast of India. Atmos Environ 28(19):3145–3154. https://doi.org/10.1016/1352-2310(94)00148-E

Kulshrestha UC, Granat L, Rodhe H (2003) Precipitation chemistry studies in India-a search for regional patterns. Report CM-99 Department of Meteorology, Stockholm University

Melidis P, Akratos CS, Tsihrintzis VA, Trikilidou E (2007) Characterization of rain and roof drainage water quality in Xanthi, Greece. Environ Monit Assess 127(1):15–27. https://doi.org/10.1007/s10661-006-9254-1

Nandakumar S, Pipil H, Ray S, Haritash AK (2019) Removal of phosphorous and nitrogen from wastewater in Brachiaria-based constructed wetland. Chemosphere 233:216–222. https://doi.org/10.1016/j.chemosphere.2019.05.240

Pipil H, Haritash AK, Reddy KR (2021) Seasonal variability and kinetics of phosphate removal in a Phragmites-based engineered wetland. Rend Fis Acc Lincei. https://doi.org/10.1007/s12210-021-01017-w

Profile (2021) National portal of India. https://www.india.gov.in/india-glance/profile. Accessed 28 Sept 2021

Reddy KR, Xie T, Dastgheibi S (2014) Removal of heavy metals from urban stormwater runoff using different filter materials. J Environ Chem Eng 2(1):282–292. https://doi.org/10.1016/j.jece.2013.12.020

Singh AK, Mondal GC, Kumar S et al (2007) Precipitation chemistry and occurrence of acid rain over Dhanbad, Coal City of India. Environ Monit Assess 125:99–110. https://doi.org/10.1007/s10661-006-9243-4

Vialle C, Sablayrolles C, Lovera M, Jacob S, Huau MC, Montrejaud-Vignoles M (2011) Monitoring of water quality from roof runoff: Interpretation using multivariate analysis. Water Res 45(12):3765–3775. https://doi.org/10.1016/j.watres.2011.04.029

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/356950540

Stumpy-Paced Green Power Generation System

Conference Paper · October 2021

DOI: 10.1109/RDCAPE52977.2021.9633633

CITATIONS
0

READS
15

9 authors, including:



Ashok Kumar Singh
Galgotias College of Engineering and Technology (GCET)
29 PUBLICATIONS **292** CITATIONS
SEE PROFILE



Phool Singh
Delhi Technological University
3 PUBLICATIONS **0** CITATIONS
SEE PROFILE



Rajendra Prasad
Galgotias College of Engineering and Technology (GCET)
16 PUBLICATIONS **62** CITATIONS
SEE PROFILE

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

Solar still View project

Solar distillation system View project

All content following this page was uploaded by Phool Singh on 28 December 2021.

The user has requested enhancement of the downloaded file.

Stumpy-Paced Green Power Generation System

Ashok Kumar Singh[a], P. Kumar[b], R. Kishore[b], Phool Singh[c], Neeraj Kant[c], Pushpendra Singh[c], Gopal Nandan[d], A. Jaiswal[a], R. Prasad[b]

[a]*Mechanical Engineering Department, Galgotias College of Engineering and Technology, Greater Noida, UP, India* [b]*Applied Sciences Department, Galgotias College of Engineering and Technology, Greater Noida, UP, India*

[c]*Mechanical Engineering Department, Delhi Technological University, Delhi, India*

[d]*Mechanical Engineering Department, Amity University Uttar Pradesh, Noida, India*

[*]Corresponding author

agashok26@gmail.com*, pratap.kumar@galgotiacollege.edu, dr.rajeevkishorepandey@gmail.com, phoolsinghdce@gmail.com, neerajkant00007@gmail.com, Dr.psinghs@gmail.com, gnandan@amity.edu, megcet097@gmail.com, rajendra.prasad@galgotiacollege.edu

2021 4th International Conference on Recent Developments in Control, Automation & Power Engineering (RDCAPE) | 978-1-6654-1429-6/21/$31.00 ©2021 IEEE | DOI: 10.1109/RDCAPE52977.2021.9633633

Abstract: **In the present study, it has been emphasized that the specified gear mechanism or techno-suitable power transmission device may help to improve the power generation capabilities of the turbo-ventilators working at low-medium speeds. This set-up is indeed required to establish a clean and ventilated space at work-stations, theaters, assembly halls, stadiums, restraints, offices, etc. In this study, a generalized observation has been overviewed based on its development, and uses. Further, with a suggested utilization of gear mechanism with other associations along with a fair discussion and conscious conclusions have been given in the present work of approach. In this way, this communication builds a indeed gate way towards the utilization of freely available clean and green natural energy for saving the human beings and society in an eco-friendly way.**

Keywords: Roof top ventilator; Electric power; Green energy; Stumpy-paced.

I. INTRODUCTION

The industries, and other indoor spaces continuously producing hot air due to absorption of heat from heat libration electrical-electronic appliances, human bodies, etc. which goes out to the atmosphere due to lighter in weight and density in the replacement of fresh air coming through a confined passage of ventilators from the outside atmosphere. The freely out going hot air through a confined passage is further utilized to run the roof top turbo ventilators from one of the primary source and another one is air velocity available from open atmosphere as a natural renewable sources. Whereas, other natural resources as solar energy is also helpful to boost up the performance of renewable energy systems, along with the solar heat receivers [1] and related other heat receiving techniques [2, 3] like solar water heat treatment [4, 5], steam production [6], electrical power production [7], under the influence of ecological patterns [8, 9] energy matrix [10, 11] with economical observation and perspectives [12, 13]. In these, the renewable energy utilizations are involved for power production without any environmental harm. Further, India installed power production system through wind renewable energy in February 2021, which is the 4th biggest renewable power generation plant [14]. This project utilizes the available wind energy at variable wind velocities to run the system. Also, this wind operated ventilators are efficiently working for different wind directions (horizontal, axial, radial, and vertical) for the production of electric power [15]. As, these wind ventilators are required to operate with sufficient wind velocities in an arid zone only as represented in Fig. 1.

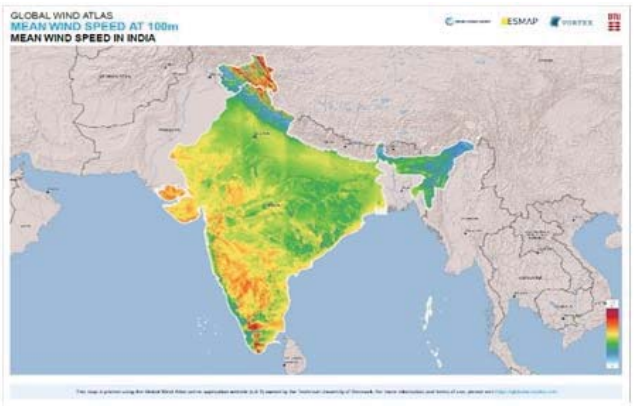


Fig. 1. Average wind velocity in Indian continent

So, it is better to use turbo-ventilators working at low-medium a speed which establishes a clean and ventilated space at work-stations, theaters, assembly halls, stadiums, restraints, offices, etc. and further, utilized for the production of electrical power also [16]. The wind operated roof top turbo ventilators are effective for high regions, arid areas for which low-medium speed wind can operate the mechanism of these ventilators to establish a clean and ventilated space at work-stations. These set up utilizes spherically designed vents made up of metals (erosion resistant, corrosion resistant with low thermal expansion coefficients). The roof top installed turbo wind ventilators are rotating with the effect of hot air coming from indoor spaces, electrical, electronic appliances, human bodies, and related objects, etc. and this rotates the turbo wind ventilators. As a result, a wind siphon effect developed that drops the air pressure significantly into the working space of work-stations, theaters, assembly halls, stadiums, restraints, offices, etc. Finally, the wind siphon effect and air pressure drop cumulatively sucks the fresh air from outside atmosphere through fresh air vent passages [17]. As per the global power generation which is mainly dependent on fossil fuels (>60%) as shown in Fig. 2. Now in the recent years, consumption of renewable energy sources are increased (>8%) globally. So the carbon savings and environmental protection are further keep better as shown in Fig. 3.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 20,2021 at 06:48:48 UTC from IEEE Xplore. Restrictions apply.

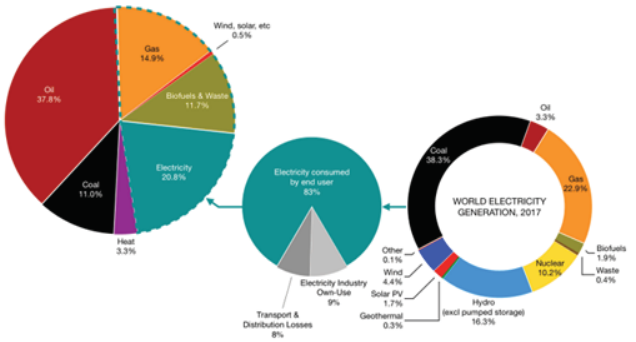


Fig. 2. Global energy facts [16]

The fundamental objectives of the present study are focusing the following frame of scheme as written herewith.

a) To study the development in the field of roof top installed wind ventilator fans and related technologies inbuilt within and the regions for which, these techniques are quite suitable for the applications.
b) To find the opportunities for low wind or low arid areas under the mono directional use of these roof top installed wind ventilator

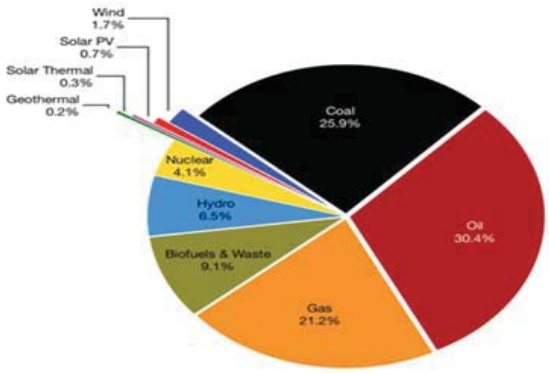


Fig. 3. Bifurcation of different energy possessions [16]

c) To apply sophisticated additional gear mechanism to improve the power output of the overall establishment (roof top installed wind ventilator) under the low speed wind capacity by increasing the rotation of wind ventilator metallic vanes.

II. CONSTRUCTIONAL DETAILS

It has been observed that the specified gear mechanism or techno-suitable power transmission device may help to improve the power generation capabilities of the turbo-ventilators working at low-medium speeds. And, such set-ups are required to establish a clean and ventilated space at work-stations, theaters, assembly halls, stadiums, restraints, offices, etc. under the wind potential capabilities. The wind ventilator regulates the warm air of occupied space from fresh air of atmospheric air through the air vent passages (metallic) as represented in Figs. 4(a)-4(b). As the hot air (heated up by accumulating heat coming out from various heat releasing sources available inside the workstation) is lighter in weight due to lowered density than the colder fresh air and this causes dual effect to run the roof top ventilators as mentioned,

a) The development of pressure drop at the ventilator vanes due to the upside movement of hot air.
b) Extra push developed by the fresh air over the hot air that provides additional up-lift thrust force and corresponding velocity to the hot air.

In this way, roof top turbo ventilators are continuously spinning throughout the day due to air exchange circulations.

In the development of turbo ventilators, computational fluid dynamic analyses is one of the important tool to simulate the air flow and heat effect for the variable vent (throat) diameter which had been taken as 300mm for greater circulation speed under given air flow [18]. Further, axial flux type permanent magnet (AFPM) as shown in Fig. 5 and Table 1 which is utilized to generate the corresponding force in between the armature winding, and during the rotation of armature, direct current (DC) is induced that can be utilized in a controlled manner to run the suitable electrical appliances [21]. All these combination with a wind turbine make the roof top turbo ventilators as a mini power plant [19]. Also, the DC power is converted accordingly based on the suitability of the available electrical appliances for the sound use of those [20]. The developed power can be utilized to improve the operating conditions and performance of other renewable energy systems, as additional benefits of such stumpy-paced green power generation systems for the particular solar operating devices such as solar desaltification systems utilized to treat the brackish water [22].

Table 1 Variation in speed of rotation against the wind velocities along with the effect or without the effect for AFPM base gears mechanism [21].

Wind velocity (m/s)	Speed of Rotation (rpm)	AFPM Rotational speed without gear (rpm)	AFPM Rotational speed with gear (rpm)
1.0	30	30	60
2.0	45	45	90
2.75	56	56	112
3.44	62	62	124
4.55	107	107	214
5.25	120	120	240
7.60	132	132	264
8.87	151	151	302
9.5	195	195	390

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 20,2021 at 06:48:48 UTC from IEEE Xplore. Restrictions apply.



Fig. 4 (a). Roof top ventilator

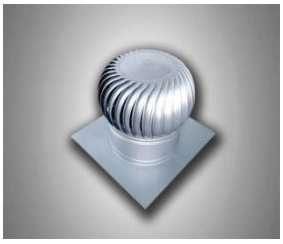


Fig. 4 (b). Roof top ventilator

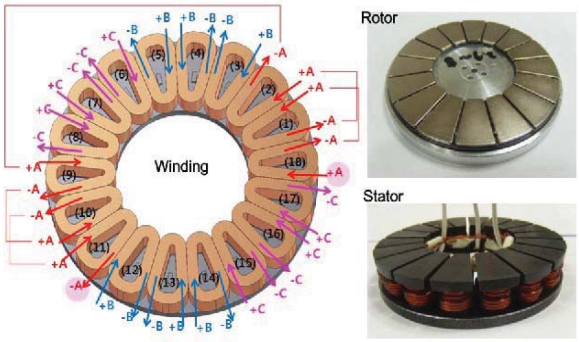


Fig. 5. Axial flux type permanent magnet generator and rectifier [21]

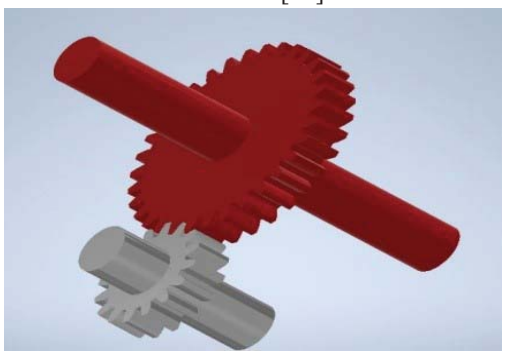


Fig. 6. Three dimensional model of gear mechanism

Also, a sophisticated gear mechanism (shown in Fig. 6) with the detailed specification and design parameters as given in Table 2 for the proper transmission of generated mechanical power towards the electric power generator. All these components are sequentially assembled as shown in Fig. 7, which makes the complete establishment as a roof top turbo ventilator with a proficient power transmission system (gear mechanism) and power generator.

Table 2 Gear mechanism association and design parameters utilized for a particular system

Design parameters	**Values**
Module (mm)	2
Pinion teeth	16
Gear teeth	32
Gear ratio	2
Pinion-gear center (mm)	48
Circular pitch of pinion (mm)	32
Circular pitch gear (mm)	64
Addendum (mm)	2
Dedendum (mm)	2.5
Tooth thickness (mm)	3.2
Bottom clearance (mm)	0.5
Whole depth (mm)	4.5

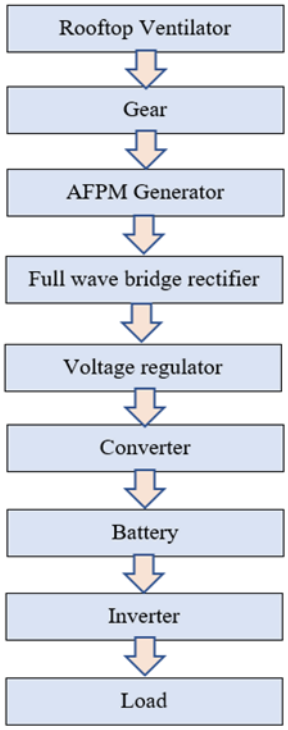


Fig. 7. Probable layout of the assembly

III. WORKING

With the help of a specified gear mechanism or techno-suitable power transmission device, one can improve the power generation capabilities of the turbo-ventilators working at low-medium speeds. And, such set-ups are required to establish a clean and ventilated space at work-stations, theaters, assembly halls, stadiums, restraints, offices, etc. under the wind potential capabilities. The wind ventilator regulates the warm air of occupied space from fresh air of atmospheric air through the air vent passages. As the hot air (heated up by accumulating heat coming out from various heat releasing sources available inside the workstation) is lighter in weight due to lowered density than the colder fresh air with the dual effect to run the roof top ventilators by developing the pressure drop at the ventilator vanes due to the upside movement of hot air and extra push offered by the fresh air over the hot air that provides additional up-lift thrust force and corresponding velocity to the hot air. In this way, roof top turbo ventilators are continuously spinning throughout the day due to air exchange circulations. In the development of turbo ventilators with better effect of air flow, the variable vent (throat) diameter may be taken as 300mm for greater circulation speed under given air flow [18]. Further, axial flux type permanent magnet (AFPM) which is utilized to generate the corresponding force in between the armature winding, and during the rotation of armature, direct current (DC) is induced that can be utilized in a controlled manner to run the suitable electrical appliances. All these combination with a wind turbine make the roof top turbo ventilators as a mini power plant. Also, the DC power is converted accordingly based on

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 20,2021 at 06:48:48 UTC from IEEE Xplore. Restrictions apply.

the suitability of the available electrical appliances for the sound use of those.

IV. DISCUSSION

It is quite necessary to have a sound association in between the integrated components of the roof top turbo ventilators with gear mechanism and electric power generator for the development of greater amount of electricity with least losses (mechanical or electrical). The gear mechanism is one of the most important linking elements which may be further tested on real background for its feasible viability and potential of utilization during the service of turbo ventilators.

Hence, the proper utilization of green energy sources along with the renewable energy sources of nature accumulatively builds an eco-friendly mechanism and sustainable technology for the production of a substantial amount of electric power. These electric powers can be utilized to run electrical appliances as fan, LED bulbs, blowers, coolers, etc. by converting it into alternating current power where ever necessary.

V. CONCLUSION

In the current study, the generalized idea of working, construction, and the developments of roof top turbo ventilators have been revealed. The conscious conclusions of the present study have been framed as written here under.

a) Sophisticated association is between the integral parts is utmost required to produce greater electric power with least thermal, mechanical, and electrical losses.
b) Gear mechanism is one of the important mechanical power transmission element, which requires on line testing for it proper and efficient functioning, that ensures smooth working of he system as well as transmits greater mechanical power to the generator.
c) Roof top turbo ventilators are quite successful for the domestic purposes also, and should be utilized by modifying the gear mechanism assembly along with the ventilator and generator system.

REFERENCES

[1] A.K. Singh, Samsher, Analytical study of evacuated annulus tube collector assisted solar desaltification system: A review, Solar Energy 207 (2020) 1404–1426.

[2] A.K. Singh, D.B. Singh, A. Mallick, Harender, S.K. Sharma, N. Kumar, V.K. Dwivedi, Performance analysis of specially designed single basin passive solar distillers incorporated with novel solar desalting stills: A review, Solar Energy, 185 (2019) 146–164.

[3] A.K. Singh, D.B. Singh, V.K. Dwivedi, N. Kumar, J.K. Yadav, A Review of Performance Enhancement in Solar Desalination Systems with the Application of Nanofluids, ICACCCN, (2018) 814 – 819.

[4] A.K. Singh, D.B. Singh, V.K. Dwivedi, G.N. Tiwari, A. Gupta, Water purification using solar still with/without nano-fluid: A review, Materials Today: Proceedings 21 (2020) 1700–1706.

[5] A.K. Singh, Samsher, A Review Study of Solar Desalting Units with Evacuated Tube Collectors, Journal of Cleaner Production 279 (2021) 123542.

[6] A.K. Singh, An inclusive study on new conceptual designs of passive solar desalting systems, Heliyon 7 (2020) e05793.

[7] A.K. Singh, R.K. Yadav, D. Mishra, R. Prasad, L.K. Gupta, P. Kumar, Active solar distillation technology: A wide overview, Desalination 493 (2020) 114652.

[8] A.K. Singh, B.K. Srivastava, V. Shankar, R. Prasad, R.K. Yadav, A.K. Shukla, G. Nandan, P. Kumar, Energy matrices and enviroeconomical perspective of solar decontamination system, Materials Today: Proceedings, (2020). https://doi.org/10.1016/j.matpr.2020.09.198

[9] A.K. Singh, D. Singh, M.K. Lohumi, B.K. Srivastava, H.P. Gupta, R. Prasad, Potable Water Production by Single Slope Active Solar Distillation Unit-A Review. Advances in Energy and Environment. Lecture Notes in Civil Engineering, Springer 142 (2021) 31-42.

[10] A.K. Singh, D.B. Singh, A. Mallick, N. Kumar, Energy matrices and efficiency analyses of solar distiller units: A review. Solar Energy, 173 (2018) 53 – 75.

[11] A.K. Singh, Samsher, Material conscious energy matrix and enviro-economic analysis of passive ETC solar still, Materials Today: Proceedings 38 (2021) 1-5.

[12] A.K. Singh, G. Dwivedi, B.K. Srivastava, B.K. Tiwari, P. Kumar, R.K. Yadav, M. Singh, A.K. Shukla, G. Nandan, Study of analytical observations on energy matrix for solar stills, Materials Today: Proceedings, (2021). https://doi.org/10.1016/j.matpr.2021.02.109

[13] A.K. Singh, D. Singh, M.K. Lohumi, B.K. Srivastava, H.P. Gupta, R. Prasad, Potable Water Production by Single Slope Active Solar Distillation Unit-A Review. Advances in Energy and Environment. Lecture Notes in Civil Engineering, Springer 142 (2021) 31-42.

[14] A. Kumar, K. Kumar, N. Kaushik, S. Sharma, S. Mishra, Renewable energy in India: Current status and future potentials, Renewable and Sustainable Energy Reviews 14 (8) (2010) 2434-2442.

[15] R. Shinde, V. Lavhale, A. Nair, S. Pawar, R. Mahajan, Generation of Electric Power using Turbo Ventilators,

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 20,2021 at 06:48:48 UTC from IEEE Xplore. Restrictions apply.

International Journal of Current Engineering and Technology 6 (2016) 89 – 93.

[16] G.K. Jadhav, P.M. Ghanegaonkar, Performance Characteristics of Turbo Ventilator: A Review, International journal of innovative technology and research volume 3 (4) (2015) 2276 – 2282.

[17] K. Dongre, A.A. Meshram, A.R. Bhoyar, R.U. Tamgade, A.R. Kapse, S. Jethani, Electricity generation by using rooftop ventilator, International Research Journal of Engineering and Technology 7(2) (2020) 3006-3008.

[18] G.K. Jadhav, P.M. Ghanegaonkar, S. Garg, Experimental and CFD Analysis of Turbo Ventilator, Journal of Building Engineering 6 (2016) 196-202.

[19] S. Dangeam, An electric generator is driven by a roof ventilator, Energy Procedia 9 (2011) 147-158.

[20] D.C. Shatri, M. Irwanto, A.N. Syafawati, S.S. Shema, Power generation roof ventilator, International Conference on Environment and Industrial Innovation 12 (2011) 183-187.

[21] M.R. Minaz, M. Celebi, Design and analysis of a new coreless PMSG with three rotor and double stator, Results in Physics 7 (2017) 183-188.

[22] A.K. Singh, Samsher, Tech-En-Econ-Energy-Exergy-Matrix (T4EM) Observations of Evacuated Solar Tube Collector Augmented Solar Desaltification Unit: A Modified Design Loom, Materials Today: Proceedings, (2021). https://doi.org/10.1016/j.matpr.2021.09.088

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 20,2021 at 06:48:48 UTC from IEEE Xplore. Restrictions apply.
View publication stats

Materials Research Bulletin 150 (2022) 111753

Contents lists available at ScienceDirect

Materials Research Bulletin

journal homepage: www.elsevier.com/locate/matresbu

Synthesis optimisation and efficiency enhancement in Eu^{3+} doped barium molybdenum titanate phosphors for w-LED applications

Pooja Rohilla, A.S. Rao *

Department of Applied Physics, Delhi Technological University, Bawana Road, New Delhi 110 042, India

ARTICLE INFO

Keywords:
Phosphors
XRD
SEM
Raman spectrum
Photoluminescence
Photoluminescence decay
CIE diagram

ABSTRACT

Barium Molybdenum Titanate (Ba_3MoTiO_8) phosphor activated with Eu^{3+} ions has been synthesized successfully by three different routes such as solid-state reaction method (SSM), combustion method (CMM), and sol-gel combustion method (SGM) to optimize relatively best synthesis technique and temperature of synthesis. X-Ray Diffraction (XRD) patterns have been used to confirm the phase and crystallinity of the as-prepared phosphor. Scanning Electron Microscope (SEM) and Raman spectra have been used to study the morphological behaviorhas and different vibrational bands present in the structure of the titled phosphor respectively. Diffuse Reflectance Spectra (DRS) have been recorded to find out the optical band gap of phosphors. Photoluminescence (PL) spectral features recorded under 393 nm and 465 nm excitation wavelengths reveal an intense red emission at 612 nm corresponding to $^5D_0 \rightarrow {}^7F_2$ transition of the doped Eu^{3+} ions. CMM synthesized sample has shown maximum luminescence intensity out of the three methods and further studies have been carried out by preparing a series of europium ions doped Ba_3MoTiO_8 phosphor using the CMM technique. The PL decay curves recorded for 612 nm red emission under 393 nm excitation show non-linear double exponential nature with a decrease in an experimental lifetime with an increase in Eu^{3+} ions concentration in the as-prepared phosphors. All the morphological and photoluminescence studies carried out on the Eu^{3+} ions doped Ba_3MoTiO_8 phosphor contemplate the suitability of the as-prepared phosphor for a red component in white light-emitting diode (w-LEDs) applications.

1. Introduction

In this digitalized world, energy plays an important role in various sectors affecting social and economic development. The ever-increasing demand for energy causes irreversible changes in our surrounding environment. Among the total energy produced from various sectors, a major share is utilized for indoor as well as outdoor lighting purposes. A huge mismatch in energy production and consumption is being caused, may be due to the inefficiency of gadgets which gave low output yield and exhaust huge amount of energy. The conventional sources of energy including fossil fuels, incandescent, and fluorescent lamps cause a large loss of energy in the form of heat and radiation, while also not being environment friendly. It is important to utilize the available energy in an efficient, sensible, and sustainable way. Thus developing alternate devices and technology such as solid-state lighting (SSL) devices are showing the possibility of excellence in next generation illumination sources when compared to fluorescent and incandescent lamps due to their high efficiency, larger lifetime, low cost of preparation, and eco-friendly behavior [1–5]. SSL technology is basically based on forward-biased light emitting p-n junction semiconductor diodes. Phosphor converted white light-emittingare diodes (pc-w-LEDs) is a core advancement in this SSL illumination technology. The commercial method that we use currently to generate white light is a coating of yellow phosphor (YAG: Ce^{3+}) on a blue LED chip. But the light generated from this method has a low color rendering index (CRI) and high correlated color temperature (CCT) due to deficiency of the red component. Another prominent method used to generate white light is coating of RGB (red, green, and blue) phosphors on n-UV LED chip or coating of R/G phosphor on blue LED chip [6–10]. Usage of these techniques can effectively improve CRI and CCT values in comparison to the usage of yellow phosphor with blue LED chip [11]. White light can also be generated by the combination of n-UV LED + RGB phosphor. This technique is not a widely used one because the commercial red phosphor ($Y_2O_2S:Eu^{3+}$) used in this technique is not a good absorbent of light generated by the n-UV chip [5,11–14]. Moreover, it has a poor lifetime under UV excitation and is chemically unstable [15]. Efforts are

* Corresponding author.
E-mail address: drsrallam@gmail.com (A.S. Rao).

https://doi.org/10.1016/j.materresbull.2022.111753
Received 2 November 2021; Received in revised form 23 December 2021; Accepted 15 January 2022
Available online 17 January 2022
0025-5408/© 2022 Elsevier Ltd. All rights reserved.

being made continuously so that a new efficient red phosphor can be made which is a very good absorbent of n-UV light [16].

Inorganic host-based phosphors doped with rare-earth (RE) ions are the main constituents of pc-w-LEDs. In general, as a host material binary or ternary compounds of oxides, nitrides, molybdates, titanates, sulfates, etc. are used [17,18]. Among all these compounds, titanate-based phosphors have attracted much attention because of the wide bandgap and high refractive index of the titanate group due to which emission intensity increases. The chromaticity of titanate-based system can be controlled by adjusting some of the synthesis parameters and by doping them with a 3d or 4f system [19]. Therefore, Ba_3MoTiO_8 (now onwards abbreviated as BMT) phosphor can act as an efficient host as it is a ternary compound and not so much explored. On the other hand, transition metal Mo present in this lattice has a broad excitation band in the UV region which affects the luminescence behavior of the BMT phosphor [20]. RE ions are used as an excellent dopant in host lattices due to their high luminescence efficacy, high solubility, and large stokes shifts [21–25]. In addition, these ions are easily get excited by absorbing UV or n-UV radiations and give emission in the visible region through down-shifting process [17,26]. Out of the various RE ions, the europium ion is one of the most promising lanthanide ions and gives red emission due to its prominent $^5D_0 \rightarrow {}^7F_2$ transition [18,27–29]. Europium ion exhibits a narrow band and has higher lifetimes as compared to other luminescent materials [5,11,30]. So, it is relevant to prepare Eu^{3+} doped BMT phosphor as an efficient red-emitting phosphor that can overcome the limitations of currently used host lattices.

In this context, an undoped BMT sample has been prepared by three different synthesis methods like SSM, SGM, and CMM to optimize the best synthesis method useful for showcasing relatively better luminescence efficiency. SSM is a commonly used synthesis method for phosphors but it has its own drawbacks such as the requirement of high sintering temperature, non-uniform mixing of precursors, vast size distribution, and ensuing grinding which out-turn in damaging of phosphor surface and reduction of luminescence intensity [15]. So, it is obligatory to focus on other chemical methods like combustion, Sol-gel, hydrothermal, wet chemical route, etc. to curb these limitations. The sol-gel process generally involves the formation of organic/inorganic network through a chemical reaction at low temperature, the transition of this solution to colloidal sol, and then the transition of this colloidal sol to gel form [31]. Pechini-type sol-gel, also known as polymerizable-complex technique, is a vastly used technique for metal-complex formation with the addition of citric acid (CA) and polyethylene glycol (PEG) [32,33]. The combustion method is an unadorned, safe, and energy frugality method. In this technique, a chemical reaction takes place at a relatively low temperature [34].

The present work includes the synthesis of BMT:Eu^{3+} phosphor for better luminescence property using three different techniques (SSM, SGM, and CMM) and characterization of the as-prepared phosphor using X-ray diffraction (XRD), Raman, SEM to check the crystallinity, morphology and phase formation. Raman spectra and SEM studies have been compared for different samples synthesized by different techniques to optimize synthesis techniques for better morphology. Further DRS, PL, and PL decay spectral analysis have been conducted on the as-prepared phosphors. The CIE coordinates have also been evaluated using the PL spectral information.

2. Experimental

2.1. Sample preparation

2.1.1. Solid-state reaction (SSM) method

Barium carbonate ($BaCO_3$), Molybdenum trioxide (MoO_3), Titanium oxide (TiO_2), and Europium oxide (Eu_2O_3) were used as a starting material to prepare a series of BMT phosphors by doping them with europium ions using the SSM method. The stoichiometric amount of these chemicals was ground in an agate mortar by using acetone as a dispersive medium for 1 hour. The mixture collected in an alumina crucible was sintered at 1100°C for 10 hours in a high-temperature programmable furnace. After that, the sample was allowed to cool naturally down to room temperature. The obtained mixture was grounded to make fine powder for further characterization. Fig. 1(a) shows all the steps involved in preparing the titled phosphor using the SSM technique.

2.1.2. Combustion Method (CMM)

High purity barium nitrate ($BaNO_3)_2$, Titanium (IV) Oxyacetylacetonate ($C_{10}H_{14}O_5Ti$), Ammonium Molybdate ($(NH_4)_6Mo_7O_{24}$*$4H_2O$), and Europium Oxide (Eu_2O_3) were used as precursors for this CMM technique. Glycine (AR Grade) has been employed as fuel to prepare BMT: Eu^{3+} phosphor in this technique. The stoichiometric amount of barium nitrate and Ammonium Molybdate were dissolved in DI water individually and $C_{10}H_{14}O_5Ti$ was dissolved in methanol at 55 °C, their mixed solution was referred to as Sol. X. Subsequently, solution of Eu_2O_3 was made by adding HNO_3 (referred to as Sol. Y). After that Sol. X and Sol. Y were mixed together and obtained solution was stirred at 100 °C for 1 hour (Sol. Z). After continuous stirring for 1 hour, an appropriate amount of glycine was added to it. Afterwards, the solution was placed again on stirring for 2 hours at 120 °C to obtain a homogeneous mixture and for dehydration. When dehydration took place, the temperature of the hot plate was raised to 500 °C. At this temperature, a grey colored foamy powder was formed with the release of a large amount of gases. Finally, this foamy powder was crushed using an agate-mortar, and the obtained powder was sintered at 1100 °C to get a pure phase.

2.1.3. Sol- Gel - Combustion (SGM) Method

Precursors used for this synthesis method were the same as used in the CMM technique and a few steps are also the same up to the preparation of Sol. Z (as described in the combustion process). After that appropriate amount of citric acid was added into the Sol. Z. Citric acid is used as a chelating agent in this reaction and the amount of citric acid was decided by keeping the metal and citric acid (CA) ratio as 2:1. PEG was used dropwise as a network former in this polymeric reaction and the pH of the solution was maintained for homogeneous gel formation. After that, the solution was kept at constant stirring for 4 hours at 120°C. This has resulted in the formation of gel and the gel was further fired at 500°C until a black foamy solid was formed. The foamy solid was sintered at 1000°C for 10h. Fig. 1(b) describes all the steps of CMM and SGM methods through a flowchart. Table 1 depicts sample coding for different concentrations of europium ion.

2.2. Characterization techniques employed

X-Ray Diffraction (XRD) was carried out using the Bruker D8 Advanced powder X-ray diffractometer via Cu-Kα radiation (1.54Å). For surface morphology, scanning electron microscope (SEM) recordings were done by using the instrument Zeiss EVO40. Diffuse reflectance spectroscopy (DRS) study was conducted by using the JASCO V–770 Spectrophotometer. To study the PL properties, JASCO-made Spectrofluorometer (FP-8300) has been employed. Raman Spectra was recorded with the help of Invia Renishaw Raman Spectroscope. To study the PL decay profiles, an Edinburgh FL920 Fluorescence spectrometer was used.

3. Results and Discussion

3.1. Structure and phase analysis

For structure and phase analysis, the XRD pattern of the BMT sample was examined for differently synthesizing methods. SSM synthesized sample was sintered at different temperatures to optimize the suitable temperature for this method (Fig. S1 in supporting information). But the

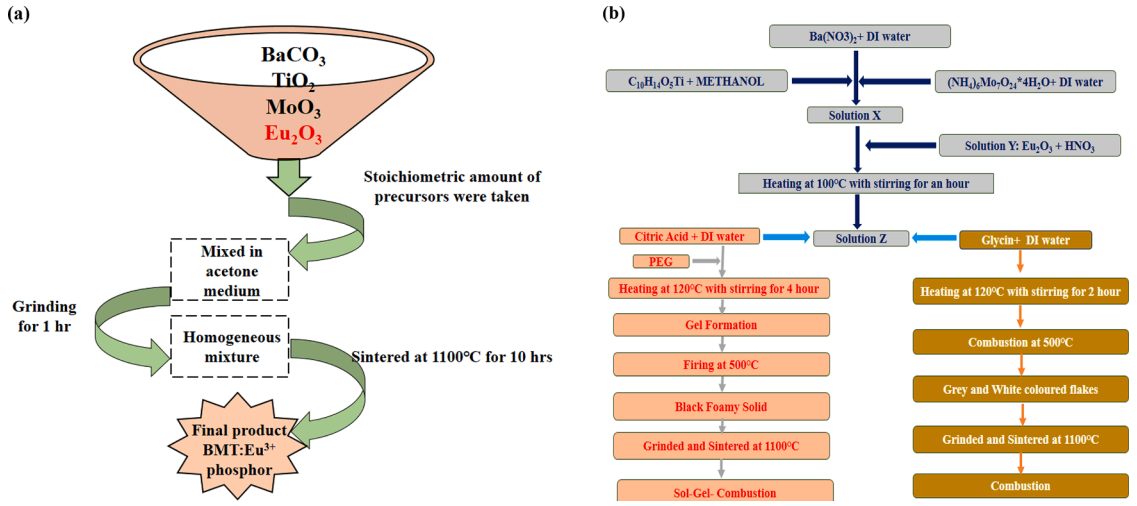


Fig. 1. (a). Flow chart for various steps involved in the synthesis process of BMT phosphor using SSM method.
Fig. 1(b). Flow chart for various steps involved in the synthesis process of BMT phosphor through CMM and SGM methods.

Table 1
Sample coding for different concentrations of Europium ion doped BMT phosphor

Synthesis Method	Eu^{3+} ion concentration (mol%)	Sample Code
Solid state Reaction (SSM)	0	SSM0
	1.0	SSM1
Sol-gel Combustion (SGM)	0	SGM0
	1.0	SGM1
Combustion (CMM)	0	CMM0
	1.0	CMM1
	2.0	CMM2
	3.0	CMM3
	4.0	CMM4

sample sintered at 1100°C has the single-phase confirmation with standard ICDD card no.00-038-0254 with a space group R-3m (166), and lattice parameters a= 5.9570Å, b=5.9570Å, and c=21.290Å. Fig. 2(a) manifests the XRD pattern of SSM0, SGM0 and CMM0 samples along with the standard JCPDS data. As shown in Fig. 2(a) the diffraction patterns of SSM0, SGM0, and CMM0 samples are perfectly matching with the standard JCPDS data (00-038-0254). Fig. 2(b) shows the XRD pattern of SSM1, SGM1, CMM1 samples with 1 mol% of europium ion. With the incorporation of europium ion, no mismatch of phase was found. This shows that Eu^{3+} ions (ionic radii 1.08 Å) are successfully replacing Ba^{2+} ions (ionic radii 1.4 Å) due to their proximity in ionic radii. In addition to this, their charges are also getting compensated ($3Ba^{2+}$ ions replaced for $2Eu^{3+}$ ions). The crystalline size was calculated by using the Debye-Scherer formula

$$D = 0.9\lambda/\beta\cos\theta$$

Where λ is the wavelength of X-Ray, θ is Bragg's diffraction angle and β is the full width at half maxima (FWHM). Here D is crystalline size. The crystalline size estimated using the Debye-Scherer formula was found to be 140 nm, 71nm, 65 nm for SSM1, SGM1and CMM1 samples respectively.

3.2. Morphological analysis from SEM

To elucidate the information pertaining to the morphology of the as-prepared phosphors, the SEM images are recorded for BMT phosphors prepared by three different techniques such as SSM, SGM, and CMM. Fig. 3 shows the SEM images of SSM0, SGM0, and CMM0 samples taken at different magnifications. Fig. 3(a(i)) shows the image of the SSM0 sample taken at 5.0 Kx magnification and it is conspicuous from the same figure that, particles are agglomerated and non-uniformly distributed. Fig. 3(a(ii)) & (a(iii)) show the images recorded for the same sample with a magnification of 10.0 Kx and 35.0 Kx respectively. Fig. 3(a(ii)) shows flowery morphology with some rod-like fibrous structure and such rod-like shape appears relatively clear in Fig. 3(a (iii)). Particle size for the SSM0 sample lies in the range of 5-7 μm. Fig. 3 (b(i)) and (c(i)) are the SEM images recorded for SGM0 and CMM0 samples respectively with 5.0 Kx magnification and illustrate that cluster formation is more for the CMM0 sample as compared to the SGM0. Due to more aggregation in CMM0 samples, they can absorb

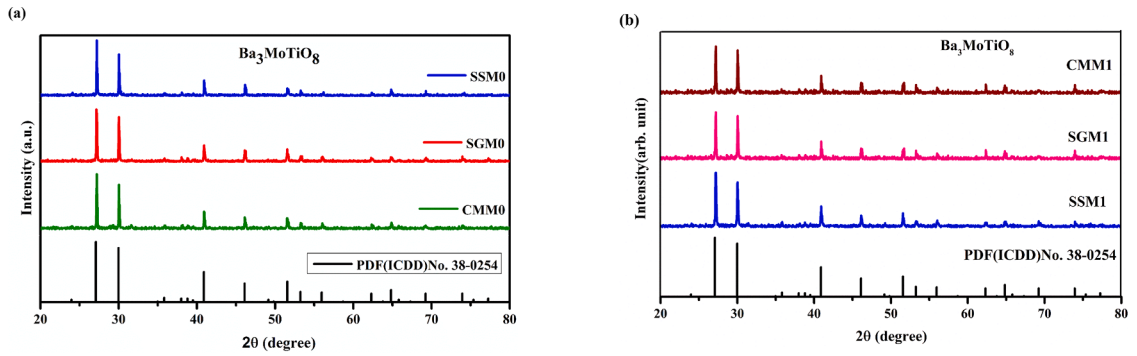


Fig. 2. Fig. 2(a). XRD pattern of BMT sample synthesized by SSM, SGM, and CMM technique along with standard diffraction pattern of BMT.
Fig. 2(b). XRD pattern of CMM1, SGM1, and SSM1 phosphor with 1 mol% concentration of Europium ion.

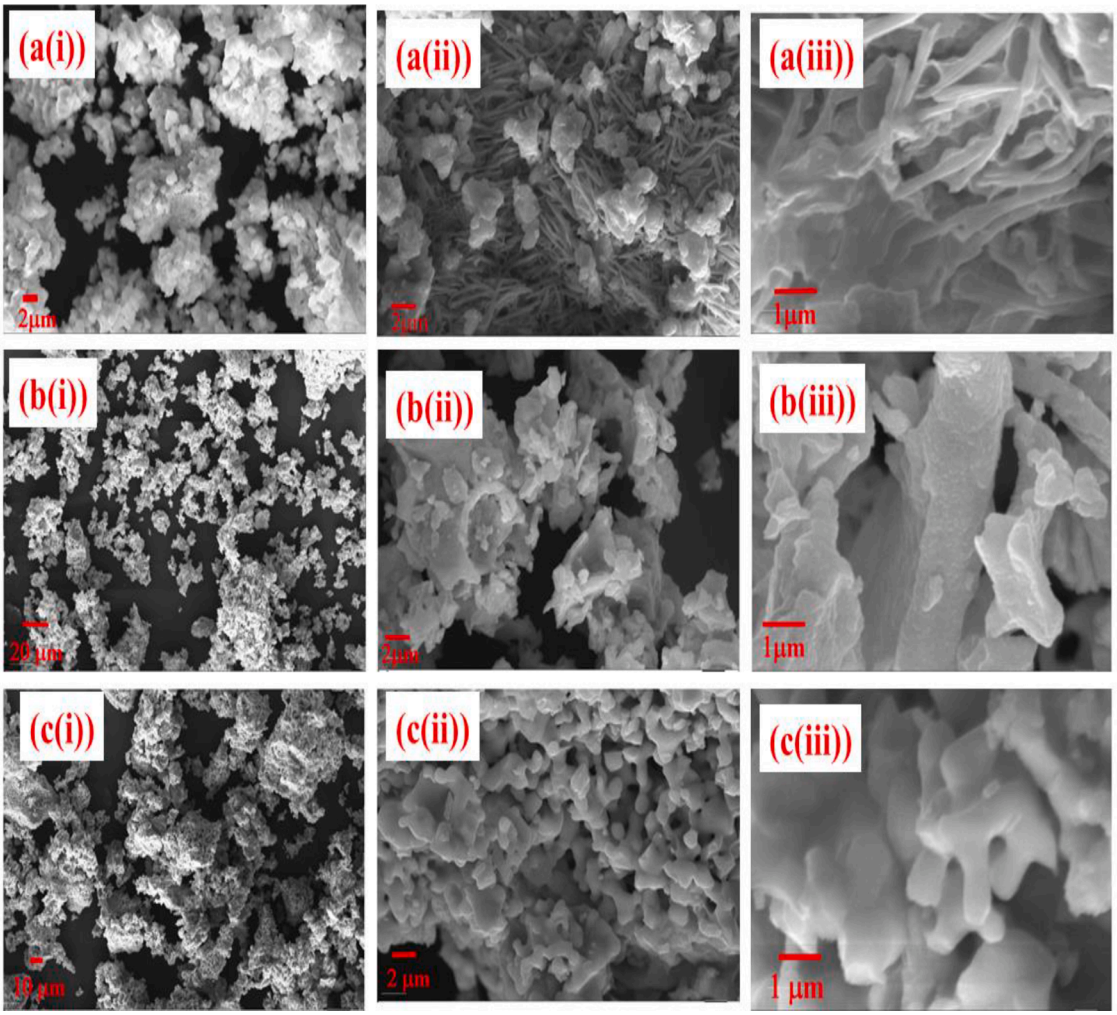


Fig. 3. SEM images recorded for BMT samples (a (i-iii)) SSM0 (b (i-iii)) SGM0 and (c (i-iii)) CMM0.

more excitation energy and are useful in showcasing relatively high luminescence [35]. Fig. 3(b(ii)) and (c(ii)) are the images recorded for SGM0 and CMM0 samples with 10.0 KX magnification and illustrate that the particles are non-uniform in size and shape in the SGM0 sample. In the case of the CMM0 sample, the particles attain dumbbell-like shape with a uniform size. Fig. 3(b(iii)) and (c(iii)) are the SEM images pertaining to SGM0 and CMM0 samples respectively with 35.0 Kx magnification. In the SEM image of the SGM0 sample, the particles are appearing with some elongation in a particular direction. On the other hand, the SEM image of the CMM0 sample shows the particles that are distributed uniformly with relatively better morphology and are useful in enhancing their luminescence performance. This difference in morphology may be due to the difference in reaction essence with varying the synthesis method [36]. The aforementioned SEM analysis performed for the BMT sample synthesized by using different techniques allows us to contemplate that, the morphology of the particles got improved gradually from the SSM to the CMM synthesis route via the SGM method. The particle size for SGM0 and CMM0 samples lie in the range of 3-4μm and 1-2μm respectively. Fig. 4 shows the SEM images and EDAX spectra for doped samples. Fig. 4(a.1), 4(b.1), and 4(c.1) show the morphology of SSM1, SGM1, and CMM1 phosphors and Fig. 4(a.2), 4 (b.2), and 4(c.2) manifest all the constituent elements present in the phosphor lattice [37,38]. The particle size for all the BMT samples prepared by using three different techniques are in micrometer range and quite suitable for coating in w-LEDs and other display devices [8, 39].

3.3. Raman spectral analysis

In the process of knowing the information pertaining to different vibrational modes present in the as-synthesized BMT samples, the Raman spectra were recorded for BMT phosphors synthesized by different techniques under 488 nm excitation. Fig. 5 manifests the different Raman vibrational peaks observed for the BMT sample prepared by different methods at wavenumbers 326, 526, 792, 841, 892, and 1042 cm^{-1}. The band at 326 cm^{-1} describes the bending vibration modes of Mo-O bonds and the band observed at 526 cm^{-1} (between the range 450-600 cm^{-1}) belongs to the bond vibrations between Ti and oxygen atoms [40,41]. Vibrational bands present at 792 and 841 cm^{-1} are in good agreement with anti-symmetric stretching bonds with tetrahedral O-Mo-O molecules [42,43]. The most intense peak that was observed at 892 cm^{-1} corresponds to the phonon energy of the as-prepared compound. It is well known that a host lattice with relatively less phonon energy gives emission more in radiative form than non-radiative. The phonon energy observed for this lattice is relatively low when we compared it with the phosphors reported earlier [40,44, 45]. Fig. 5 represents the comparison of Raman spectra observed for SSM0, SGM0, and CMM0 samples. Among all the three samples prepared by using three different methods, a sample prepared by the CMM method show relatively intense PL when compared with the other methods. As the particle size decreases from solid-state to combustion method, shifting in Raman peaks are observed towards higher wavenumber side (as shown in the inset of Fig. 5) and intensity of the peaks

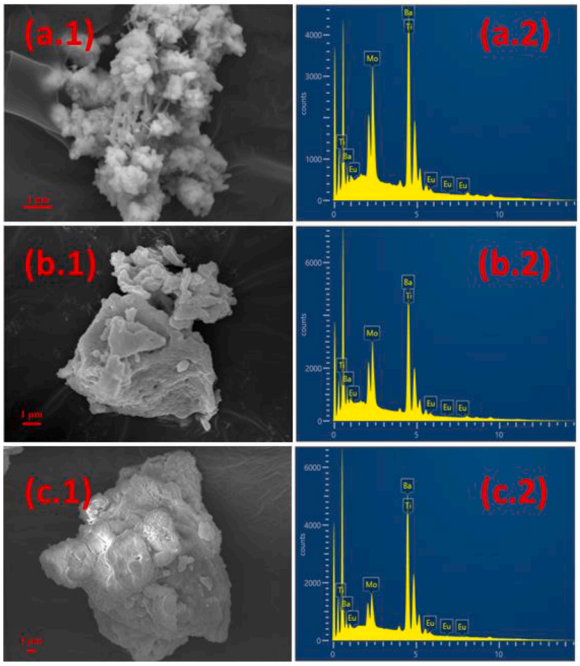


Fig. 4. SEM images for doped samples (a.1) SSM1 (b.1) SGM1 (c.1) CMM1 and EDAX spectra for doped BMT phosphor (a.2) SSM1 (b.2) SGM1 (c.2) CMM1

decreased gradually [46].

3.4. Diffuse Reflectance Spectra (DRS) analysis for bandgap measurement

Fig. 6 displays the DRS of CMM0 and CMM2 phosphor recorded in the range 200-800 nm. To find out the optical band gap the DRS spectral data was converted into the Kubelka–Munk function using the following equation [47]:

$$F(R) = (1 - -R)^2/2R = \alpha/S$$

In this equation, R is Reflectance, α is absorption coefficient and S is the Scattering coefficient. The tauc equation gives the relation between energy bandgap (E_g) and absorption coefficient (α) as shown below:

$$\alpha h\upsilon = C\left(h\upsilon - -E_g\right)^n$$

Here hυ belongs to photon energy, C is energy independent constant and n can take values depending upon the type of transition which is 1/2 and 2 for direct and indirect allowed transitions. Since F(R) is directly proportional to α, the above equation can be rewritten as:

$$F(R)h\upsilon = C\left(h\upsilon - -E_g\right)^n$$

To calculate the bandgap of the aforementioned mentioned phosphors (CMM0 and CMM2), a graph has been plotted between the $[F(R)h\upsilon]^{1/2}$ and hυ as shown in the inset of Fig. 5 [48,49]. The direct bandgap for CMM0 and CMM2 was found to be 3.60 and 3.46 eV respectively by extrapolating the slope as shown in the inset of Fig. 6

3.5. PL excitation and emission spectral studies

To study the luminescent properties of europium doped BMT phosphors, PL spectral features were investigated at room temperature. PL excitation spectra of BMT phosphors with 1 mol% of Eu^{3+} ions (prepared by different synthesizing techniques) are shown in Fig. 7 The spectra were recorded in the wavelength range 260-500 nm. The spectra comprise of one broad charge transfer band (CTB) between the range 260-290 nm and several peaks in the region between 350-500 nm at wavelengths 360, 381, 393, 414, 465 nm, and the transitions corresponding to these peaks are $^7F_0 \rightarrow {}^5D_4$, $^7F_0 \rightarrow {}^5L_7$, $^7F_0 \rightarrow {}^5L_6$, $^7F_0 \rightarrow {}^5D_3$, $^7F_0 \rightarrow {}^5D_2$ respectively. CTB is in agreement due to the transition from the 2p orbital of O^{2-} to the empty 4f shell of the Eu^{3+} ion [18]. Out of these peaks, two main peaks of comparatively high intensity are present at wavelengths 393nm and 465nm and were used as an excitation wavelength to record the PL emission spectra. The PL emission spectra have been recorded for 1 mol% Eu^{3+} doped samples (SSM1, SGM1, CMM1) under 393 and 465 nm excitation wavelengths. For comparison purposes, the PL emission spectra recorded under the two intense excitation wavelengths are shown in Fig. 8(a) (under 393 nm excitation) and Fig. 8

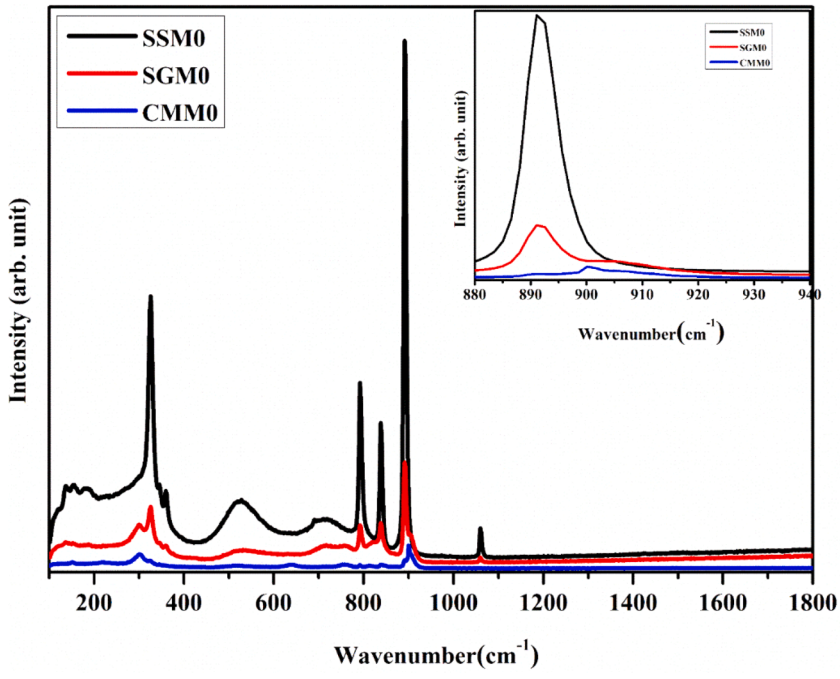


Fig. 5. Raman spectra of SSM0, SGM0, and CMM0 sample.

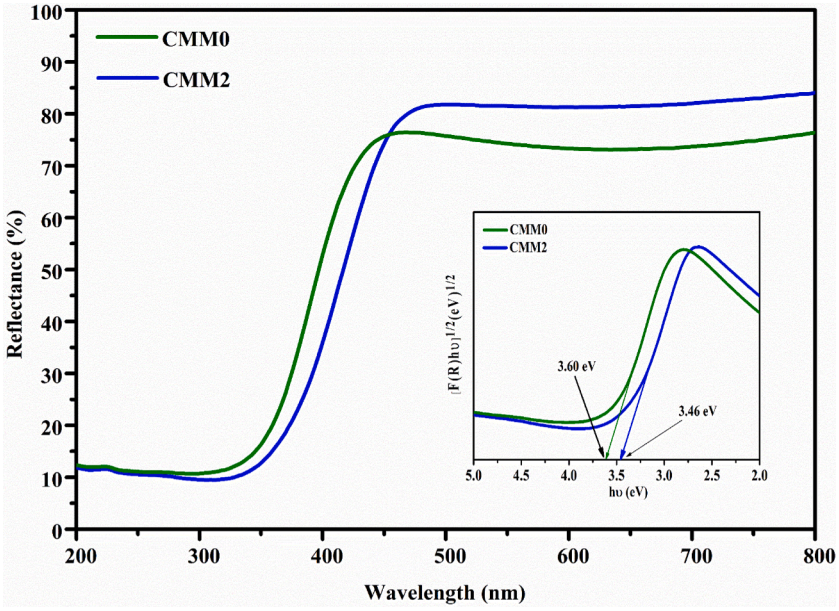


Fig. 6. Diffuse Reflectance Spectrum of CMM0 and CMM2 sample. The inset image shows the assesthe sment of the direct bandgap using Kubelka- Munk plot.

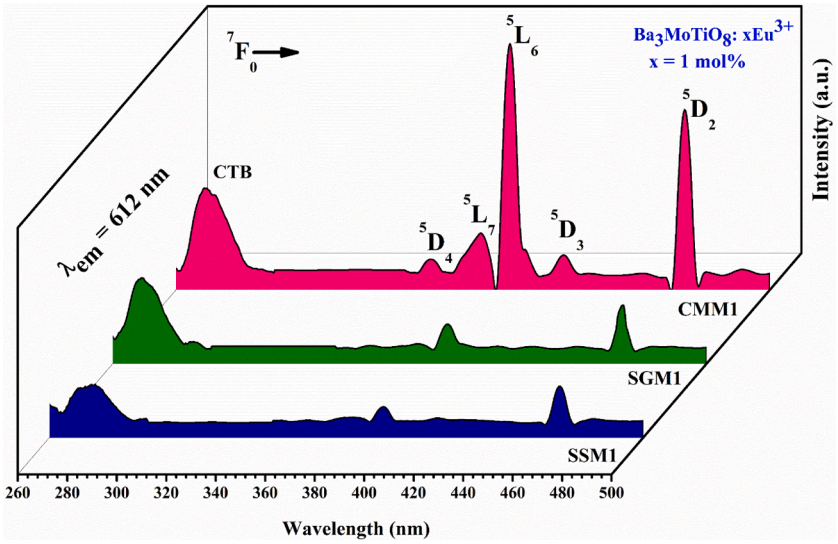


Fig. 7. Excitation spectra of SSM1, SGM1, and CMM1 phosphor with 1 mol% of Eu^{3+} ions at 612 nm emission wavelength

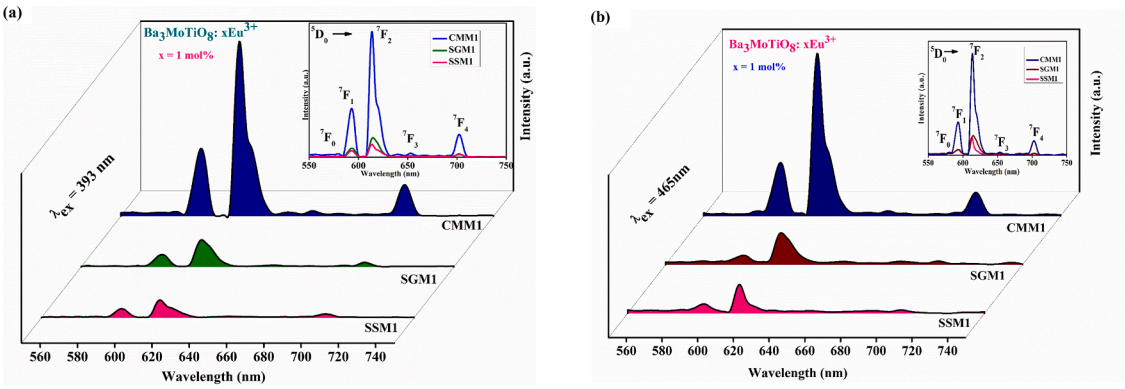


Fig. 8. Fig. 8(a). Emission spectra of SSM1, SGM1, and CMM1 phosphor with 1 mol% Eu^{3+} ion at 393 nm excitation wavelength.
Fig. 8(b). Emission spectra of SSM1, SGM1, and CMM1 phosphor with 1 mol% Eu^{3+} ion at 465 nm excitation wavelength.

(b) (under 465 nm excitation) between the region 550-750 nm. These two PL emission spectra comprise the same peaks at wavelengths 582, 593, 612, 658, and 704 nm accompanying the transitions $^5D_0 \rightarrow {}^7F_0$, $^5D_0 \rightarrow {}^7F_1$, $^5D_0 \rightarrow {}^7F_2$, $^5D_0 \rightarrow {}^7F_3$, and $^5D_0 \rightarrow {}^7F_4$ respectively [25,50,51]. A comprehensive analysis of both the figures (Fig. 8(a) & 8(b)) allows us to contemplate that, the BMT phosphor prepared by employing combustion technique is yielding good luminescence when compared with the remaining two techniques. Relatively high luminescence intensity in the case of the CMM1 sample may be due to the improved morphology and smaller size particles that can be seen from the SEM images. Particles with relatively smaller size can minimize scattering and yields relatively intense luminescence [52]. Low luminescence intensity in the case of SSM synthesized sample may be due to high agglomeration in particles due to which luminescence intensity decreased. So, on that behalf, it is contemplated that, CMM technique is the optimized one among the three techniques used to prepare BMT: Eu^{3+} phosphor. Fig. S2 in supporting information shows the comparative emission spectra recorded for the CMM1 sample at different excitation wavelengths 393nm and 465nm. The spectra recorded at 393 nm showed a little higher intensity as compared to 465 nm. Intense luminescence possessed by BMT: xEu^{3+} phosphor under 393 nm excitation wavelength promotes its usage in visible red photonic device applications under n-UV LED/blue LED excitation sources.

Relatively higher luminescence exhibited by BMT phosphor prepared by using the CMM technique prompted us to prepare a series of europium ion-doped BMT: xEu^{3+} phosphors (x= 1, 2, 3, 4 mol%) using the CMM technique for the optimization of the doped RE ion concentration for better luminescence. Emission spectra were recorded for this CMM synthesized series to check the optimized concentration of activator ion. Spectra were recorded between the wavelength range 550-750 nm under 393 nm and 465nm excitation wavelengths. All the peak positions and transitions are the same as mentioned above and there is no shifting in the peaks by changing the excitation wavelength. For both the excitation wavelengths, the highest luminescence intensity was observed for x = 2 mol% concentration of europium ion. Beyond 2 mol% of europium ion concentration, the PL intensity decreases for both the cases as shown in Fig. 9(a) and Fig. 9(b). This may be due to concentration quenching produced by a non-radiative energy transfer that took place between the activator ions. Therefore, it is concluded that 2 mol% of europium ion concentration is optimum in BMT phosphor to showcase relatively intense luminescence. An energy level diagram has been made to show various transitions of excitation and emission spectra and is shown in Fig. 10. For the non-radiative relaxation, three mechanisms are responsible like radiative re-absorption, exchange interaction, and electric multipolar interaction between the activator ions. Critical distance R_c is the average shortest distance on which energy transfer depends between activator ions. To find out the critical distance (distance between two neighboring ions) the following formula is used [53].

$$R_C = 2[3V/(4PX_CN)]^{1/3}$$

Where X_C is mol concentration of activator ion at which quenching took place. Here X_C=0.02 (2 mol%), V is the volume of a unit cell and N is the number of cations per unit cell. For this case- V=654.28Å^3, N=3, and Xc = 0.02. The average critical distance calculated by using these values is 27.51Å which is very large when compared to 5Å designated for exchange interaction. This allows us to contemplate that, exchange interaction is not responsible for energy transfer. This made us apply Dexter's theory further to check the type of multipolar interaction responsible for energy transfer.

The relation between emission intensities and doping concentration is given as [48,54]-

$$I/x = K\left[1+\beta(x)^{Q/3}\right]^{-1}$$

Here x is the mole concentration of europium ion, Q determines the type of interaction (Q=6, 8, and 10 are corresponding to electric dipole-dipole interaction, dipole-quadrupole interaction, and quadrupole - quadrupole interaction respectively), I is the luminous intensity, K and β are constants. According to Huang's theory, the relation between luminous intensity and mole fraction of activator ions is given as [55]-

$$log\frac{I}{x} = A - \frac{Q log x}{3}$$

where A is independent of doping concentration. Fig. 11(a) and 11(b) show the graph between log (I/x) and log(x) using Dexter theory under 393 nm and 465 nm excitation wavelength. The Q value estimated from the slope of the graph was found out 4.617 and 4.253 which are close to 6 indicating electric dipole-dipole interaction as the responsible mechanism for energy transfer and the subsequently observed concentration quenching effect.

3.6. Colorimetric analysis

Commission Internationale de I'Eclairage was used to find out the colorimetric performance of BMT: Eu^{3+} phosphor by using the emission spectral data recorded under 393 nm and 465 nm excitation wavelength. The CIE coordinates calculated for CMM synthesized BMT: xEu^{3+} phosphor (x= 1, 2, 3, 4 mol%) were depicted in Table 2. The CIE coordinates of CMM synthesized BMT: xEu^{3+} phosphor is falling in the pure red region as shown in Fig. 12(a) and 12(b). The correlated color temperature (CCT) was calculated using the McCamy empirical theoretical equation [7,56]:

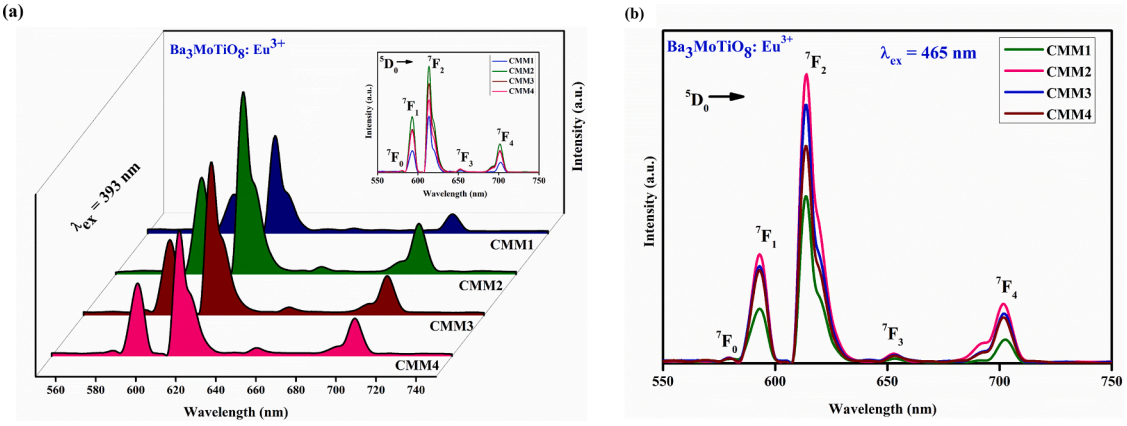


Fig. 9. (a). Emission spectra of CMM synthesized BMT: xEu^{3+} (x=1, 2, 3 and 4 mol%) phosphor under 393 nm excitation wavelength. Fig. 9(b). Emission spectra of CMM synthesized BMT: xEu^{3+} (x=1, 2, 3 and 4 mol%) phosphor under 465 nm excitation wavelength.

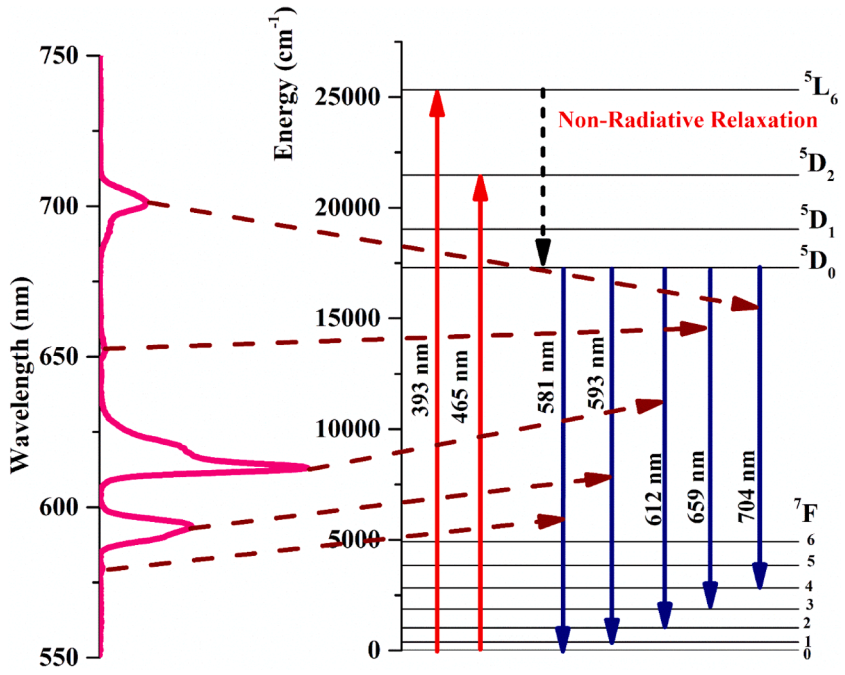


Fig. 10. Energy level diagram for various transitions in europium ion.

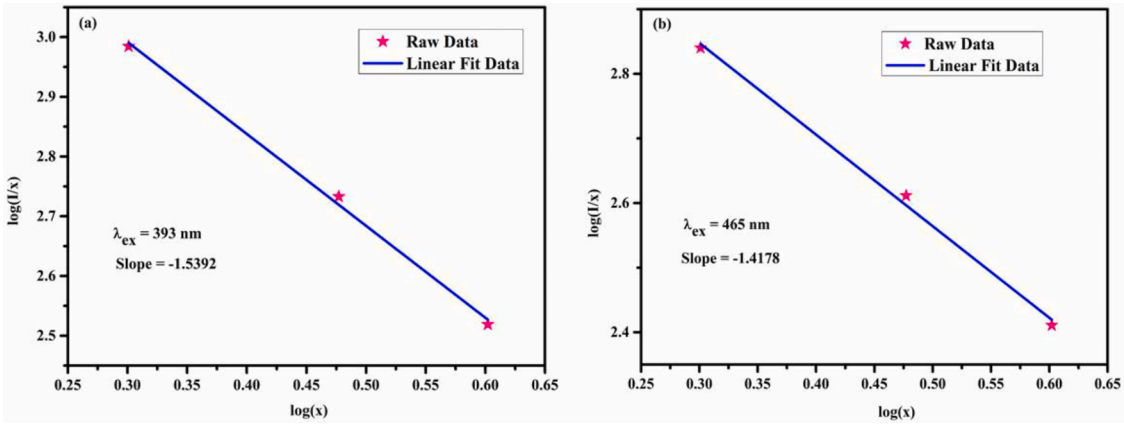


Fig. 11. Dexter plot of log (I/x) vs. log (x), (a) under 393 nm (b) under 465 nm, excitation wavelengths.

Table 2
CIE Coordinates for BMT phosphors prepared by CMM technique

Sample	CIE Co-ordinates (x, y) under 393 nm excitation	CIE Co-ordinates (x, y) under 465 nm excitation
CMM1	(0.6097, 0.3895)	(0.6454, 0.3542)
CMM2	(0.6182, 0.3811)	(0.6501, 0.3495)
CMM3	(0.6178, 0.3816)	(0.6455, 0.3540)
CMM4	(0.6146, 0.3847)	(0.6438, 0.3558)

$$CCT = -473n^3 + 3601n^2 - -6861n + 5514.31$$

Where $n = (x\text{-}x_e)/(y\text{-}y_e)$ gives the inverse slope value and x_e (0.3320) and y_e (0.1858) are the epicenter coordinates. The CCT values determined for BMT phosphor were 1756.17, 1827.94, 1822.68, 1791.86 K for the PL spectra recorded under 393 nm excitation wavelength and for 465 nm excitation wavelength were 2410.72, 2582.46, 2415.68, and 2357.34 K corresponding to 1.0, 2.0, 3.0, and 4.0 mol% of Eu^{3+} ion concentration, respectively. This allows us to authenticate the superior nature of BMT: xEu^{3+} phosphor for intense red emission required for the fabricating w-LEDs.

3.7. PL decay analysis

PL decay measurements were carried on combustion synthesized BMT: xEu^{3+} phosphor (x= 1, 2, 3, 4 mol%) for 612 nm red emission under 393 nm excitation and is shown in Fig. 13. From Fig. 13 it is conspicuous that, all the decay curves recorded are showing non-linear bi-exponential behavior and obeys the following by-exponential equation [25,57]:

$$I = A_1 \exp(-t/t_1) + A_2 \exp(-t/t_2)$$

Where "I" is the PL intensity at time t, A_1 and A_2 are the constant parameters, τ_1 and τ_2 are the fast and slow lifetimes for bi-exponentially fitted data. On substitution of these values the following equation gives the average experimental lifetime:

$$t = \left(A_1 t_1^2 + A_2 t_2^2\right)/\left(A_1 t_1 + A_2 t_2\right)$$

The average experimental lifetime values are found to be 1.89, 1.77,

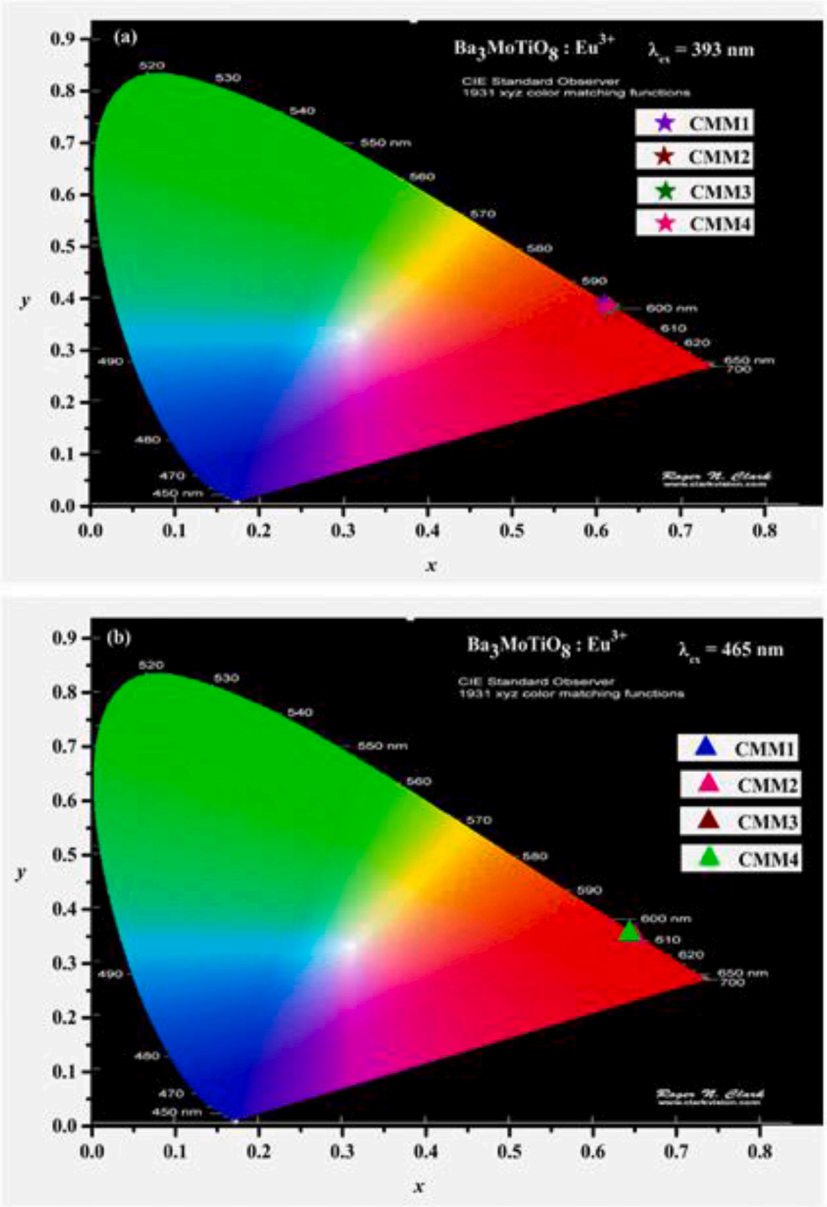


Fig. 12. CIE chromaticity diagram of BMT:xEu^{3+} (x= 1, 2, 3, 4 mol%) phosphor synthesized by CMM method, under the excitation wavelength of (a) 393 nm (b) 465 nm.

1.55, and 1.35 ms, for CMM1, CMM2, CMM3, and CMM4 phosphors respectively [58]. These results showed that, with an increase in the concentration of europium, experimental decay times are decreasing. This happened because of an increase in non-radiative recombination resulting in the energy transfer between the activator ions itself [58–60].

4. Conclusions

BMT: xEu^{3+} phosphor was successfully synthesized by three different methods- Solid-state, Sol-gel, and Combustion method. A comparative Raman Analysis was also carried out for undoped samples of three techniques which shows that by varying the synthesis route, peaks position shifts toward higher wavenumber side and this happens because of the decrease in particle size when we go from SSM technique to SGM and then to CMM. The XRD pattern confirmed the phase and crystallinity of all the samples. A comparison of morphology was done by SEM micrographs. In the case of combustion synthesized samples, the morphology of particles gets improved, and this improved morphology enhances the luminescence intensity as evident from the PL spectra. A comparative PL spectrum recorded for 1 mol% europium ion-doped BMT phosphor prepared by using SSM, SGM, and CMM techniques demonstrates the superiority of the CMM technique in showcasing relatively intense red emission. This prompted to have further investigations on a BMT phosphor prepared by the CMM technique with different concentrations of europium. All the BMT phosphors prepared by CMM techniques are displaying red emission under 393 and 464nm as excitation wavelengths. This has been further confirmed by the CIE chromaticity diagram also. And lastly, time decay analysis was carried out which shows that 2 mol% is the optimum concentration of uuropium ion for combustion synthesized series and after that energy, transfer takes place between the activator ions itself through dipole-dipole interaction which can be confirmed from the Dexter theory. CIE points calculated for combustion synthesized samples lie in the pure red region. So above results indicate that the combustion method synthesized BMT phosphor can act as an efficient red phosphor in w-LEDs applications under n-UV or Blue radiation excitation.

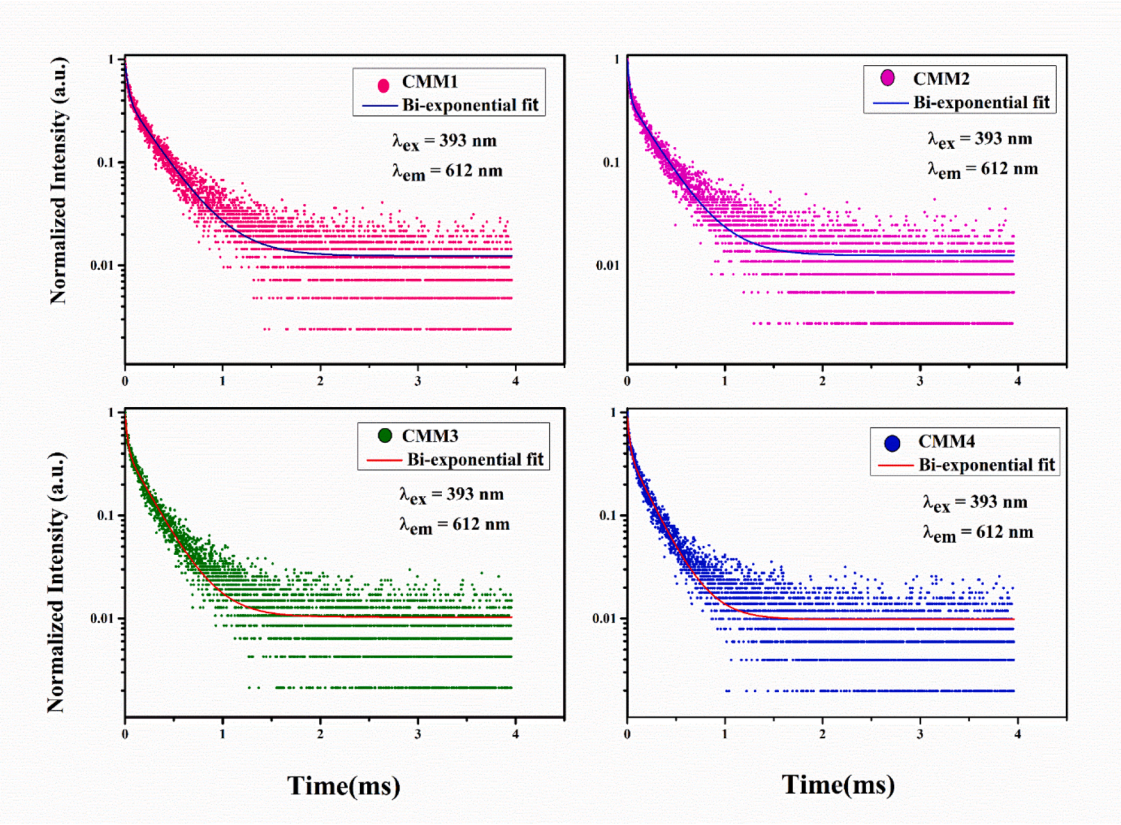


Fig. 13. Decay profile of Eu^{3+} ions doped BMT phosphors synthesized by CMM route.

Credit author statement

Ms. Pooja Rohilla: work plan, experiments, calculation, and writing the manuscript.

Prof. A.S. Rao: work plan, supervision, validation, correction of the manuscript.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

The authors are very much thankful to the Delhi Technological University for the encouragement to carry out the present work. The work reported in this paper has been partially supported by a Major Research Project (EMR/2016/007766) sanctioned to Prof. A.S. Rao by SERB, Department of Science and Technology, Govt. of India, New Delhi. We want to acknowledge the Department of Chemistry, Dayalbagh Educational Institute, Deemed University, Agra for making the SEM facility available to us.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.materresbull.2022.111753.

References

[1] M. Janulevicius, P. Marmokas, M. Misevicius, J. Grigorjevaite, L. Mikoliunaite, S. Sakirzanovas, A. Katelnikovas, Luminescence and luminescence quenching of highly efficient $Y_2Mo_4O_{15}:Eu^{3+}$ phosphors and ceramics, Sci. Rep. 6 (2016) 1–12, https://doi.org/10.1038/srep26098.

[2] S. Ye, F. Xiao, Y.X. Pan, Y.Y. Ma, Q.Y. Zhang, Phosphors in phosphor-converted white light-emitting diodes: Recent advances in materials, techniques and properties, Mater. Sci. Eng. R Reports. 71 (2010) 1–34, https://doi.org/10.1016/j.mser.2010.07.001.

[3] X. Li, J.D. Budai, F. Liu, J.Y. Howe, J. Zhang, X.J. Wang, Z. Gu, C. Sun, R.S. Meltzer, New yellow $Ba_{0.93}Eu_{0.07}Al_2O_4$ phosphor for warm-white light-emitting diodes through single-emitting-center conversion, Light Sci. Appl. 2 (2013) 1–8, https://doi.org/10.1038/lsa.2013.6.

[4] Z. Xia, Z. Xu, M. Chen, Q. Liu, Recent developments in the new inorganic solid-state LED phosphors, Dalt. Trans. 45 (2016) 11214–11232, https://doi.org/10.1039/c6dt01230b.

[5] G.H. Li, N. Yang, J.G. Guo, Z.L. Wang, G.M. Cai, X.J. Wang, Efficient and stable $Sr_3Eu_2B_4O_{12}$ red phosphor benefiting from low symmetry and distorted local environment, Dalt. Trans. 49 (2020) 3260–3271, https://doi.org/10.1039/c9dt04769g.

[6] L. Zhao, Y. Liu, C. Zhai, F. Liao, Y. Gao, Photoluminescence properties of Tb-doped and (Zn,Tb) co-doped barium strontium titanate crystalline powders, J. Alloys Compd. 694 (2017) 721–725, https://doi.org/10.1016/j.jallcom.2016.09.332.

[7] G.L. Bhagyalekshmi, D.N. Rajendran, Luminescence dynamics of Eu^{3+} activated and co-activated defect spinel zinc titanate nanophosphor for applications in WLEDs, J. Alloys Compd. 850 (2021), 156660, https://doi.org/10.1016/j.jallcom.2020.156660.

[8] Y. Wang, C. Yu, Y. Zhou, E. Song, H. Ming, Q. Zhang, Mn^{4+} doped narrowband red phosphors with short fluorescence lifetime and high color stability for fast-response backlight display application, J. Alloys Compd. 855 (2021), 157347, https://doi.org/10.1016/j.jallcom.2020.157347.

[9] P. Singh, R.S. Yadav, P. Singh, S.B. Rai, Upconversion and downshifting emissions of Ho^{3+}-Yb^{3+} co-doped $ATiO_3$ perovskite phosphors with temperature sensing properties in Ho^{3+}-Yb^{3+} co-doped $BaTiO_3$ phosphor, J. Alloys Compd. 855 (2021), 157452, https://doi.org/10.1016/j.jallcom.2020.157452.

[10] G.-H. Li, P.-F. Wu, B. Ye, G.-M. Cai, Enhancement of Eu^{2+} photoluminescence behavior in $NaBaB_9O_{15}$ based on the K^+ doping, J. Lumin. 243 (2022), 118613, https://doi.org/10.1016/j.jlumin.2021.118613.

[11] Y. Zhang, J. Xu, Q. Cui, B. Yang, Eu^{3+} -doped $Bi_4Si_3O_{12}$ red phosphor for solid state lighting : microwave synthesis, characterization, photoluminescence properties and thermal quenching mechanisms, Nat. Publ. Gr. 7 (2017) 1–12, https://doi.org/10.1038/srep42464.

[12] C. Guo, T. Chen, L. Luan, H.K. Yang, B.C. Choi, J.H. Jeong, Synthesis and characterization of the intense red phosphor $AgEu(MoO_4)_2$ for blue GaN-based LED chips, J. Korean Phys. Soc. 54 (2009) 720–724, https://doi.org/10.3938/jkps.54.720.

[13] L. Zhang, B. Sun, C. Shao, F. Zhen, S. Wei, W. Bu, Q. Yao, Z. Jiang, H. Chen, Preparation, band-structure and luminescence of double perovskite Ba_2MgMoO_6: Eu^{3+} orange-red phosphor for white LEDs, Ceram. Int. 44 (2018) 17305–17312, https://doi.org/10.1016/j.ceramint.2018.06.192.

[14] Y. Li, X. Liu, Sol-gel synthesis, structure and luminescence properties of Ba_2ZnMoO_6:Eu^{3+} phosphors, Mater. Res. Bull. 64 (2015) 88–92, https://doi.org/10.1016/j.materresbull.2014.12.041.

[15] C. Guo, H.K. Yang, Z. Fu, L. Li, B.C. Choi, J.H. Jeong, A potential red-emitting phosphor $BaGd_2(MoO_4)_4$:Eu^{3+} for near-UV white LED, J. Am. Ceram. Soc. 92 (2009) 1713–1718, https://doi.org/10.1111/j.1551-2916.2009.03076.x.

[16] J. Tang, J. Si, X. Fan, Y. Liu, G. Li, G. Cai, Tunable emission, energy transfer and thermal stability of Ce^{3+}, Tb^{3+} co-doped $Na_2BaCa(PO4)_2$ phosphors, J. Rare Earths. (2021), https://doi.org/10.1016/j.jre.2021.05.016.

[17] G.B. Nair, H.C. Swart, S.J. Dhoble, A review on the advancements in phosphor-converted light emitting diodes (pc-LEDs): Phosphor synthesis, device fabrication and characterization, Prog. Mater. Sci. 109 (2020), 100622, https://doi.org/10.1016/j.pmatsci.2019.100622.

[18] Y. Liu, G. Li, R. Wang, G. Cai, Dependence of luminous performance on Eu^{3+}site occupation in $SrIn_2(P_2O_7)_2$: the effect of the local environment, Inorg. Chem. 60 (2021) 17219–17229, https://doi.org/10.1021/acs.inorgchem.1c02550.

[19] M. Mohapatra, Y.P. Naik, V. Natarajan, T.K. Seshagiri, Z. Singh, S.V Godbole, Rare earth doped lithium titanate (Li_2TiO_3) for potential phosphor applications, J. Lumin. 130 (2010) 2402–2406, https://doi.org/10.1016/j.jlumin.2010.08.001.

[20] A. Verma, S.K. Sharma, Rare-earth doped/codoped $CaMoO_4$ phosphors: a candidate for solar spectrum conversion, Solid State Sci 96 (2019) 105945, https://doi.org/10.1016/j.solidstatesciences.2019.105945.

[21] K.N. Kumar, L. Vijayalakshmi, J. Su, Enhanced red luminescence quantum yield from Gd^{3+} /Eu^{3+} : $CaLa_2ZnO_5$ phosphor spheres for photonic applications, Mater. Res. Bull. 103 (2018) 234–241, https://doi.org/10.1016/j.materresbull.2018.03.045.

[22] V. Balaram, Geoscience frontiers rare earth elements : a review of applications, occurrence, exploration, analysis, recycling, and environmental impact, Geosci. Front. 10 (2019) 1285–1303, https://doi.org/10.1016/j.gsf.2018.12.005.

[23] D. Chiriu, L. Stagi, C.M. Carbonaro, P.C. Ricci, pss to replace critical raw materials, 997 (2016) 989–997. https://doi.org/10.1002/pssc.201600116.

[24] Y. Tian, Development of phosphors with high thermal stability and efficiency for phosphor-converted LEDs, (2014) 1–15.

[25] L. Su, Y. Liu, G. Li, Z. Wu, G. Cai, L. Ma, X. Wang, Multicolor emission leading by energy transfer between Dy^{3+} and Eu^{3+} in wolframite InNbTiO6, J. Lumin. 227 (2020) 117578, https://doi.org/10.1016/j.jlumin.2020.117578.

[26] D. Halmurat, T. Yusufu, Q. ling Wang, J. He, A. Sidike, Rare earth ion Tb^{3+} doped natural sodium feldspar (NaAlSi3O8) Luminescent properties and energy transfer, Sci. Rep. 9 (2019) 1–8, https://doi.org/10.1038/s41598-019-51171-3.

[27] P. Du, J.S. Yu, red-emitting phosphors with ultrabroad excitation band for white light-emitting diodes, Sci. Rep. 3 (2017) 1–10, https://doi.org/10.1038/s41598-017-12161-5.

[28] C.F. Yang, C.Y. Lin, Investigation of luminescent properties of Eu^{3+} doped double perovskite Ba_2ZnMoO_6 phosphors by using solid-state reaction method, Microsyst. Technol. 24 (2018) 4067–4074, https://doi.org/10.1007/s00542-017-3624-1.

[29] Z. Ci, Y. Wang, J. Zhang, $Ca_{1-x}Mo_{1-y}NbyO4:Eu_x^{3+}$: A novel red phosphor for white light emitting diodes, J. Phys. Conf. Ser. 152 (2009) 3–9, https://doi.org/10.1088/1742-6596/152/1/012091.

[30] M.K. Sahu, H. Kaur, B.V Ratnam, J.S. Kumar, M. Jayasimhadri, Structural and spectroscopic characteristics of thermally stable Eu^{3+} activated barium zinc orthophosphate phosphor for white LEDs, Ceram. Int. 46 (2020) 26410–26415, https://doi.org/10.1016/j.ceramint.2020.07.263.

[31] J. Lin, M. Yu, C. Lin, X. Liu, Multiform oxide optical materials via the versatile pechini-type sol-gel process: synthesis and characteristics, J. Phys. Chem. C. 111 (2007) 5835–5845, https://doi.org/10.1021/jp070062c.

[32] C. Lin, C. Zhang, J. Lin, Phase transformation and photoluminescence properties of nanocrystalline ZrO2 powders prepared via the pechini-type sol-gel process, J. Phys. Chem. C. 111 (2007) 3300–3307, https://doi.org/10.1021/jp066615l.

[33] C. Guo, F. Gao, Y. Xu, L. Liang, F.G. Shi, B. Yan, Efficient red phosphors $Na_5Ln(MoO_4)_4$: Eu^{3+} (Ln = La, Gd and Y) for white LEDs, J. Phys. D. Appl. Phys. (2009) 42, https://doi.org/10.1088/0022-3727/42/9/095407.

[34] A. Kumar Choubey, Mechanoluminescence & thermoluminescence of SrAL2O4: Eu nano-phosphors, Open Nanosci. J. 5 (2011) 41–44, https://doi.org/10.2174/1874140101105010041.

[35] M.M. Kolte, V.B. Pawade, A.B. Bhattacharya, S.J. Dhoble, Tunable emission in Ln^{3+} (Ce^{3+}/Dy^{3+}, Ce^{3+}/Tb^{3+}) doped KNa3Al4Si4O16 phosphor synthesized by combustion method, J. Phys. Chem. Solids. 116 (2018) 338–346, https://doi.org/10.1016/j.jpcs.2018.02.004.

[36] Z.M. Chen, S.X. Jiang, R.H. Guo, B.J. Xin, D.G. Miao, Microstructure and photoluminescent properties of Y_2O_3:Eu^{3+} phosphors synthesised by precipitation and combustion methods, Mater. Technol. 29 (2014) 198–203, https://doi.org/10.1179/1753555713Y.0000000123.

[37] A. Prasad, A.S. Rao, G.V. Prakash, A study on up-conversion and energy transfer kinetics of $KGdF_4$:Yb^{3+}/Er^{3+} nanophosphors, J. Mol. Struct. 1205 (2020), 127647, https://doi.org/10.1016/j.molstruc.2019.127647.

[38] A. Prasad, A.S. Rao, G.V. Prakash, Up-conversion luminescence and EPR properties of $KGdF_4$:Yb^{3+}/Tm^{3+} nanophosphors, Optik 208 (2020), 164538, https://doi.org/10.1016/j.ijleo.2020.164538.

[39] C. Yang, C. He, X. Huang, G. Chen, G. Leng, R. Mi, X. Min, light-emitting diodes, Ceram. Int. 6 (2021) 2–10, https://doi.org/10.1016/j.ceramint.2021.01.087.

[40] V.V. Popov, A.P. Menushenkov, A.A. Yastrebtsev, A.Y. Molokova, S.G. Rudakov, R. D. Svetogorov, N.A. Tsarenko, K.V. Ponkratov, N.V. Ognevskaya, O.N. Seregina, The effect of the synthesis conditions on the structure and phase transitions in $Ln_2(MoO_4)_3$, Solid State Sci 112 (2021), 106518, https://doi.org/10.1016/j.solidstatesciences.2020.106518.

[41] S. Behara, K. Ikeda, T. Thomas, Amphoteric behavior of Dy in Na0.5Bi0.5TiO3: Neutron diffraction and Raman studies, Ceram. Int. 47 (2021) 12870–12878, https://doi.org/10.1016/j.ceramint.2021.01.149.

[42] N. Jain, R. Paroha, R.K. Singh, S.K. Mishra, S.K. Chaurasiya, R.A. Singh, J. Singh, Synthesis and rational design of europium and lithium doped sodium zinc molybdate with red emission for optical imaging, Sci. Rep. 9 (2019) 1–14, https://doi.org/10.1038/s41598-019-38787-1.

[43] M. Song, Y. Liu, Y. Liu, L. Wang, N. Zhang, X. Wang, Z. Huang, C. Ji, Sol-gel synthesis and luminescent properties of a novel $KBaY(MoO_4)^{3+}$: Dy^{3+} phosphor for white light emission, J. Lumin. 211 (2019) 218–226, https://doi.org/10.1016/j.jlumin.2019.03.052.

[44] B. Verma, R.N. Baghel, D.P. Bisen, N. Brahme, A. Khare, Synthesis and concentration dependent luminescent characterization of $BaMgSiO_4$: Eu^{3+} phosphor, J. Alloys Compd. 838 (2020), 155326, https://doi.org/10.1016/j.jallcom.2020.155326.

[45] R. Mahajan, R. Prakash, S. Kumar, V. Kumar, J. Choudhary, D.M. Phase, Optik Surface and luminescent properties of $Mg_3(PO_4)_2$: Dy^{3+} phosphors, Optik (Stuttg) 225 (2021), 165717, https://doi.org/10.1016/j.ijleo.2020.165717.

[46] H.C. Choi, Y.M. Jung, S. Bin Kim, Size effects in the Raman spectra of TiO^2 nanoparticles, Vib. Spectrosc. 37 (2005) 33–38, https://doi.org/10.1016/j.vibspec.2004.05.006.

[47] M.K. Sahu, M. Jayasimhadri, K. Jha, B. Sivaiah, A.S. Rao, D. Haranath, Synthesis and enhancement of photoluminescent properties in spherical shaped Sm^{3+}/Eu^{3+} co-doped $NaCaPO_4$ phosphor particles for w-LEDs, J. Lumin. 202 (2018) 475–483, https://doi.org/10.1016/j.jlumin.2018.06.002.

[48] S. Kaur, M. Jayasimhadri, A.S. Rao, A novel red emitting Eu^{3+}doped calcium aluminozincate phosphor for applications in w-LEDs, J. Alloys Compd. 697 (2017) 367–373, https://doi.org/10.1016/j.jallcom.2016.12.150.

[49] B.V. Ratnam, M.K. Sahu, A.K. Vishwakarma, K. Jha, H.J. Woo, K. Jang, M. Jayasimhadri, Optimization of synthesis technique and luminescent properties in Eu^{3+}-activated $NaCaPO_4$ phosphor for solid state lighting applications, J. Lumin. 185 (2017) 99–105, https://doi.org/10.1016/j.jlumin.2017.01.006.

[50] M.K. Sahu, M. Jayasimhadri, Conversion of blue emitting thermally stable $Ca_3Bi(PO_4)_3$ host as a color tunable phosphor via energy transfer for luminescent devices, J. Lumin. 227 (2020), 117570, https://doi.org/10.1016/j.jlumin.2020.117570.

[51] X. Gao, F. Song, A. Khan, Z. Chen, D. Ju, X. Sang, Room temperature synthesis, Judd Ofelt analysis and photoluminescence, J. Lumin. 230 (2021), 117707, https://doi.org/10.1016/j.jlumin.2020.117707.

[52] H. Kaur, M. Jayasimhadri, Optimization of structural and luminescent properties with intense red emitting thermally stable Sm^{3+} doped $CaBiVO_5$ phosphors for w-LED applications, Opt. Mater. (Amst). 107 (2020), 110119, https://doi.org/10.1016/j.optmat.2020.110119.

[53] B. Li, X. Huang, J. Lin, phosphors with high-e ffi ciency : Photoluminescence, energy transfer and application in near-UV-pumped white LEDs, J. Lumin. 204 (2018) 410–418, https://doi.org/10.1016/j.jlumin.2018.08.044.

[54] M. Jiao, Q. Xu, Y. Zhao, D. Wang, L. Liu, C. Yang, Photoluminescence properties and energy transfer of high thermal stable Na2GdPO4F2 : RE (RE¼ Sm3+,Ce 3+, Tb 3+) phosphor for solid-state lighting, 226 (2020).

[55] Y. Il Jeon, L.K. Bharat, J. Su, phosphors for solid-state lighting applications, J. Lumin. 166 (2015) 93–100, https://doi.org/10.1016/j.jlumin.2015.05.010.

[56] X. Fan, J. Si, M. Xu, G. Li, J. Tang, G. Cai, Photoluminescence and energy transfer of efficient and thermally stable white-emitting $Ca_9La(PO_4)_7$:Ce^{3+}, Tb^{3+}, Mn^{2+} phosphors, Ceram. Int. 47 (2021) 12056–12065, https://doi.org/10.1016/j.ceramint.2021.01.049.

[57] V. Uma, M. Vijayakumar, K. Marimuthu, G. Muralidharan, Luminescence and energy transfer studies on Sm^{3+} /Tb^{3+} codoped telluroborate glasses for WLED applications TBLT2S, J. Mol. Struct. 1151 (2018) 266–276, https://doi.org/10.1016/j.molstruc.2017.09.053.

[58] S. Kaur, A.S. Rao, M. Jayasimhadri, B. Sivaiah, D. Haranath, Synthesis optimization, photoluminescence and thermoluminescence studies of Eu^{3+} doped calcium aluminozincate phosphor, J. Alloys Compd. 802 (2019) 129–138, https://doi.org/10.1016/j.jallcom.2019.06.169.

[59] Y. Chang, Z. Shi, Y. Tsai, S. Wu, H. Chen, The effect of Eu3+ -activated InVO4 phosphors prepared by sol – gel method, 33 (2011) 375–380. https://doi.org/10.1016/j.optmat.2010.09.018.

[60] A. Prasad, A.S. Rao, M. Gupta, G.V. Prakash, Morphological and luminescence studies on $KGdF_4$:Yb^{3+}/Tb^{3+} up-conversion nanophosphors, Mater. Chem. Phys. 219 (2018) 13–21, https://doi.org/10.1016/j.matchemphys.2018.07.056.

Arabian Journal for Science and Engineering
https://doi.org/10.1007/s13369-021-06379-9

RESEARCH ARTICLE-ELECTRICAL ENGINEERING

Systematic Realization of CFOA Based Rössler Chaotic System and Its Applications

Kriti Suneja[1] · **Neeta Pandey**[1] · **Rajeshwari Pandey**[1]

Received: 2 April 2021 / Accepted: 8 November 2021
© King Fahd University of Petroleum & Minerals 2021

Abstract
In this paper, systematic realization of Rössler's chaotic system is put forward. The governing equations of Rössler chaotic system are realized with current feedback operational amplifier (CFOA) and an analog multiplier. Existing Rössler chaotic systems use voltage mode active block namely operational amplifier where additional active blocks are required for addition/subtraction of signals. On the other hand, CFOA has capability of processing both voltage and current signals which results in design flexibility leading to component saving. The functionality of the proposed system is examined through LTspice simulations and experimentation; and observing the chaotic outputs, their frequency spectrum, and local attractors in phase trajectories in different planes. The minimum frequency range for noise-like behavior of state variables is observed to be 50 Hz to 10 kHz, thus it can be used for audio frequency applications. The robustness of the proposed circuit is examined against component variation through Monte Carlo analysis. The usability of the proposed system is illustrated in adaptive synchronization between two Rössler chaotic systems and signal masking. Simulation results for both the schemes are also included for completeness. The impact of variation in master and slave design parameter is studied and it is found that the percentage error is within 1% when same parameter values are used for master and slave design. Further, the effect of message signal amplitude and its frequency is examined in masking and it is observed that the signal is reproduced at receiver with a minor error of 3.33% in amplitude.

Keywords Chaos · Synchronization · Nonlinear circuits · Analog building blocks

1 Introduction

CHAOTIC systems find wide range of applications in secure communication [1–3], speech encryption [4], image encryption [5, 6], text encryption [7, 8], weather prediction, [9] and solar irradiance forecasting [10] due to their high sensitivity on initial conditions, nonlinearity and non-repeatability of trajectory. These systems are represented by set of differential equations and differ from each other in number of terms and types of nonlinearities. It is pertinent to mention here that number of nonlinear terms infers complexity in hardware realization of the system. Some of the popular chaotic systems having two nonlinear terms are Lorenz [11], Sprott case A-E [12], and Pehlivan [13] while Chua's [14], Sprott case F-S [12], and Rössler [15] are representatives of chaotic systems with single nonlinear term.

The actual physical implementation of an attractor started with the design of Chua's circuit [14, 16], which consisted of two capacitors, one inductor, one passive resistor, and one nonlinear negative resistor, also known as Chua's diode. Subsequently, researchers proposed Operational Amplifier (OpAmp) based hardware designs of Lorenz's [17], Pehlivan's [13], Rössler's [18], and Chua's [19, 20] chaotic circuits. The OpAmp processes voltage signals and requires additional active block for addition/subtraction of signals while current mode and mixed-mode building blocks provide flexibility for processing both voltage and current signals and have been gainfully used to design chaotic circuits in the recent past. Some of the representative examples are—current feedback operational amplifiers (CFOAs) [21, 22], current conveyors (CCII + s) [23], Voltage Differencing Transconductance Amplifier (VDTA) [24], operational trans resistance amplifier (OTRA) [25], differential voltage current conveyor transconductance amplifier (DVCCTA) [26] based Chua's circuit, CCII + s based Lorenz circuit [27], CCII based RC chaotic generators [28] and Sprott H circuit

✉ Neeta Pandey
n66pandey@rediffmail.com

[1] Delhi Technological University, New Delhi, India

Published online: 17 January 2022

[29]. The hardware realization of Chua's [21–26] chaotic system has been explored more than that of Lorenz's and Pehlivan's which may be due to presence of single nonlinear term. Rössler is yet another variant of chaotic systems based on single nonlinear term. The available Rössler chaotic system [18, 30] is built around operational amplifier (OpAmp). The circuit reported in [18] uses four OpAmps, one analog multiplier (AM), ten resistors, and three capacitors while [30] employs six OpAmps, two analog multiplier (AM), thirteen resistors, and three capacitors. A variant of Rossler chaotic system is presented in [31, 32] wherein the governing equations are modified while keeping the nature of attractors intact. Both realizations employ five OpAmps, one diode, ten resistors, and three capacitors. Current mode building block-,based realization of Rossler chaotic system is not available in open literature to the best of authors' exposure in the field. To fill this gap, a CFOA based Rössler chaotic system is presented in this paper. The proposed realization utilizes four CFOAs, one Analog multiplier, seven resistors, and three grounded capacitors. Thus, overall component count is minimum in comparison to its OpAmp based counterparts [18, 30–32]. Further, a complete adaptive synchronization system has been designed as a standalone circuit for potential applications in communication and encryption. To the best of our knowledge, the CFOA based circuit design of adaptive control synchronization between two identical chaotic systems has not yet been proposed. The circuitry for masking scheme has also been designed with self-synchronization, which utilizes eight CFOAs, two AMs, eleven resistors, and five capacitors. This count is significantly less than that of Chua's circuit-based masking scheme proposed in [23], which utilizes twenty current conveyors (CCII +) embedded in AD844s, seventeen resistors, and six capacitors.

The paper is organized as follows: Sect. 2 describes the governing equations of Rössler chaotic system and their systematic realization using CFOA. This section also includes simulation and experimental results for the state variables and phase trajectories in various planes. The mathematical formulation for adaptive synchronization of Rössler chaotic system is described in Sect. 3 which is followed by CFOA realization of complete system. The chaotic masking scheme is briefly described in Sect. 4 first which is followed by its CFOA based realization. The function verification of the CFOA based circuits in Sects. 3 and 4 are placed in corresponding sections. The paper is finally concluded in Sect. 5.

2 Proposed Circuit

In this section, CFOA based realization of Rössler chaotic system is proposed. The governing equations of Rössler chaotic system [13] are given as

$$\dot{x} = -(y + z) \tag{1a}$$

$$\dot{y} = x + ay \tag{1b}$$

$$\dot{z} = b + z(x - c) \tag{1c}$$

Here, the dots above the state variables x, y, and z represent their derivatives respectively which evolve with time. The a, b and c are positive constants whose values for the system to behave chaotically, as studied by Rössler for the first time, are 0.2, 0.2, and 5.7 respectively.

The block diagrammatic representation of CFOA (AD844) and its internal details is given in Fig. 1a. There are two voltage buffers which are used to transfer voltages of Y and Z ports to ports X and O respectively. The current flows through Y port is zero and the current flowing through X port is conveyed to Z port. Thus, constituent port relations of CFOA are given as $V_X = V_Y$; $V_O = V_Z$, $I_Y = 0$, and $I_Z = I_X$. The realization of (1) requires three capacitors (C_x, C_y, and C_z) and corresponding voltages V_x, V_y, and V_z may be treated as three state variables. An analog multiplier AD633, as shown in Fig. 1b, is used for producing multiplication of state variables x and z in (1c). The internal details of AD633 are given in Fig. 1b. It uses two voltage subtractors for calculating difference between two inputs (V_{x1}-V_{x2}, V_{y1}-V_{y2}), one multiplier for providing$((V_{x1} - V_{x2})(V_{y1} - V_{y2})/10)$, one adder to give $((V_{x1} - V_{x2})(V_{y1} - V_{y2})/10) + V_z$ and one voltage buffer that provides multiplier output at low impedance. The port relation of AD633 is described by (2).

$$V_w = \frac{(V_{x1} - V_{x2})(V_{y1} - V_{y2})}{10} + V_z \tag{2}$$

The CFOA based implementation of (1a–1c) is given in Fig. 2a–c respectively.

The analysis of the proposed circuit gives the following equations

$$\frac{dV_x}{dt} = \frac{-V_y}{R_1 C_x} - \frac{V_z}{R_2 C_x} \tag{2a}$$

$$\frac{dV_y}{dt} = \frac{V_y}{R_4 C_y} + \frac{V_x}{R_3 C_y} \tag{2b}$$

$$\frac{dV_z}{dt} = \frac{V_x V_z}{10 R_5 C_z} - \frac{V_z V_c}{10 R_5 C_z} - \frac{V_b}{R_7 C_z} \tag{2c}$$

Comparing (1) and (2), the set of equations obtained are shown in (3).

$$\begin{aligned} &\frac{1}{R_1 C_x} = 1,\ \frac{1}{R_2 C_x} = 1,\ \frac{1}{R_3 C_y} = 1,\ \frac{1}{R_4 C_y} = 0.2, \\ &\frac{V_b}{R_7 C_z} = -0.2,\ \frac{1}{10 R_5 C_z} = 1 \text{and} V_c = 5.7 \end{aligned} \tag{3}$$

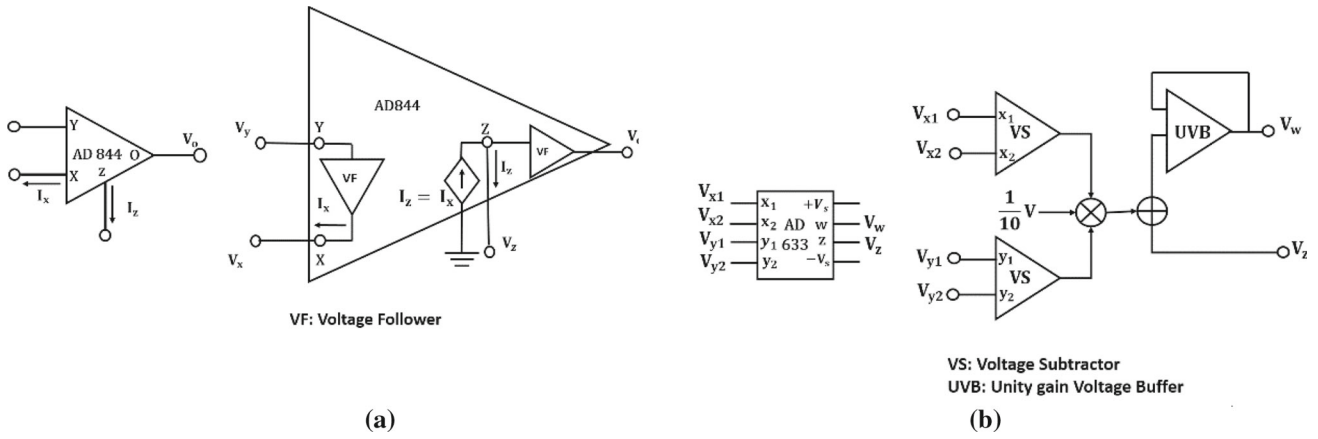


Fig. 1 Block diagram and internals details of **a** CFOA **b** AD633

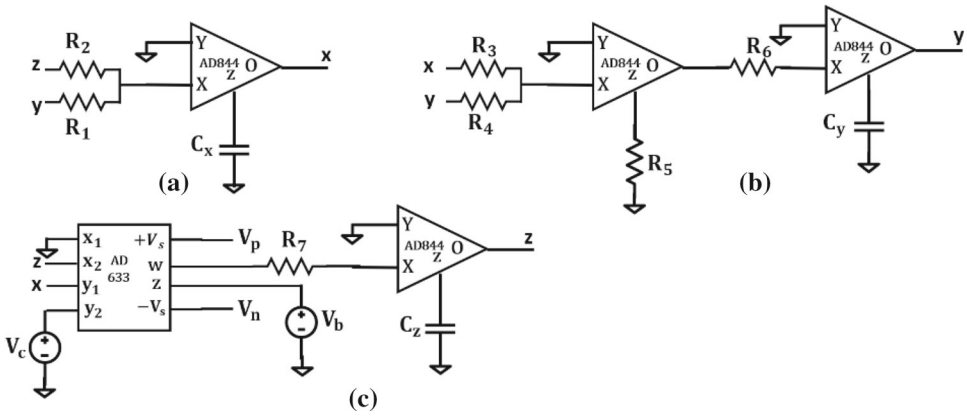


Fig. 2 Proposed CFOA based Rössler system

In order to validate the proposed design, simulations have been carried out in LTspice design suite. The simulation steps include problem definition, system definition, model formulation and translation, verification and validation, and experimentation and analysis in order. The methodology used for incorporating these steps are illustrated below:

Step 1 For the designed Rössler chaotic circuit, the port relationships were equated with the governing equations to find the components' values. A scaling factor of 1000 is used for obtaining the component values. The resulting components are $R_1 = R_2 = R_3 = R_5 = R_6 = 10\,\text{k}\Omega$, $R_4 = 50\,\text{k}\Omega$, $R_7 = 1\text{k}\Omega$ and $C_x = C_y = C_z = 100\text{nF}$.

Step 2 The passive components Resistors, Capacitors, etc. are available in LTspice. The macro models of the active blocks AD844 and AD633 [33, 34] have been imported as netlist from Analog Devices. The models' netlist includes Bipolar Junction Transistors and diode with the reverse saturation current IS = 10-15 A. The power supplies are chosen as ±15V from the permissible supply voltages for AD844 and AD633 which are (± 4.5V to ± 18V) and (±8V to ±18V), respectively. The voltages across capacitors representing the state variables, and the values of capacitors and resistors equating with the coefficients of the differential equations, we plotted the time series, phase space diagrams, and FFTs to study the behavior of the circuit. The initial conditions were (0,0,0).

Step 3 The circuit was designed and simulated to obtain time series, phase space diagrams, and FFT plots. The phase plots of the Rössler chaotic system obtained from the designed circuit confirms the correctness of the proposed design. Also, to validate the design, it has been shown theoretically the equivalence between the representative equations of the system and the KVL equations of the circuit.

Figures 3, 4 and 5 show simulated results for state variables V_x, V_y, and V_z, all three are aperiodic with different shapes. The Fourier transform of the plots of Fig. 3 is placed in Fig. 4. The observed spectrums for state variables x, y, and z are noise-like in frequency ranges of (50 Hz to 10 kHz), (50 Hz to 10 kHz), and (50 Hz to 80 kHz) respectively. Thus, the proposed circuit can be used for audio applications.

Further, Monte Carlo (MC) analysis is also performed by considering 10% tolerance in the passive component values. The simulated phase trajectories in $V_x - V_y$, $V_y - V_z$, and V_x

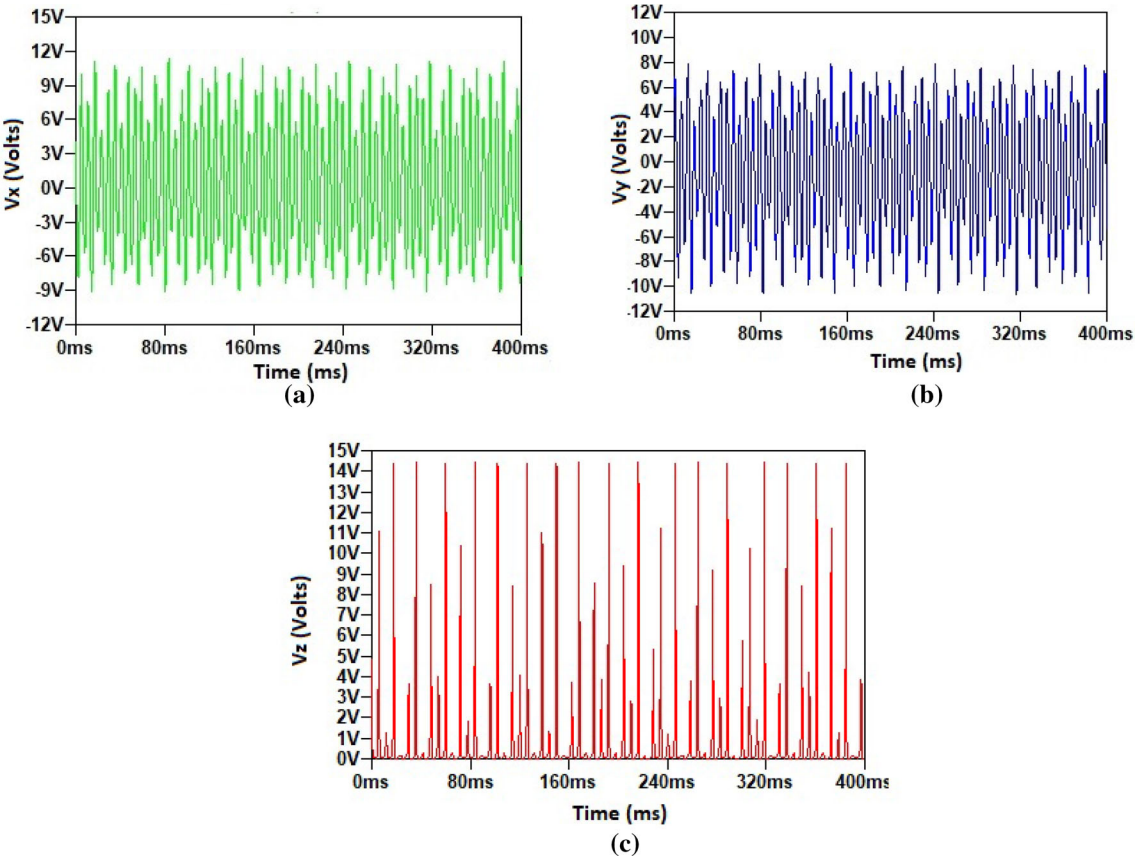


Fig. 3 Simulated chaotic outputs at **a** V_x, **b** V_y, **c** V_z

$-V_z$ plane, shown in Fig. 6. It may be noted that all trajectories have strange local attractors and are of similar nature as those without MC analysis, suggesting the robustness of the design.

The proposed realization is compared with the existing ones [18, 30–32] through simulations. The component values for existing circuits [18, 30–32] are taken from the respective references and the power supplies are kept same as used in the proposed circuit ($\pm$ 15 V) for fair comparison. The power consumption is found to be 1.87 W, 2.78 W, 2.13 W, 2.13 W, and 1.1 W, for the circuits reported in [18, 30–32]and the proposed one respectively. The proposed circuit consumes the least power under same supply voltage condition.

Experimental verification of the proposed circuit in Fig. 2 is also done by breadboarding it using commercially available ICs AD844 and AD633. The supply voltages and component values are kept same as those used in simulative investigations. The experimental time-domain outputs for state variables (V_x V_y), (V_y V_z), and (V_x,V_z) are shown in Fig. 7a–c respectively and corresponding phase trajectories are shown in Figs. 7d–f. It may be noted that the experimental results are in good agreement with the LTspice simulations.

3 Synchronization Between Rössler Systems

In order to utilize chaotic signals in private communications, the receiver's chaotic signals must match with that of transmitter's. In other words, synchronization is required between transmitter and receiver. Carroll and Pecora were first ones to introduce synchronization among chaotic systems [35]. Different coupling schemes can be used to synchronize receiver with the transmitter, such as active control, backstepping method, sliding mode control, adaptive control, etc. [36–41]. Adaptive control synchronization is used when the system parameters are initially uncertain or unknown. Thus, in this section, step by step process of adaptive synchronization as adopted in [42] with its circuit design has been elaborated.

Let master and slave Rössler systems are characterized by state variables (x_1, y_1, z_1) and (x_2, y_2, z_2) respectively, and their corresponding governing equations are given by (4) and (5).

$$\begin{aligned} \dot{x_1} &= -(y_1 + z_1) \\ \dot{y_1} &= x_1 + a y_1 \\ \dot{z_1} &= b + z_1(x_1 - c) \end{aligned} \tag{4}$$

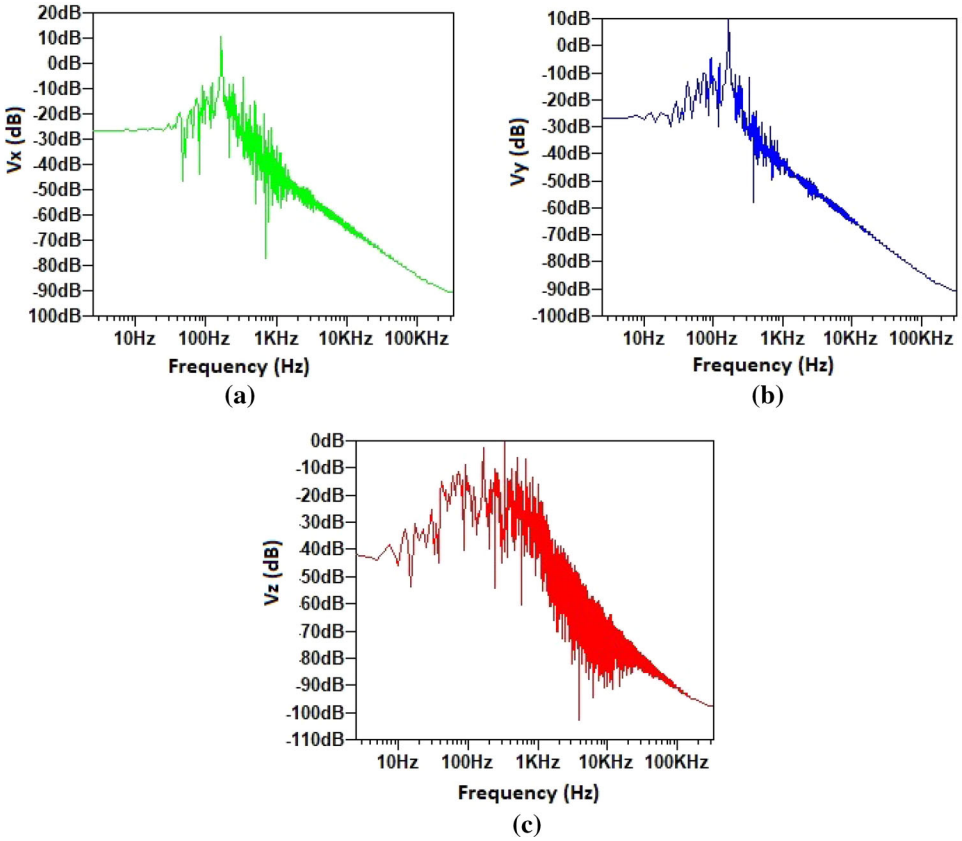


Fig. 4 Fourier transform plot of **a** V_x, **b** V_y, **c** V_z

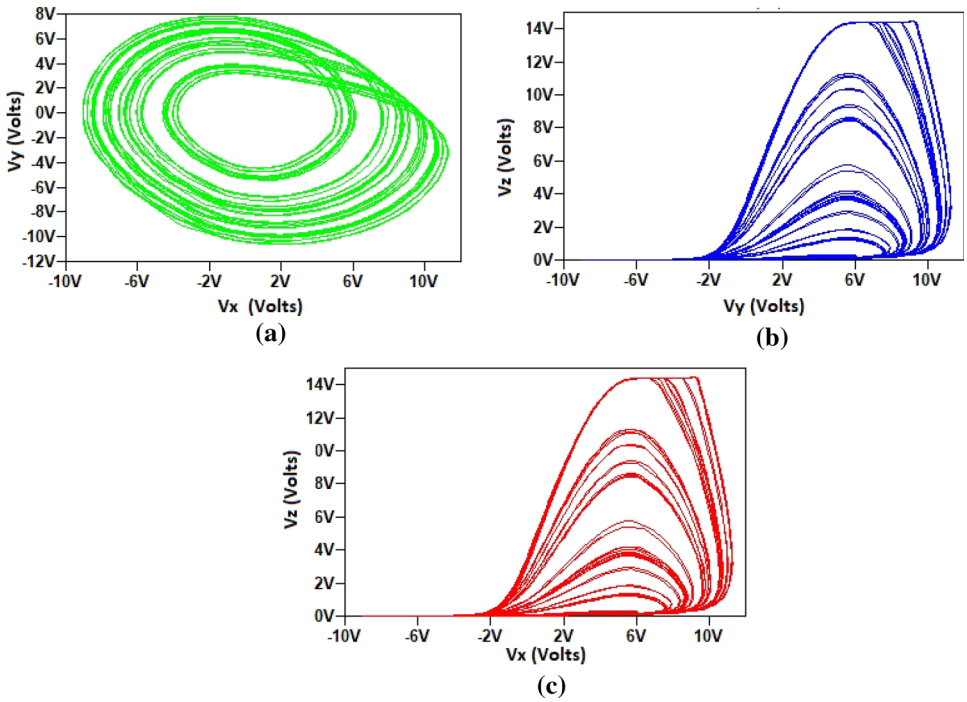


Fig. 5 Simulated phase trajectories in **a** $V_x - V_y$, **b** $V_y - V_z$, and **c** $V_x - V_z$ plane

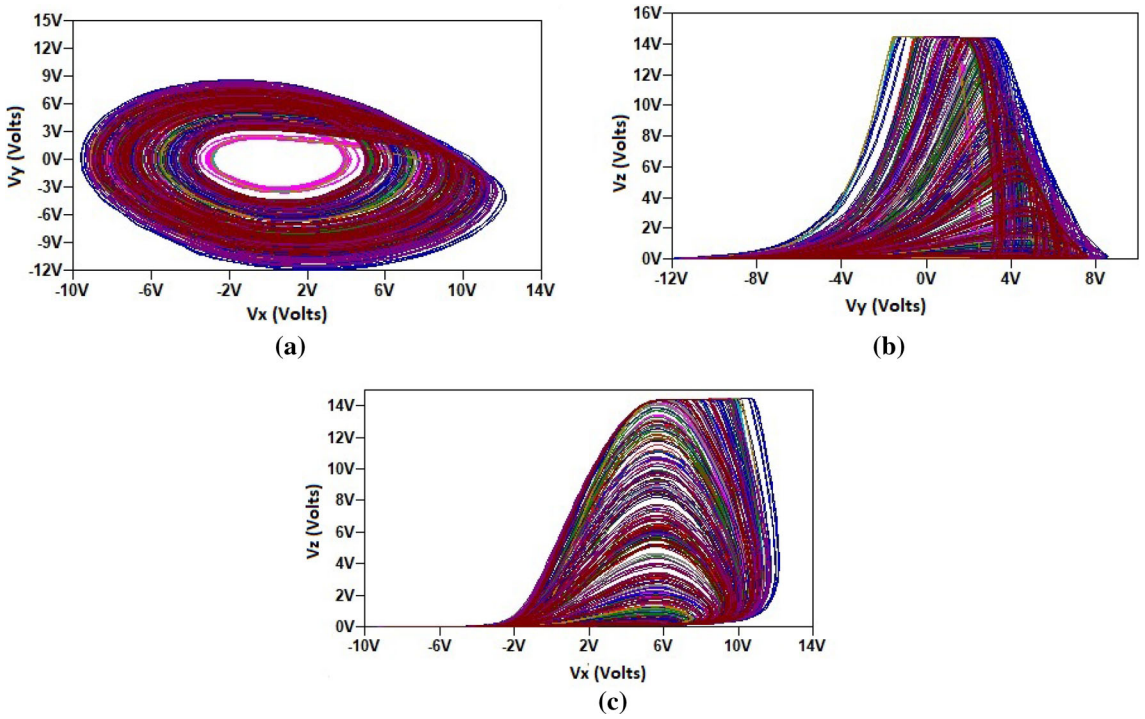


Fig. 6 Simulated phase trajectories in **a** $V_x - V_y$, **b** $V_y - V_z$, and **c** $V_x - V_z$ plane with MC analysis

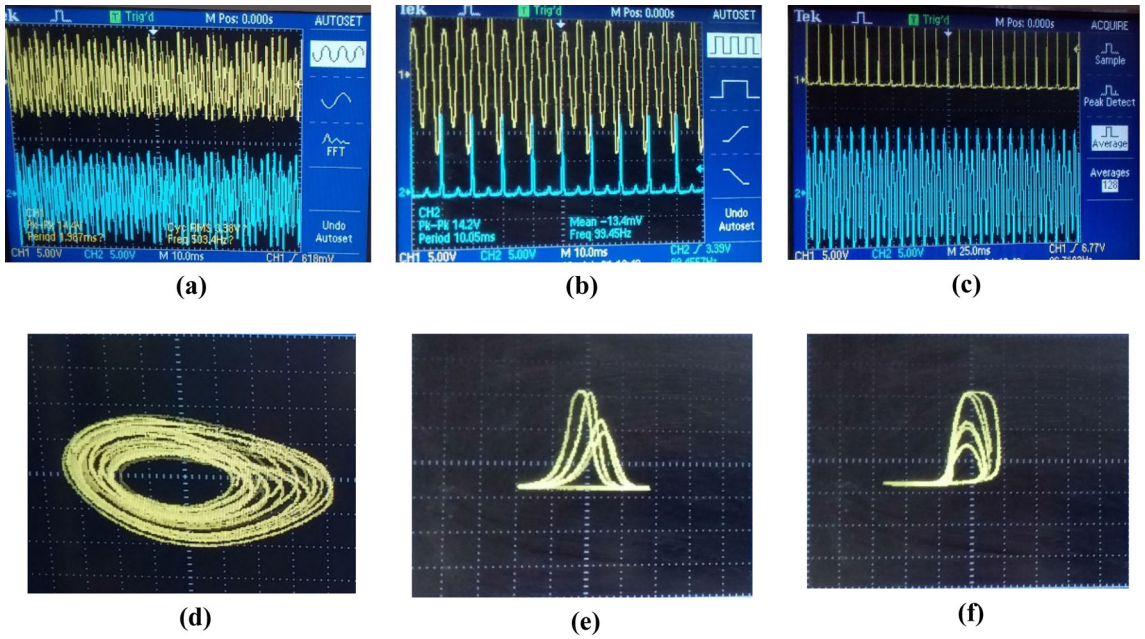


Fig. 7 Experimental time-domain outputs for state variables **a** *Vx,Vy* **b** *Vy,Vz* **c** *Vx,Vz* and phase trajectories in **d** $Vx - Vy$, **e** $Vy - Vz$ and **f** $Vx - Vz$ plane

$$\begin{aligned} \dot{x_2} &= -(y_2 + z_2) + u_1 \\ \dot{y_2} &= x_2 + a y_2 + u_2 \\ \dot{z_2} &= b + z_2(x_2 - c) + u_3 \end{aligned} \tag{5}$$

where u_1, u_2 and u_3 are adaptive functions controlling the synchronization between master and slave.

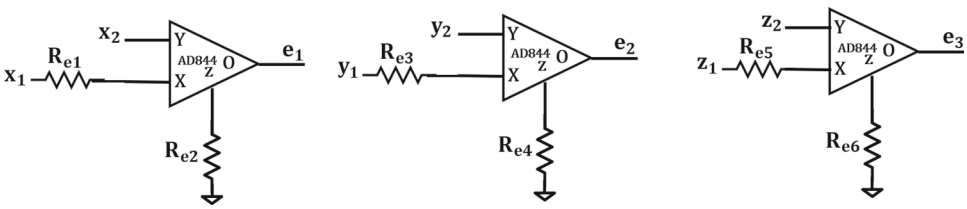


Fig. 8 Error functions of adaptive control synchronization

The errors in synchronization between master and slave are given in (6) and the corresponding CFOA based circuit is shown in Fig. 8 ($R_{e1} = R_{e2}, R_{e3} = R_{e4}, R_{e5} = R_{e6}$).

$$\begin{aligned} e_1 &= x_2 - x_1 \\ e_2 &= y_2 - y_1 \\ e_3 &= z_2 - z_1 \end{aligned} \tag{6}$$

Differentiating (6) with respect to time and putting the values of time differentiated state variables from (4) and (5), the error dynamics is obtained as:

$$\begin{aligned} \dot{e}_1 &= -e_2 - e_3 + u_1 \\ \dot{e}_2 &= e_1 + ae_2 + u_2 \\ \dot{e}_3 &= ce_3 + x_2z_2 - x_1z_1 + u_3 \end{aligned} \tag{7}$$

From (7), the adaptive functions controlling the synchronization between master and slave are computed as:

$$\begin{aligned} u_1 &= -k_1e_1 + e_2 + e_3 \\ u_2 &= -k_2e_2 - e_1 - \widehat{a}e_2 \\ u_3 &= -k_3e_3 + \widehat{c}e_3 - x_2z_2 + x_1z_1 \end{aligned} \tag{8}$$

where, $\widehat{a}$ and $\widehat{c}$ are the estimates of parameter a and c respectively; and k_1, k_2, and k_3 are positive constants.

Putting u_1, u_2 and u_3 from (8) in (7), we obtain:

$$\begin{aligned} \dot{e}_1 &= -k_1e_1 \\ \dot{e}_2 &= (a - \widehat{a})e_2 - k_2e_2 \\ \dot{e}_3 &= (c - \widehat{c})e_3 - k_3e_3 \end{aligned} \tag{9}$$

In order to maintain the stability of the synchronized system, the control parameters are updated using quadratic scalar function of (10) [42].

$$V(e_1, e_2, e_3, e_a, e_b, e_c) = \frac{1}{2}({e_1}^2 + {e_2}^2 + {e_3}^2 + {e_a}^2 + {e_b}^2 + {e_c}^2) \tag{10}$$

where, e_a, e_b, $and e_c$ are the errors in the parameters $(a - \widehat{a})$, $\left(b - \widehat{b}\right)$, $and (c - \widehat{c})$ whose time derivatives are given by (11).

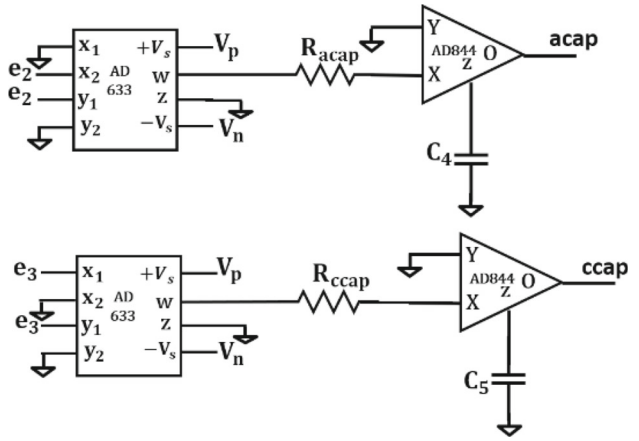


Fig. 9 Estimated parameters a and c

$$\begin{aligned} \dot{e_a} &= -\dot{\hat{a}} \\ \dot{e_b} &= -\dot{\hat{b}} \\ \dot{e_c} &= -\dot{\hat{c}} \end{aligned} \tag{11}$$

Differentiating (10) with respect to time to determine the stability of the system, we obtain:

$$\begin{aligned} \dot{V} = &-k_1{e_1}^2 - k_2{e_2}^2 - k_3{e_3}^2 \\ &+ e_a({e_2}^2 - \dot{\hat{a}}) - e_b\left(\dot{\hat{b}}\right) - e_c({e_3}^2 + \dot{\hat{c}}) \end{aligned} \tag{12}$$

For the system to be stable, $\dot{V}$ has to be negative. The first three terms of (12) are negative. If the last three terms are equated to zero, the estimated parameters are updated by the following law:

$$\begin{aligned} \dot{\hat{a}} &= {e_2}^2 \\ \dot{\hat{b}} &= 0 \\ \dot{\hat{c}} &= -{e_3}^2 \end{aligned} \tag{13}$$

The circuit design of (13) using CFOAs and AD633 is shown in Fig. 9.

The complete set of equations for adaptively synchronized slave system is given by (14).

$$\dot{x_2} = -(y_2 + z_2) - k_1e_1 + e_2 + e_3$$

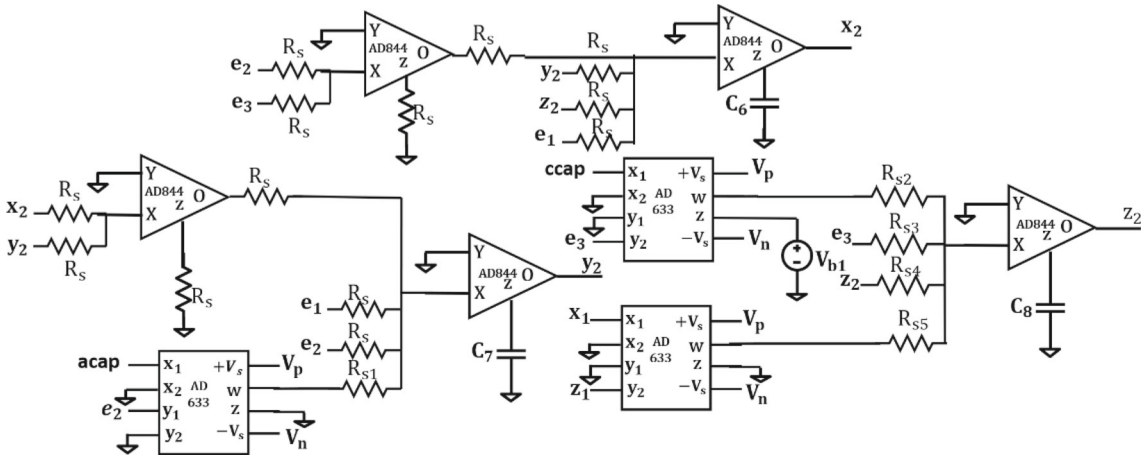


Fig. 10 Proposed CFOA slave circuit

$$\dot{y_2} = x_2 + ay_2 - k_2e_2 - e_1 - \widehat{a}e_2$$
$$\dot{z_2} = b + z_2(x_2 - c) - k_3e_3 + \widehat{c}e_3 - x_2z_2 + x_1z_1 \quad (14)$$

and its complete hardware realization is given in Fig. 10. The e_i's (i = 1,2 3), $\widehat{a}$ and $\widehat{c}$ are obtained from circuit of Figs. 8 and 9.

The functionality of complete synchronized design is verified using proposed Rössler chaotic circuit of Fig. 2 as master and Fig. 10 as slave circuit. Following steps are used to obtain the simulation results:

Step 1 The component values for slave are computed so that the governing equations are satisfied. A scale of 1000 provides the following values of components $R_s = R_{s3} = 10$ kΩ; $R_{s1} = R_{s2} = R_{s5} = 1$ kΩ, $R_{s4} = 1.695$ kΩ,$R_{ei\ (i=1,2,\ldots,6)} = 10$ kΩ, $R_{acap} = R_{ccap} = 1$ kΩ and $C_{i\ (i=4,5\ldots,8)} = 100$ nF. The component values for master circuit are taken same as those given in Sect. 2.

Step 2 The time-domain response for state variables (x_1, x_2), (y_1, y_2), and (z_1, z_2) are plotted and the graphs for corresponding master and slave state variables are extracted (Fig. 11) to measure the degree of closeness between the respective state variables.

It is observed that x and y waveforms of master and slave have linear relationship supporting their synchronization. The z waveforms also have linear relationship but with 2.18% error in amplitude which may be attributed to this mismatch in the value of parameter 'c' in master (5.7) and slave (5.9) circuit.

To confirm the assertion that the error is due to mismatch in value of parameter 'c', the adaptively synchronized system, extensive simulations have been carried out to obtain synchronization error between the master and slave system for 'a' = 'b' = 0.2 and 'c' = 5, 5.7 5.9 and 6 as Rössler system behaves chaotically for these values [43]. The percentage error with respect to parameter 'c' is listed in Table 1. It may be noted from Table 1 that the percentage error when both master and slave are designed with same value of parameter 'c' is within 1% which is negligibly small to affect synchronization. An increasing trend in percentage error is observed for larger deviations in the values of parameter 'c' of master and slave.

Table 1 Percentage error in adaptive synchronization with respect to parameter 'c'

Transmitter ('c') ↓	Receiver ('c') →			
	Error (%)			
	5	5.7	5.9	6
5	0.96	6.19	6.39	6.67
5.7	6.27	0.98	2.18	3.59
5.9	6.84	2.72	0.96	2.17
6	7.48	3.51	1.65	0.93

Theoretically, the error reduces to zero with time in adaptive control synchronization and whole dynamics of receiver may become unstable if error increases with time. The simulation results of the designed circuit show a minor synchronization error in state variable z at receiver end and remain constant with time leading to stable system.

Experimental verification of the proposed circuits of Figs. 2 and 10 is also done by breadboarding them using commercially available ICs AD844 and AD633. The supply voltages and component values are kept same as those used in simulative investigations except R_{s4} = 1.695k Ω. The value of R_{s4} is taken as 1.68 kΩ. The experimental results for synchronization of state variable V_z are shown in Fig. 12. It may be noted that there is close agreement in state variables of master and slave system, thus experimental results are in line with the LTspice simulations.

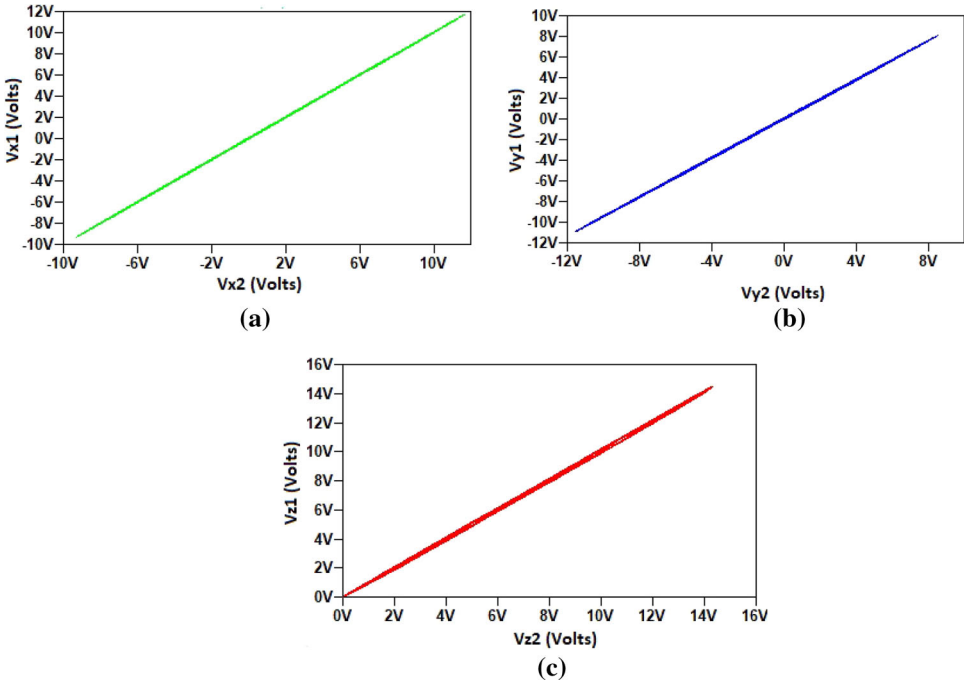


Fig. 11 Simulated plots between chaotic signals of master and slave in adaptive synchronization: **a** V_{x1} versus V_{x2}, **b** V_{y1} versus V_{y2}, **c** V_{z1} versus V_{z2}

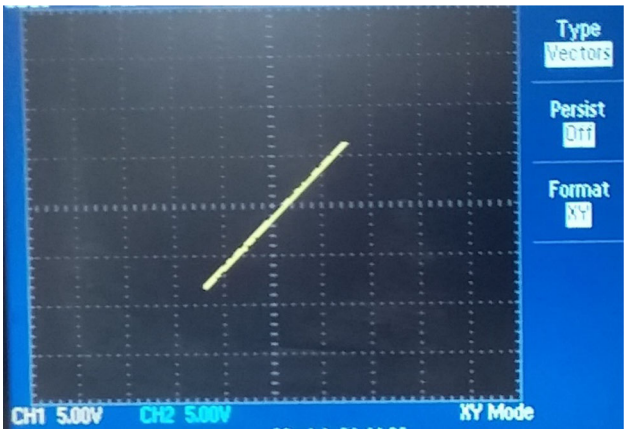


Fig. 12 Experimental plot between chaotic signals of master and slave in adaptive synchronization for state variable 'z'

4 Chaotic Masking in Communication Systems

Chaotic circuits have always been a choice for secure communication because of their noise-like spectrum. Different researches over the years have shown the applications of chaotic circuits in masking applications of communication [44–47]. In chaotic masking scheme, the chaotic signal $x(t)$, which acts as a key, is added to the message signal $m(t)$ at the transmitter side and the resultant signal $s(t)$ is transmitted to the receiver. To decode this encrypted signal $x_2(t)$, the same key needs to be generated at the receiver side, which is possible only if the transmitter and receiver are synchronized. The key at receiver end may be recovered by subtracting $x_2(t)$ from $s(t)$. Two identical dynamic systems are said to be in self-synchronization when one is driven by the other. The overall masking scheme is depicted in Fig. 13 where self-synchronization is implemented at receiver end.

It is pertinent to mention here that all the variables cannot be used to drive the other system and the choice of variable is decided by the conditional Lyapunov exponent [44]. The Rössler system has a positive conditional Lyapunov exponent for x and z drive, therefore, it cannot be synchronized with x and z drive. So, the state variable y_1 of the transmitter is used for driving the receiver in which resulting system is a unified self- synchronized system of five variables (x_1, y_1, z_1, x_2, z_2). If this unified system is to be converted into a circuit, it will lead to lesser number of components as compared to adaptive control synchronization method (detailed in Sect. 2).

The complete design for chaotic masking scheme is put forward in Fig. 14 wherein state variable y of transmitter is used as drive signal. The transmitter and receiver circuit comprises of Rössler chaotic circuit of Fig. 2 except the one required for y_2. Additionally, transmitter circuit uses a CFOA based adder (CFOA4) to add the chaotic signal x_1 to the message signal m while the chaotic signal x_2 is subtracted from the transmitted signal s to recover the message m_r using CFOA based subtractor (CFOA7) in receiver circuit. The masking scheme was implemented using self-synchronization, and the designed system was simulated to observe if the transmitted signal is completely recovered at the receiver. Following steps are used:

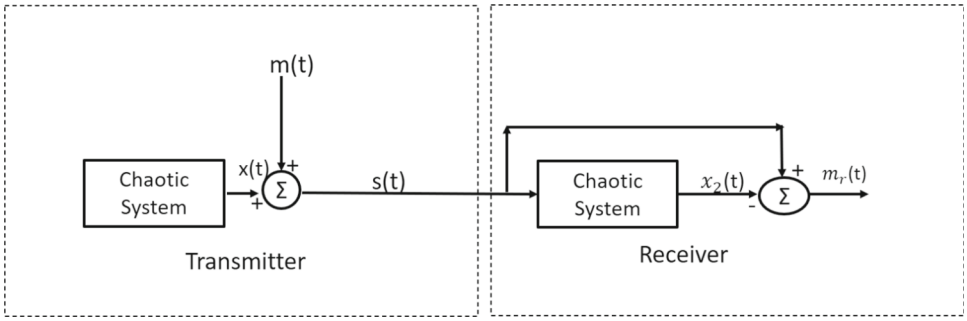


Fig. 13 Block diagram of chaotic masking scheme [46]

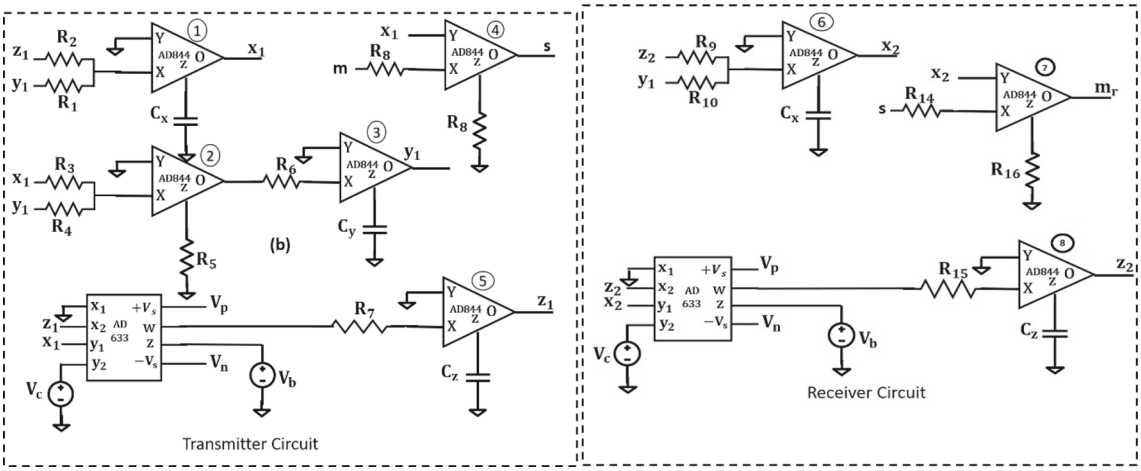


Fig. 14 Transmitter and receiver circuit in masking scheme

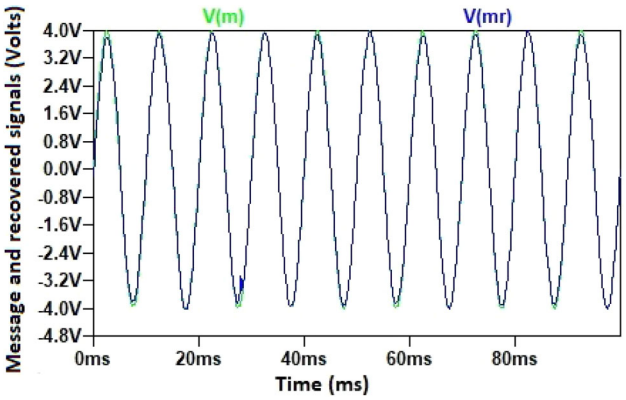


Fig. 15 Message and recovered signals in chaotic masking scheme

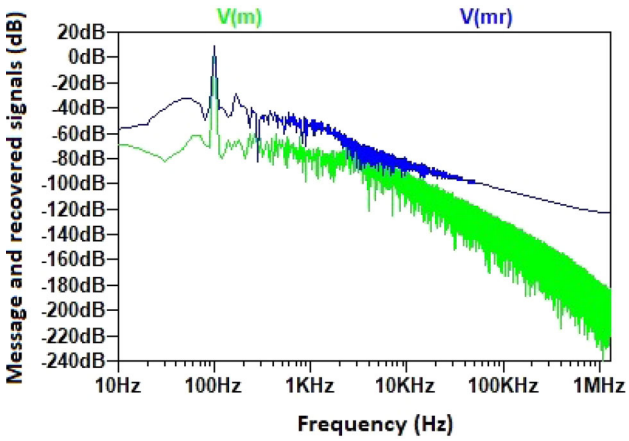


Fig. 16 Fast Fourier transform of message and recovered signal in chaotic masking scheme

Step 1 The component values for receiver are taken similar to those used in Sect. 2. The resistors R_8, R_{14}, and R_{16} which are used to add/ subtract message signal with chaotic signal are taken as 1kΩ.

Step 2 The time-domain simulations are carried out and both transmitted and recovered signals are plotted along with corresponding FFTs.

Step 3 The error in magnitude of recovered signal is observed for two conditions- (i) magnitude of message signal is varied while keeping frequency fixed; (ii) frequency of message signal is varied while keeping magnitude fixed.

The simulation results for transmitted signal m and received signal m_r are depicted in Fig. 15 and the corresponding spectrum is placed in Fig. 16. As can be seen from the spectrum that though there has been a change of around 25 dB in the amplitude but the peaks of the transmitted and

received signals in frequency spectrum coincide, thus making chaotic masking a feasible scheme for transmission of coded data. Further, the effect of message signal amplitude and its frequency is examined in masking and it is observed that the signal is reproduced at receiver with a minor error of 3.33% in amplitude.

5 Conclusion

In this work, a Rössler chaotic system has been designed systematically using CFOAs and analog multiplier. The proposed design uses four CFOAs, one analog multiplier, seven resistors, and three capacitors. The passive component count in proposed circuit is considerably low as compared to existing Rössler chaotic system. The workability of the proposed design is examined through LTspice simulations and experimentation using CFOA IC AD844 and analog multiplier IC AD633. The observed outputs for state variables are aperiodic and having strange local attractors for phase trajectories in different planes, which confirms the correctness of the proposed design. The power consumption of the proposed circuit is 1.1 W, at power supplies of ± 15 V. Two applications namely adaptive control synchronization between two Rössler chaotic systems and masking scheme for transmission are also put forward to illustrate the usability of the proposed circuit. In adaptive control synchronization, the master and slave systems are found to be in synchronization with a maximum error under 1%. The self-synchronized receiver in masking scheme is also being able to reproduce the message signal with minute error of 3.33% in amplitude at a sinusoidal signal of amplitude 3 V and frequency 100 Hz was transmitted. Also, the designed Rössler chaotic circuit is robust within 10% tolerance of passive components. This is confirmed from MC analysis done for the designed circuit. Thus, we propose to use this chaotic circuit design in communication applications as carrier signal generator.

References

1. Alvarez, G.; Montoya, F.; Romera, M.; Pastor, G.: Breaking two secure communication systems based on chaotic masking. IEEE Trans. Circuits Syst **II**(51), 505–506 (2004). https://doi.org/10.1109/TCSII.2004.836047
2. Wang, T.; Wang, D.; Wu, K.: Chaotic adaptive synchronization control and application in chaotic secure communication for industrial internet of things. IEEE Access. **6**, 8584–8590 (2018). https://doi.org/10.1109/ACCESS.2018.2797979
3. Zhengguo, L.; Kun, L.; Changyun, W.; Yeng Chai, S.: A new chaotic secure communication system. IEEE Trans. Comm. **51**(8), 1306–1312 (2003). https://doi.org/10.1109/TCOMM.2003.815058
4. Hato, E.; Shihab, D.: Lorenz and Rossler chaotic system for speech signal encryption. Int. J. Comput. Appl. **128**, 25–33 (2015)
5. Avasare, M.G.; Kelkar, V.V: Image encryption using chaos theory. In: Proc. Int. Conf. Communication, Information & Computing Technology (ICCICT). Mumbai, pp.1–6 (2015)
6. Paliwal, A.; Mohindroo, B.; Suneja, K.: Hardware design of image encryption and decryption using CORDIC based chaotic generator. In: Proc. Int. Conf. Recent advances and innovations in engineering (ICRAIE). Jaipur, pp.1–5 (2020)
7. Giakoumis, A.; Volos, C.K.; Munoz-Pacheco, J.M.; Gomez-Pavon, L.C.; Stouboulos, I.N.; Kyprianidis, I.M.: Text encryption device based on a chaotic random bit generator. In: IEEE 9th Latin Amer. Symp. Circuits & Systems (LASCAS). Puerto Vallarta, pp. 1–5 (2018)
8. Negi, A.; Saxena, D.; Suneja, K.: High level synthesis of chaos based text encryption using modified hill cipher algorithm. In: Proc. Int. Conf. India Council (INDICON). New Delhi, pp.1–5 (2020)
9. Shum, D.T.F.; Lee, R.S.T.; Liu, J.N.K.: Chaotic weatherman: the design and implementation of a chaotic weather prediction system. In: IEEE Congr. Evolutionary Computation (IEEE World Congress on Computational Intelligence). Hong Kong, pp. 2186–2193 (2008)
10. Fathima, T.A.; Nedumpozhimana, V; Lee, Y.H.; Winkler, S; Dev, S: A chaotic approach on solar irradiance forecasting. Photonics and Electromagnetics Research Symposium - Fall (PIERS - Fall). Xiamen China, pp. 2724–2728 (2019).
11. Lorenz, E.N.: Deterministic non-periodic flows. J. Atmos. Sci. **20**, 130–141 (1963)
12. Sprott, J.C.: Some simple chaotic flows. Phys. Rev. E Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top. **50**(2), R647–R650 (1994). https://doi.org/10.1103/physreve.50.r647
13. Pehlivan, I.; Uyaroğlu, Y.: A new chaotic attractor from general Lorenz system family and its electronic experimental implementation. Turk. J. Electr. Eng. Comput. Sci. **18**(2), 171–184 (2010). https://doi.org/10.3906/elk-0906-67
14. Chua, L.O.: Lin, G: Canonical realization of Chua's circuit family. IEEE Trans. Circuits Syst. **37**(7), 885–902 (1990)
15. Rössler, O.E.: An equation for continuous chaos. Phys. Lett. A. **57**(5), 397–398 (1976)
16. Cruz, J.M.; Chua, L.O (1993) An IC chip of Chua's circuit. IEEE Trans. Circuits Syst. II: Analog Digit. Signal Process. **40**(10), 614–625.
17. Cuomo, K.M.; Oppenheim, A.V.; Strogatz, S.H.: Synchronization of Lorenz-based chaotic circuits with applications to communications. IEEE Trans. on Circuits and Systems II: Analog and Digital Signal Processing.**40**(10), 626–633 (1993). https://doi.org/10.1109/82.246163.
18. Ibrahim, K.M.; Jamal, R.K.; Ali, F.H.: Chaotic behaviour of the Rossler model and its analysis by using bifurcations of limit cycles and chaotic attractors. J. Phys. Conf. Ser. **1003**, 012099 (2018). https://doi.org/10.1088/1742-6596/1003/1/012099
19. Bao, B.; Wang, N.; Chen, M.; Xu, Q.; Wang, J.: Inductor-free simplified Chua's circuit only using two op-amp based realization. Nonlinear Dyn. **84**, 511–525 (2016). https://doi.org/10.1007/s11071-015-2503-5
20. Gopakumar, K.; Premlet, B.; Gopchandran, K.G.: Implementation of Chua's circuit using simulated inductance. Int. J. of Electronics. **98**(5), 667–677 (2011)
21. Cuautle, E.T.; Hernndez, A.G.: Implementation of a chaotic oscillator by designing chua's diode with CMOS CFOAs. Analog Integr. Circuits and Signal Process. **48**, 159–162 (2006)
22. Jothimurugan, R.; Suresh, K; Ezhilarasu, P.M.; Thamilmaran, K: Improved realization of canonical Chua's circuit with synthetic inductor using current feedback operational amplifiers. Int. J Electron Commun (AEÜ).**68**(5), 413–421 (2014).
23. Trejo-Guerra, R.; Tlelo-Cuautle, E.; Cruz-Hernández, C.; Sanchez-Lopez, C.: Chaotic communication system using Chua's oscillators realized with CCII+s. Int. J. Bifurc. Chaos. **19**, 4217–4226 (2009)

24. Choubey, C.K.; Paul, S.K.: Implementation of chaotic oscillator by designing a simple Chua's diode using a single VDTA. Int. J. Electron. Commun. (2020). https://doi.org/10.1016/j.aeue.2020.153360
25. Kushwaha, A..K..; Paul, S.K.: Chua's oscillator using operational transresistance amplifier. Revue Roumaine des Sciences Techniques- Serie Electrotechnioque et Energetique. **6**, 299–303 (2016)
26. Kushwaha, A.; Paul, S: Inductorless realization of Chua's oscillator using DVCCTA. Analog Integrated Circuits and Signal Processing. **88** (2016). https://doi.org/10.1007/s10470-016-0746-9.
27. Deniz, H.I.; Cam Taskiran, Z.G.; Sedef, H: An analog chaotic lorenz circuit based on CCII+ and Multiplier. In: 41st Int. Conf. Telecommunications and Signal Processing (TSP). Athens; pp.1–5 (2018).
28. Elwakil, A.S.; Soliman, A.M.: Current conveyor chaos generators. IEEE Trans. Circuits Syst. I Fundam. Theory Appl.; **46**(3), 393–398 (1999)
29. Çiçek, S; Ferikoğlu, A; Pehlivan, I: Electronic circuit design of Sprott H chaotic system with CCII+. In: 22nd Signal Processing and Communications Applications Conference (SIU). Trabzon. pp. 2015–2018 (2014).
30. Itoh, Y; Adachi,M: Reconstruction of bifurcation diagrams using time-series data generated by electronic circuits of the Rössler equations. In: International Symposium on Nonlinear Theory and Its Applications (NOLTA) (IEICE).pp. 439–442 (2017).
31. Honglad, S; San-Um, W :Automatic stand-alone liquid mixer with chaotic PWM control using diode-based Rössler system. In: International Electrical Engineering Congress (iEECON). Thailand.pp.1–4 (2014).
32. Larptwee, S; San-Um, W: Implementation of Rössler chaotic system through inherent exponential nonlinearity of a diode with two-channel chaotic synchronization applications. In: Fourth International Conference on Intelligent Control and Information Processing (ICICIP). Beijing, China. pp. 787–791 (2013). https://doi.org/10.1109/ICICIP.2013.6568179.
33. https://www.analog.com/en/products/ad844.html#product-overview
34. https://www.analog.com/en/products/ad633.html#product-overview
35. Carroll, T.L.; Pecora, L.M.; Rachford, F.J.: Chaotic transients and multiple attractors in spin-wave experiments. Phys. Rev. Lett. **59**(25), 2891–2894 (1987). https://doi.org/10.1103/PhysRevLett.59.2891
36. Bai, E.W.; Lonngren, K.E.: Synchronization of two Lorenz system using active control. Phy. Rev. Lett. **8**(1), 51–58 (1997)
37. Tan, X.; Zhang, J.; Yang, Y.: Synchronizing chaotic systems using backstepping design. Chaos Solitrons Fractals. **16**, 37–45 (2003)
38. Vaidyanathan, S.; Rasappan, S.: Global chaos synchronization of n-scroll Chua circuit and Lur'e system using backstepping control design with recursive feedback. Arab J Sci Eng. **39**, 3351–3364 (2014). https://doi.org/10.1007/s13369-013-0929-y
39. Akar, M.; Ozguner, U.: Decentralized sliding mode control design using overlapping decompositions. Automatica **38**(10), 1713–1718 (2002)
40. Maghraoui, N; Belghith, S: Synchronization of two Rossler chaotic systems via adaptive backstepping with modular design. In: 1st Int. Symp. Control, Communications and Signal Processing. Hammamet, Tunisia. Pp. 131–134 (2004)
41. Haris, M.; Shafiq, M.; Ahmad, I., et al.: A nonlinear adaptive controller for the synchronization of unknown identical chaotic systems. Arab. J. Sci. Eng. (2021). https://doi.org/10.1007/s13369-020-05222-x
42. Sundarapandian, V.; Pehlivan, I.: Analysis, control, synchronization, and circuit design of a novel chaotic system. Math. Comput. Model. **55**(7–8), 1904–1915 (2012)
43. Gosar, Z: Chaotic dynamics - Rössler system. Datoteka: seminar_kaoticna_dinamika_Rossler_koncna, Zadnja spremememba. https://doi.org/10.13140/RG.2.2.27897.62564.
44. Pecora, L.M.; Carroll, T.L.: Synchronization of chaotic systems. Chaos Interdiscip J. Nonlinear Sci (2015). https://doi.org/10.1063/1.4917383
45. Liu, J; Zhang, Y: The application of Chaotic masking and chaotic switching in communication. In: Second International Conference on Mechanic Automation and Control Engineering. Hohho. p. 7781–7784 (2011)
46. Mobini, M.; Zahabi, M.: Masking communication using Sprott94 case a chaotic system in AWGN channel. J. World. Elect. Eng. Technol. **7**(2), 09–16 (2018)
47. Aromataris, G.; Annovazzi-Lodi, V.: Enhancing privacy of chaotic communications by double masking. IEEE J. Quantum Electron. **49**(11), 955–959 (2013). https://doi.org/10.1109/JQE.2013.2283584

Hindawi
International Journal of Photoenergy
Volume 2021, Article ID 4517923, 18 pages
https://doi.org/10.1155/2021/4517923

Research Article

Thermal Performance Improvement of Solar Parabolic Dish System Using Modified Spiral Coil Tubular Receiver

Rajkumar Malviya,[1] **Prashant V. Baredar**,[1] **and Anil Kumar**[2,3]

[1]*Energy Centre, Maulana Azad National Institute of Technology, Bhopal, 462003 M.P., India*
[2]*Department of Mechanical Engineering, Delhi Technological University, Delhi 110042, India*
[3]*Centre for Energy and Environment, Delhi Technological University, Delhi 110042, India*

Correspondence should be addressed to Rajkumar Malviya; er.rajkmalviya@gmail.com

Received 5 October 2021; Accepted 17 November 2021; Published 30 December 2021

Academic Editor: Alberto Álvarez-Gallegos

Copyright © 2021 Rajkumar Malviya 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.

The present research intends to design an efficient receiver for solar thermal applications with a solar dish concentrator system. Thermal and dynamic analysis is carried out for different convolutions of a spiral coil, and experiments are performed for testing the modified absorber. Experimental results are validated for the spiral absorber with numerical results. Three receivers of different numbers of convolutions are analyzed, and simulation steps are performed for these receivers to make improvements in the system efficiency. Finally, 5 convolutions of a spiral coil tubular absorber are taken for the modified design of the system. Absorber position for every spiral convolution is kept at the focus of the concentrated solar dish collector to achieve maximum efficiency. Material used for the reflective surface is anodized aluminum and copper for the absorber. The diameter of the aperture for the parabolic dish collector is 1.4 m. The maximum absorber temperature for May month comes out to be 296°C, and the maximum working fluid outlet temperature is found to be 294.2°C which is near to simulating temperature of 289.59°C and 288.15°C, respectively. This innovative design of the absorber consists of a feature of a 5 mm extension to the spiral tube at the exit and entry; hence, the turbulence effect could be overcome. Experimental thermal efficiency was found the highest (i.e., $\eta_{th}(\max) = 75.98\%$) for May. This work emphasizes on improving thermal performance by obtaining optimum absorber size using convolution strategy. Investigation of 5 convolutions of spiral coil tubular absorber with extended ends for obtaining optimum performance than existing work is the superiority of this work.

1. Introduction

Solar thermal receiver (i.e., solar absorber) is the vital component in solar dish concentrator technologies. It behaves as a connecting linkage for the insolation and the heat transfer working fluid (HTWF) flowing in the absorber. HTWF transfers the heat to the system that would be further transformed for a usable form of energy. The three most important points related to thermal absorber have to keep in mind. First, the power of absorbing the solar radiant heat with minimum heat losses is the most important characteristic of a receiver. Second, the optical properties and the structural parameters related to the thermal absorber. Third, the role of conductive, convective, and radiation heat losses from the concentrator and receiver system. To discover a cost-effective receiver for a solar thermal PD system with improved performance is great anxiety for researchers.

There are various geometries of receivers, i.e., cylindrical, conical, and spherical. Apart from it, many other geometries are also there, which can be studied. Spiral coil absorber, cylindrical tube receiver, modified cavity monotube boiler receiver, conical cavity spiral receiver, cylindrical shape spiral coil receiver, all-glass evacuated tube receiver, etc., are the most used receivers nowadays. Many researchers work on the heat losses and the different shapes of the thermal absorber [1]. Scientists always underestimate the role of absorber shape in reducing heat losses. Absorber size should be as small as possible for reducing heat losses in a thermal system. Daabo et al. [2] compared optical and thermodynamic behavior of three receiver geometries, i.e., conical,

cylindrical, and spherical. The capacity of conical shape absorber to receive and absorb the highest useful energy among the three absorbers with the highest thermal efficiency ($\eta_{th} = 77.05\%$). The cylindrical shape receiver had the best flux distribution. Pavlovic et al. [3] determined thermal and optical efficiency of solar parabolic dish 34% and 35%, respectively, using stainless steel corrugated spiral conical coil absorber located inside aluminum housing, and they also compared the three different heat transfer working fluids. The geometric properties of the thermal absorber decide the optimum thermal performance of the SPDC system. Bellos et al. [4] compared the conical spiral and flat spiral design of solar thermal absorber. Due to the shape of the conical absorber and the minimum number of convolutions resulting in conical absorber, it has the minimum effective area. This describes the heat losses due to shape are minimum. Thus, the performance and efficiency of the conical design were the best [2, 5, 6]. Spiraled tube conical cavity receiver welded on a hollow copper cylinder was developed using a storage carrier, i.e., molten salt. A thermocouple was used with a cylindrical component to obtain the temperature at a particular point in time. Thermal performance in the range of 58% to 77% was obtained for optimum flow characteristics [7]. A lightweight and low-cost solar dish collector was innovated to optimize the volumetric flow rate and working fluid outlet temperature using a spiral coil absorber to achieve optimum exergetic and thermal performances of the design. The idea behind the spiral absorber geometry was to get evenly distributed heat flux which improves the heat transfer rates in the absorber [8]. This concept was proved numerically by keeping the volume of the receiver constant with changing aspect ratio [9, 10]. A modified U-shaped spiral air channel inside the solar thermal receiver was developed to minimize the convective heat losses up to 58% [11]. The conical absorber design has a very high performance which leads to higher absorbed solar energy [2, 3, 5–7]. Tian et al. [12] investigated the temperature range for seven working fluids to achieve the best and efficient performance level as mentioned in Table 1.

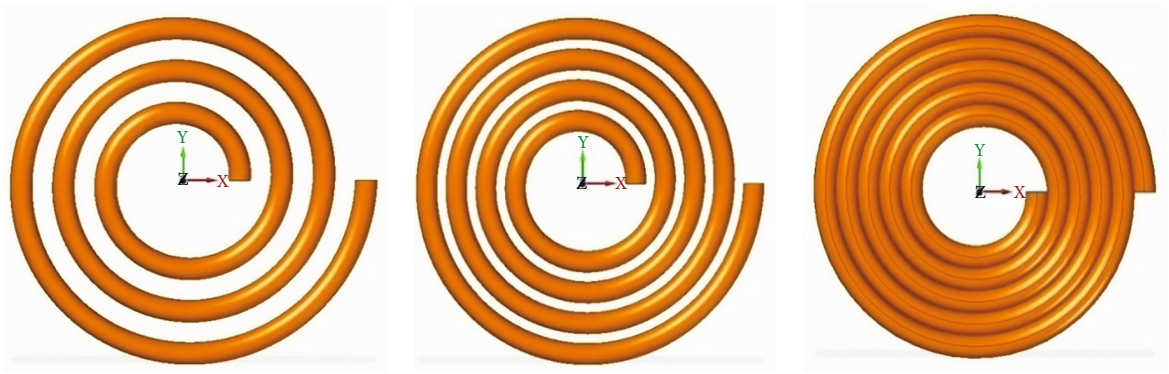
Order of thermal and exergetic efficiencies is as $\eta_{th/exe} = \eta_H > \eta_{Cu} > \eta_{Cy}$ (where η_H, η_{Cu}, and η_{Cy} are efficiencies for hemispherical, cubical, and cylindrical cavities, respectively) for inlet temperature of working fluid. The efficiency of cubical and cylindrical receivers was compared using low viscous oil working fluid, and thermal efficiencies were 65.14% and 56.44%, respectively [13]. Dry working fluids have the worst performance over wet working fluids. Water is the most suitable heat transfer working fluid at low temperatures, and at higher temperatures, therminolis the best [12–16]. Zou et al. [17] determined an enhancement in thermal and hydraulic performance using a tubular flat spiral ribbed receiver for PTC. Materials that have higher absorption and lower reflections are used in solar still construction to increase distilled water produced by the system [18]. Ali et al. [19] suggested the honeycomb receiver of alumina by using a reasonable surface layer of the thermal receiver material as the absorption of the material affects the optical and thermodynamic character of the absorber. Li et al. [20] investigated a collector of low concentrating solar thermal power by comparing the thermal efficiencies of darkening chrome-layered copper tubular absorber and multiwalled carbon nanotube receiver. Auti et al. [21] suggested stainless steel 312 for absorber material as its properties of fair corrosion resistance and of withstanding high temperatures. Although aluminum has the highest corrosion resistance, it is difficult for aluminum to bear high temperature. Satpute et al. [22] found an increment in HTC between HTWF and absorber at an increase in mass flow rate using an aluminum triangular absorber of aluminum in CPC. Impinging the receiver behaves more sensitive for high absorptive surface material. Total optical efficiency for cavity surface cylindrical receiver increases with increasing material absorptivity [23]. Wang et al. [24] examined four-receiver radius, $R = 23$ mm, 25 mm, 27.5 mm, and 30 mm. The smallest 23 mm receiver radius was chosen for a porous media receiver. However, little spillage occurred with this size of the receiver.

Table 1: Temperature range for working fluids [10].

Working fluid	Temperature level
Water and Therminol VP-1	Up to 550 K
Molten salt	Up to 870 K
Liquid sodium	Up to 1100 K
Air, CO_2, and He	Greater than 1100 K

Pavlovic et al. [25] found the perfect location of the receiver was 2.1 m from the concentrator surface smaller than focal distance, 2.26 m, to optimize the optical performance and to devise evenly distributed heat flux for the receiver. Figaj et al. [26] analyzed a hybrid system of solar chiller for the cooling unit by using an aluminum heat receiver. Zou et al. [27] investigated a novel design of a cylindrical hollow chamber receiver. The effect of hollow chamber length and the effective area of aperture based on the number of loops were studied for obtaining maximum efficiency.

In line with the literature, most of the researchers have been worked on absorbers' different geometries, absorbers' materials, and working fluid flowing inside the absorber, intending to optimize the performance by improving the SPDC system's efficiency. Considering all the literature review, it is the first time to design such an innovative absorber using the "convolution strategy." To design and analyze the thermal absorber in this way is to introduce an absorber having 5 convolutions of the copper tube by performing simulation for 3 convolutions and then 4 convolutions. This innovative design of the absorber consists of a novel feature of a 5 mm extension to the spiral tube at the exit and entry; hence, the turbulence effect could be overcome. Contact surface area of receiver/unit volume of HTWF is more for a tubular flat spiral receiver when compared to the conventional short cylinder receiver, which acts as a mechanism to enhance heat transfer. No gap in consecutive turns of coil and even distribution of heat flux on flat spiral coil tubular absorber reduce heat losses [28].



(a) (b) (c)

FIGURE 1: Spiral coil tubular copper absorber modeled on Solid Edge ST9 software as (a) 3 convolutions, (b) 4 convolutions, and (c) 5 convolutions.

It optimizes the thermal performance of the SPDC system [29]. Solid Edge ST9 software is used for constructing the optimized absorbers' geometry. The velocity of working fluid inside the absorber, steady-state thermal analysis, and pressure drop behavior are analyzed with the help of ANSYS 16.1 software. Dynamic behavior of the working fluid is analyzed with the object of using the kinetic energy of the working fluid in the operating mechanisms of various applications. Extension of the copper tube at entry and exit level minimizes the turbulence effect to prevent damages to the mechanism. Tube surface and working fluid temperature, working fluid velocity, pressure drop, heat flux distribution, and directional heat flux distribution for different convolutions are analyzed to design this absorber. After designing the absorber numerically, the construction of the copper tube spiral absorber is done. Testing results are used to validate the simulating results. Finally, a light weight and compact coiled tubular receiver in flat spiral form for a parabolic dish concentrator of 1.4 m is developed and tested.

2. Thermal Performance Indicators

Parametric definitions and energy balance equations used for developing an optimized thermal absorber are discussed in this section.

2.1. Concentration Ratio and Shape of Receiver. It is basically defined as the amount of solar radiant rays accumulated by a particular concentrator. There are 2 differently defined concentration ratios in use as follows:

(1) Optical Solar Concentration Ratio (CR_{opts}). It is the ratio of the incoming beam radiant energy on the effective reflective surface area (I_{sur}) to radiant energy that reaches the receiver/absorber (I_{rec}) [2, 30]:

$$CR_{opts} = \frac{I_{sur}}{I_{rec}} \quad (1)$$

(2) Geometric Concentration Ratio (CR). The ratio between the effective concentrator aperture (A_a) area and the effective receiver/absorber area (A_{rec}) [2, 5, 30–33]:

$$CR = \frac{A_a}{A_{rec}} \quad (2)$$

Calculations of the thermal efficiency related to the thermal absorber are done by using this geometric concentration ratio.

Shape of spiral coil tubular receiver is based on the following equation (3), as [34]

$$R = R_0 e^{a\theta}. \quad (3)$$

2.2. Performance Calculations for the SPDC System. The efficiency of the SPDC system is determined by dividing the useful heat energy generated (Q_u) to the accessible solar irradiation heat energy on effective PD aperture surface (Q_s), as [3–5, 30, 35–37]

$$\eta_{th} = \frac{Q_u}{Q_s}. \quad (4)$$

Balancing equation in the HTWF volume is given as [3–5, 28, 35, 37–40]

$$Q_u = m \cdot C_p \cdot \left(Tf_{(out)} - T_{f(in)}\right), \quad (5)$$

where

$$m = \frac{\rho \cdot \nu}{3600}\ (\text{kg/s}), \quad (6)$$

where ρ is in kg/L, ν is in L/h, and 3600 is in s/h.

And the available solar irradiation heat energy on effective PD aperture surface is determined as [3–5, 35]

$$Q_s = A_a \cdot G_b. \quad (7)$$

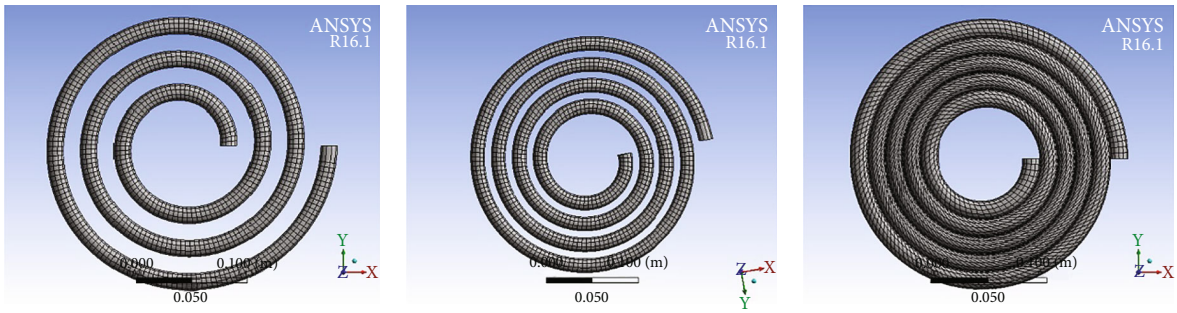


Figure 2: Spiral coil tubular copper absorber mesh on ANSYS 16.1 software as (a) 3 convolutions, (b) 4 convolutions, and (c) 5 convolutions.

Table 2: Description of mesh data.

Number of convolutions of coil	No. of elements	No. of nodes	ANSYS modules
3 convolutions	34704 & 6678	42716 & 33034	
4 convolutions	42372 & 8010	52056 & 39670	ANSYS CFX and steady-state thermal analysis
5 convolutions	41080 & 11592	52428 & 56530	

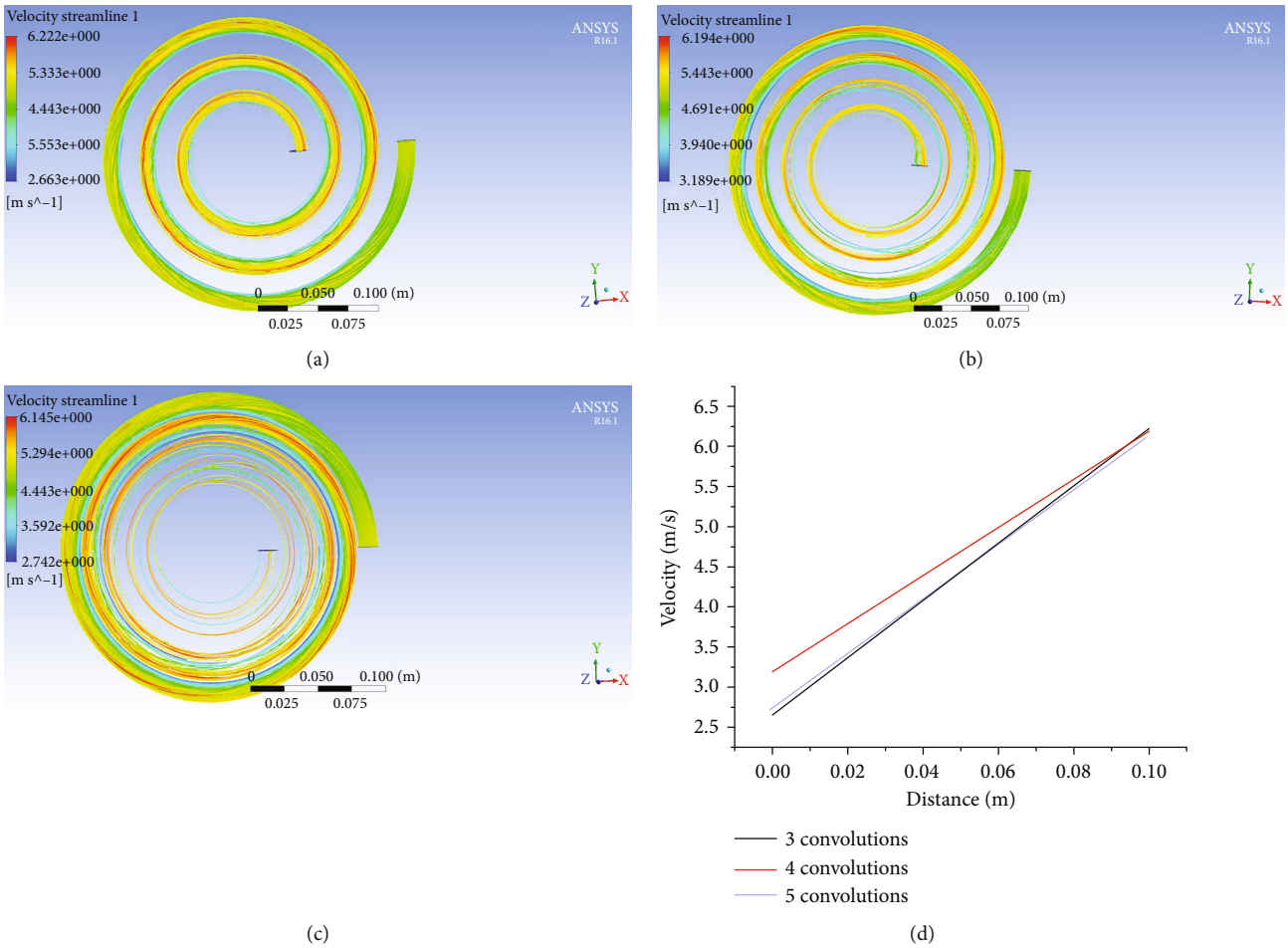


Figure 3: Streamlines of flow through coil of (a) three convolutions, (b) four convolutions, (c) five convolutions, and (d) velocities of streamlines at different positions of coil with varying number of convolutions.

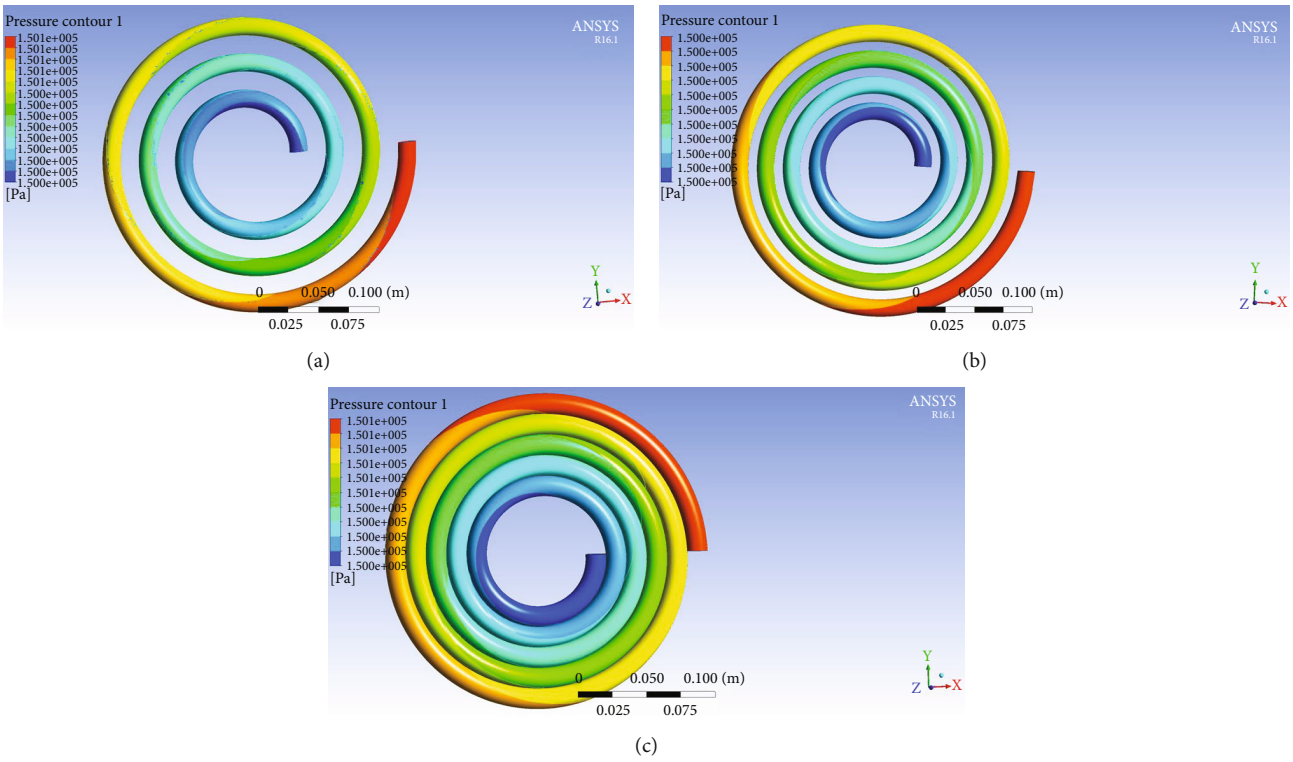


FIGURE 4: Pressure contours for coil of (a) three convolutions, (b) four convolutions, and (c) five convolutions.

However, some of the radiations do not concentrate on the absorber due to design constraints. These account for the radiation losses.

2.3. Energy Balance Equations for Spiral Coil Tubular Absorber. The absorbed solar energy (Q_{abs}) is given as [3, 5, 41, 42]

$$Q_{abs} = Q_u + Q_{loss}, \tag{8}$$

where [36]

$$Q_{loss} = Q_{rad} + Q_{conv,in} + Q_{conv,ex} \tag{9}$$

Also, the absorbed energy may be defined as [2–5, 28, 30, 40]

$$Q_{abs} = \eta_{opts} \cdot Q_s, \tag{10}$$

where

$$\eta_{opts} = \text{Reflectance} \cdot \text{Absorbance} \cdot \text{Intercept Factor}. \tag{11}$$

Intercept factor ("γ") is calculated by determining the actual amount of energy concentrated on the absorber and the radiation losses caused by design constraints of PD as [3]

$$\gamma = \frac{\text{Energy intercepted by the spiral coil tubular receiver}}{\text{Energy reflected by the parabolic dish collecter}}. \tag{12}$$

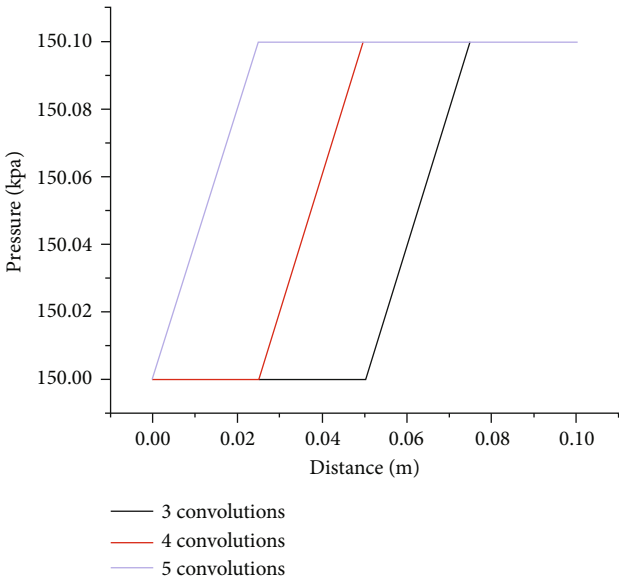


FIGURE 5: Pressure at different positions of coil with varying number of convolutions.

2.4. Heat Transfer Equations for Spiral Coil Tubular Absorber. The energy transfer from the receiver to the HTWF is given as [3, 4]

$$Q_u = A_{rec} \cdot h_f \cdot (T_{rec} - T_{fm}), \tag{13}$$

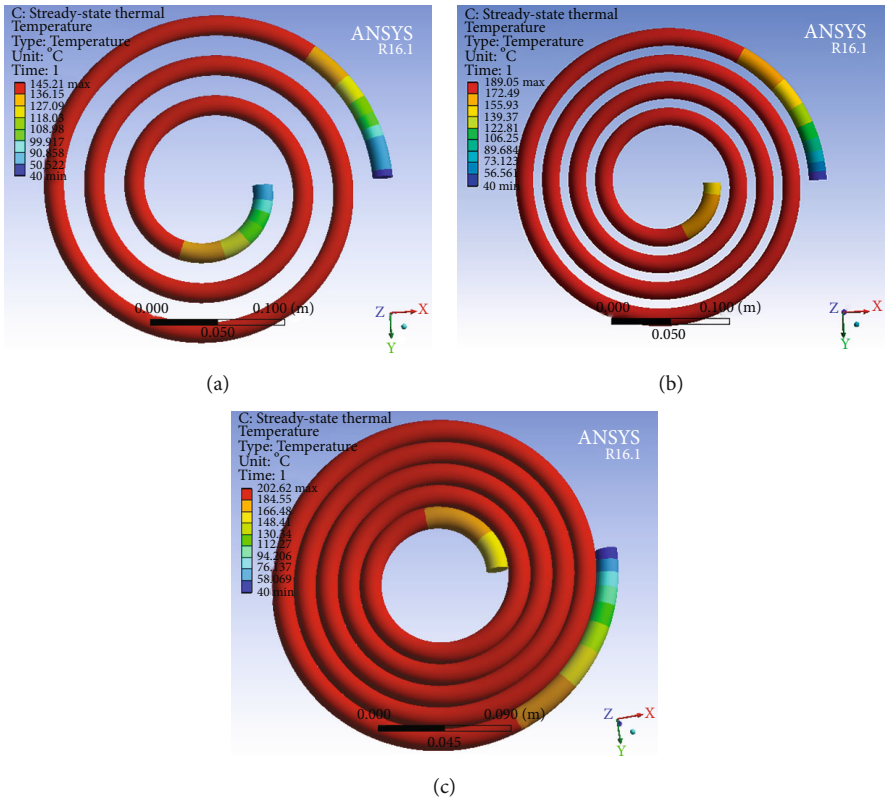


Figure 6: Temperature distribution through (a) three convolutions of coil, (b) four convolutions of coil, and (c) five convolutions of coil.

where

$$h_f = \left(\frac{d_{rec}}{k_f}\right) \cdot \frac{(f_r/8) \cdot R_e \cdot P_r}{1 + 12.8 \cdot \sqrt{f_r/8} \cdot (P_r^{0.68} - 1)}. \quad (14)$$

2.5. Characteristics of Fluid Flow inside the Spiral Coil Tubular Receiver. On account of curvature of the coil centrifugal force is found. With this force, secondary flow is set up owing to the curvature nature of the spiral coil tube. This secondary flow has the characteristic of escalating the mixing of fluid particles rapidly. Swirling motion within the fluid due to mixing activity enhances the heat and mass transfer as well as fluid transfer across cross section. This concept is used to shift HTF from inlet boundary to outlet boundary with mean flow velocity (v_{θ_i}) at cross section. M_x is the shifting parameter defined by equation (15), as [34]

$$M_x = \sum_{i=0}^{n} \left\{ \frac{1}{r_i} \cdot \left(\frac{\Delta A_i v_{\theta_i} x_i}{\Delta A_i v_{\theta_i}} \right) \right\}. \quad (15)$$

The friction factor f is used to determine the consequence of pressure drop in the spiral coil tubular absorber [3–5, 34]

$$f = \frac{-(\Delta P/L) d_h}{(1/2)\rho v_\theta^2}. \quad (16)$$

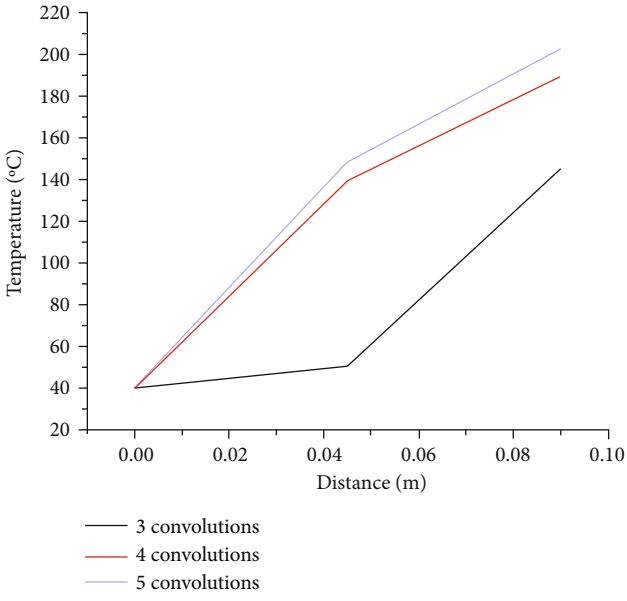
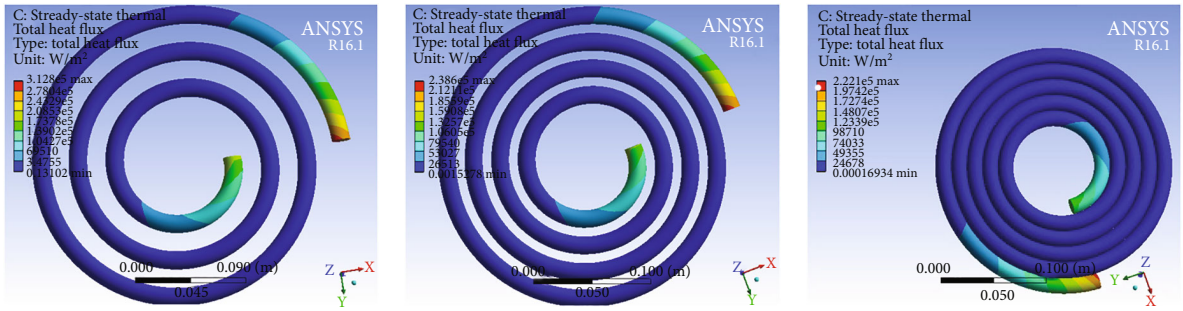


Figure 7: Temperature values at different positions of coil with varying number of convolutions.

Modified critical Reynolds' number for turbulence flow can be calculated by equation (17), as [4, 34]

$$R_{e,cr} = 2300 \left[1 + 8.6 \left(\frac{d_i}{2R} \right)^{0.45} \right]. \quad (17)$$



(a) (b) (c)

FIGURE 8: Heat flux distribution through (a) three convolutions, (b) four convolutions, and (c) five convolutions of coil.

Range of R_e depends on the mode of the application and type of fluid flow. The main equations are described as follows [34, 43]:

General continuity equation

$$\frac{\partial}{\partial x_i}(\rho u_i) + \frac{\partial}{\partial y_j}(\rho v_j) + \frac{\partial}{\partial z_k}(\rho w_k) + \frac{\partial}{\partial t}(\rho) = 0. \quad (18)$$

Momentum equation

$$\frac{\partial}{\partial x_i}(\rho u_i u_j) = -\frac{\partial p}{\partial x_i} + \frac{\partial}{\partial x_i}\left[\mu\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right)\right]. \quad (19)$$

The three applied boundary constraints for this flow field are described as follows:

(i) Flow rate of mass at entry is kept constant

(ii) The exit condition is the second type (Neumann)

(iii) At the peripheral walls, there are no slip conditions

3. Numerical Simulation of Spiral Coil Tubular Absorber

Graphical contours and plots are obtained for velocity, pressure, temperature, heat flux, and directional heat flux by ANSYS CFX, ANSYS Fluent, and steady-state thermal analysis. Model is developed in Solid Edge ST9 software. After transporting the model geometry in ANSYS 16.1, meshing is performed. The final solution is obtained by defining the boundary. In the entry of water, the flow mass rate and the uniform temperature were selected. After this, the static pressure in the outlet of the tube was set to be environmental. The last boundary condition is the heat convection between the coil outer surface and the environment. It is important to say that for determining a different operating condition, the water inlet temperature was changed in the proper boundary condition. After this step, the radiation surfaces were selected. The reflector was set to be a symmetrical surface in order to reflect the sun rays.

A copper absorber of spiral coil shape is to be modeled on Solid Edge ST9 software. For efficient and optimal performance, three coil models are modeled for 3, 4, and 5 convolutions, respectively. These are depicted in Figure 1.

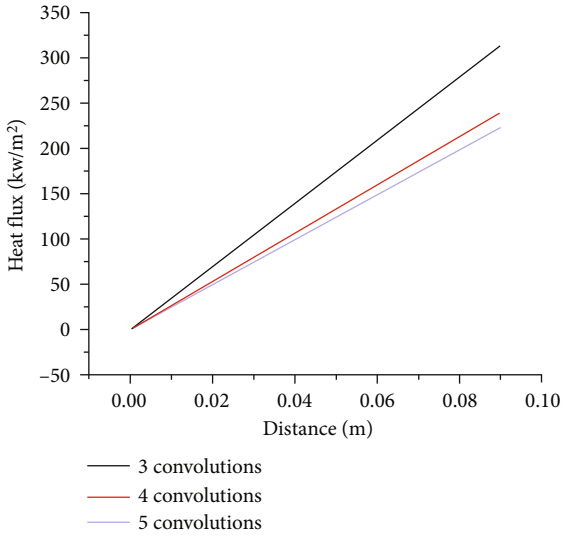


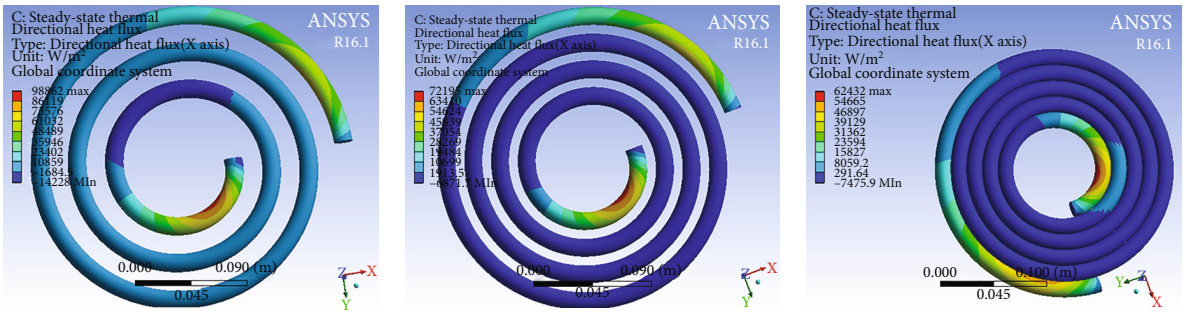
FIGURE 9: Heat flux at different positions of coil with varying number of convolutions.

The length of the coil tube is variable, and the inner and outer diameters of these coils are kept constant. The gap of the coils, as shown in Figures 1(a) and 1(b) is responsible for energy (heat) loss, and Figure 1(c) is depicted for no loss of energy (heat) due to gap because in this coil, there is no gap between the convolutions.

After creating a geometric model, the geometry of the model has been divided into small elements for flow analysis; this process of making small elements is called meshing, as shown in Figure 2. This step is followed by geometry construction for performing detailed and microscopic analysis for more detail. The process involves breaking up the geometry into nodes and elements for a detailed analysis of the entire geometry. For part meshing, different mesh size has been given to different individual parts.

Mesh data for different parts and their different conditions has been given in Table 2.

3.1. Velocity Distribution for Absorber. Obtained contour for spiral coil absorber for three, four, and five convolutions are



(a) (b) (c)

Figure 10: Directional heat flux distribution for (a) three convolutions, (b) four convolutions, and (c) five convolutions of coil.

depicted in Figure 3. The fewer convolution coils have more gaps, and more convolution coils have less gaps. After performing a practical approach, five convolution coil gives the best result among all. The flow of working fluid is streamlined and smooth. From these contours, it is clear that the velocities of streamlines at the outlet of coil increases with an increasing number of convolutions which in turn will operate efficiently the CSP-based irrigation water pump and/or other mechanisms/devices. But, due to consideration of the variable (increasing) length of the tube, velocity decreases as an increasing number of convolutions. Hence, finally, it gives nearly equal velocity at outlet in all three cases of absorber coil as depicted in Figure 3(d).

There is turbulence at outlet in actual working condition, which can damage the mechanism/system due to the sudden impact of pressurized steam; therefore, a 5 mm extension is given both at inlet and outlet of the coil.

3.2. Pressure Distribution for Absorber. Outlet pressure of coil should be matched with the inlet condition of the inlet steam chamber of the system, which is found in the simulation is 150 kPa in the condition of 100% inlet opening as depicted in Figure 4.

Suppose steam pressure changes its values inside the tube. In that case, there may be possibilities of impacts, and due to these impacts and jerks, the coil tube may deform, which in turn is responsible for increasing the pump's cost because copper coil may deform quickly. To overcome this drawback, steam pressure inside the coil tube must be constant through the journey of working fluid inside it.

It is clear that the pressure at different positions of coil with varying number of convolutions is almost constant throughout the coil as shown in Figure 5.

3.3. Temperature Distribution for Absorber. As the heat loss decreases, temperature increases on an increasing number of convolutions of absorber coil because the gap decreases as depicted in Figures 6(a)–6(c). The coil shape is spiral, resulting in even distribution of the temperature and heat, and getting that desire results in simulation of the absorber coil. The chances of heat loss are very much reduced compared to other shapes of the absorber. The five convolution coil is closely packed, and it has no space between nearby convolutions.

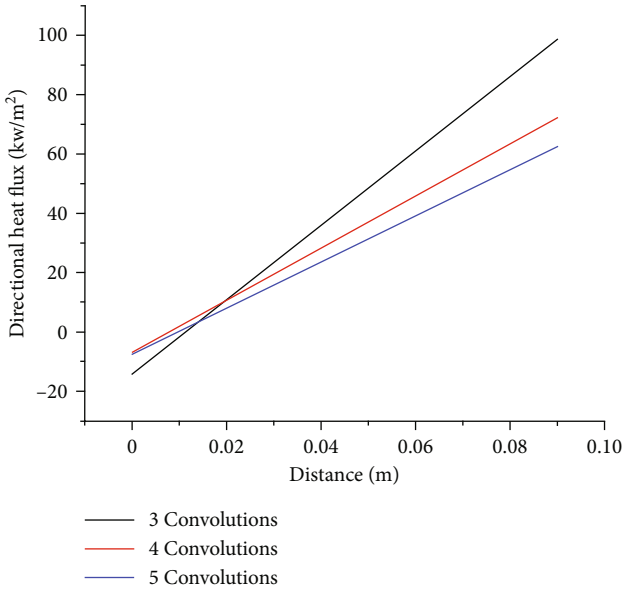


Figure 11: Directional heat flux at different positions of coil with varying number of convolutions.

It is clear that temperature values at the outlet of the coil increase with an increasing number of turns and will be sufficient to convert water into steam of very high temperature as shown in Figure 7.

3.4. Heat Flux Distribution. As the heat absorbing area increases, the heat flux decreases, and the following results of simulation for heat flux distribution are also favoured the same as shown in Figures 8(a)–8(c).

Simulated contours of heat flux are depicted in Figure 9; heat flux at the coil outlet decreases with an increasing number of the coil.

3.5. Directional Heat Flux Distribution. As the heat absorbing area increases, the directional heat flux decreases, and the following results of simulation for directional heat flux distribution are also favoured the same as shown in Figures 10(a)–10(c).

Simulated contours of directional heat flux are depicted in Figure 11; directional heat flux at the coil outlet decreases with an increasing number of convolutions of coil.

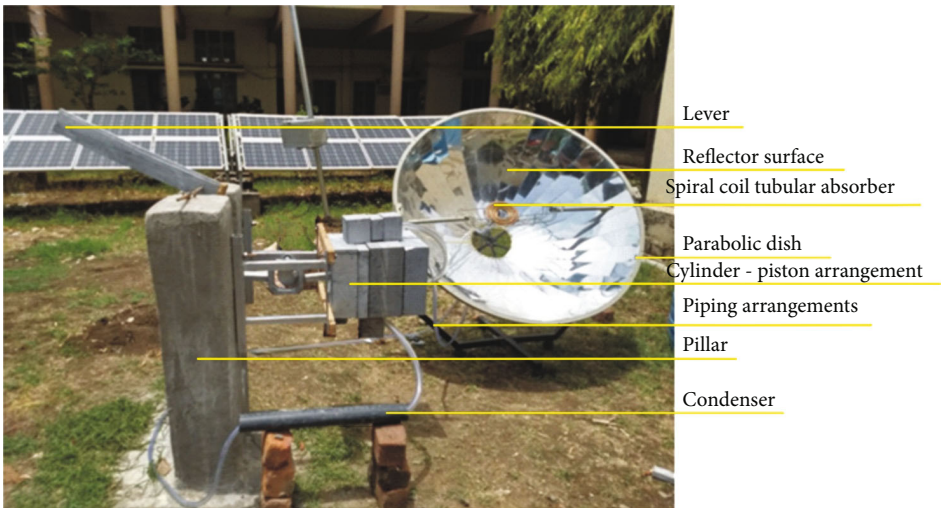


Figure 12: Modified solar thermal absorber installed in SPDC for solar water thermal pump system.

4. Experimental Setup and Instrumentation

4.1. Experimental Setup. The experimental setup has been installed on the lawn of the Energy Centre of Maulana Azad National Institute of Technology, Bhopal (India). According to NREL, Bhopal has 23.250 latitudes and 77.450 longitudes. Annual average direct normal irradiance is 5.61 kWh/m^2/day which is maximum in April and May, 7.27 kWh/m^2/day and 7.51 kWh/m^2/day, respectively.

The modified spiral coil tubular absorber is set on the focus point of the SPDC, as shown in Figure 12. This combination of PD and absorber may be used in various applications. Here, it is used for driving an irrigation water pump system using its kinetic energy and may also be used for heating and drying purposes. Pumped water enters into the absorber and takes heat through a copper tube.

The spiral coil tubular absorber is shown in Figure 13. Copper is used for making this spiral coil tubular absorber. The thermal conductivity of copper is 385 W/mK. Therefore, it is an excellent absorber of heat and may be used for heating and drying applications. Here, this absorber coil is used as a device to convert pressurized and pumped working fluid from the liquid stage to the vapor stage (steam). The coil has 5 convolutions and a 5 mm extension at entry and exit. The specifications of the spiral coil tubular absorber and its material properties are given in Table 3. The innovation of this system is the spiral absorber geometry which is an alternative way to create a low-cost system. This configuration is simple, and it helps to create a reactively uniform heat flux distribution.

The readings of measured and experimental parameters were taken on 21st January, 17th February, 23rd March, 25th April, and 12th May in the year 2020 for the duration from 8:30 A.M. to 6:30 P.M. for every hour. Then, mean was calculated for whole day readings of a particular month, and this mean value was considered for that month.

A schematic diagram of the experimental setup developed for performance tests is presented in Figure 14. The working principle of this pumping system can be explained



Figure 13: Modified spiral coil tubular thermal absorber.

Table 3: Design parameters, material properties, and dimensions of collector and modified absorber.

Description	Value (all dimensions in mm)
Diameter of PD	1400.0
Focal length of the PD	272.2
Depth of the PD	450.0
Rim angle	45°
Material of PD	Anodized aluminium
Collector reflectivity	0.82
Intercept factor	0.98
Length of spiral tube	2286.0
Diameter of spiral tube	10.0
Inner diameter of copper coil	76.2
Outer diameter of copper coil	152.4
Emittance value for absorber	0.11
Absorbance value for absorber	0.836
Tube material	Copper
Thermal conductivity of copper	385 W/mK

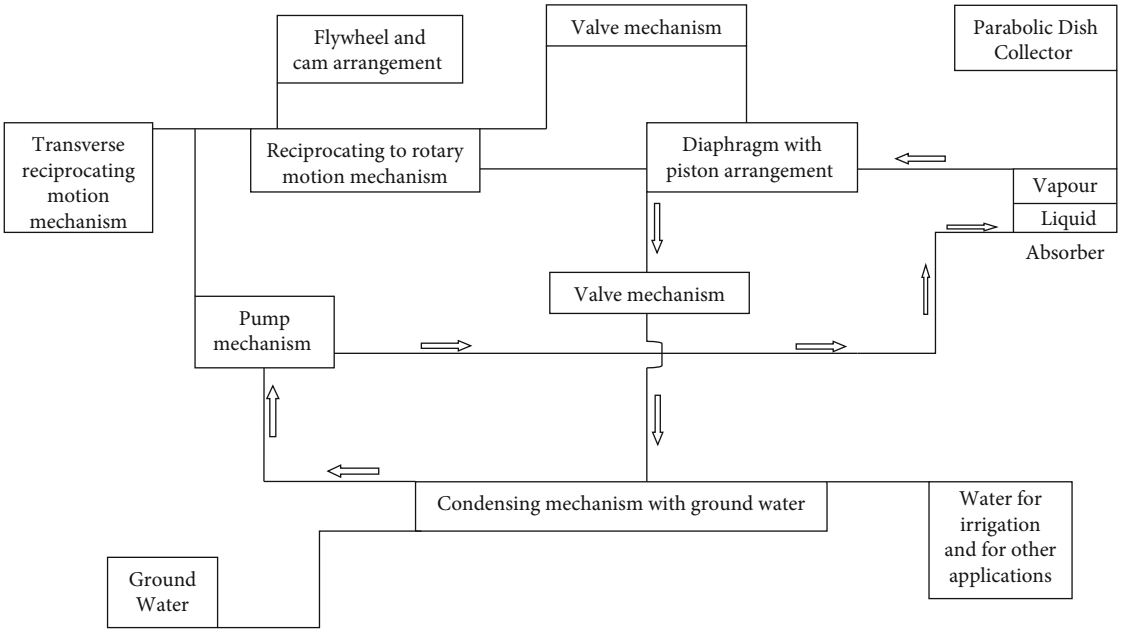


Figure 14: HTWF circulation layout along with water pumping mechanism.

Table 4: Output obtained from SPDC system using modified absorber.

Month	Measured parameters				Experimental			Numerical			Deviation		
	$T_{f(in)}$ (°C)	$V_{f(out)}$ (m/s)	ν (L/h)	G_b (kWh/m^2/day)	$T_{abs(max)}$ (°C)	$T_{f(out)}$ (°C)	η_{th} (%)	$T_{abs(max)}$ (°C)	$T_{f(out)}$ (°C)	η_{th} (%)	$T_{abs(max)}$	$T_{f(out)}$	η_{th} (%)
Jan	30	6.145	6948	10.9	220	218.6	57.93	202.62	199.32	54.21	0.079	0.088	6.42
Feb	34	6.8	7690	11.4	234	232.9	62.83	219.32	216.65	59.61	0.063	0.070	5.12
March	40	7.2	8140	11.8	258	256.4	66.94	244.12	242.13	64.16	0.053	0.056	4.15
April	42	7.3	8260	12.2	272	269.8	64.94	264.74	262.20	66.17	0.027	0.028	1.89
May	44	8.1	9160	12.7	296	294.2	75.98	289.59	288.15	74.86	0.022	0.021	1.47

clearly by considering a simple Rankine cycle. The working fluid is water which remains water for half cycle and converts into steam for the remaining half of the cycle [36]. Here, an arrangement of piston-cylinder and diaphragm is used to convert the steam enthalpy into the mechanical energy. In this work, the collector is taken as a parabolic solar dish concentrator along with a solar thermal spiral tubular coil absorber. The working fluid (water) is pressurized in pumping arrangement going through the collector and got heated up to around 150°C to 300°C depending on the available solar beam radiation at particular time interval. The flow of working fluid is also shown in the figure in which the Rankine cycle is shown with the help of arrows. Pressurized water enters first in the absorber coil. With the help of a solar concentrator, the sun rays are focused on the absorber coil and the temperature will increase. Range of temperature in absorber coil is 150°C-300°C. Now, water is converted into steam. Pressurized steam enters first into the pressure chamber, where steam pressure is to be measured. From then, it enters to valve-block arrangement. These valves are attached to the cam arrangement with the help of this; the valves of inlet and outlet open and close at the particular timing. When the valve comes in open condition, steam enters to the piston-cylinder and diaphragm-chamber arrangement. Pressurized steam pushes the diaphragm and piston outwards and exits by the outlet valve into the condenser. The piston is pushed by the pressurized steam in outward direction. With the help of this mechanism and due to inertia effect, the piston again comes towards the cylinder head and makes a complete rotation. This moves the lever arrangement and then a belt and rope are there for lifting the underground water. When the lever upper arm revolves, this helps to reciprocate the pumping arrangement and this creates the pressure of the water (working fluid) which is flowing into absorber coil. Then, this working fluid exits from the piston-cylinder arrangement and enters into the condenser. The working fluid gets cooled and then enters to pumping arrangement and then to absorber coil which completes the working cycle. The solar energy first gives its heat to water, this water is converted into steam, then pressurized steam runs the mechanism, and with the help of mechanical energy of this mechanism, pump is operated [29].

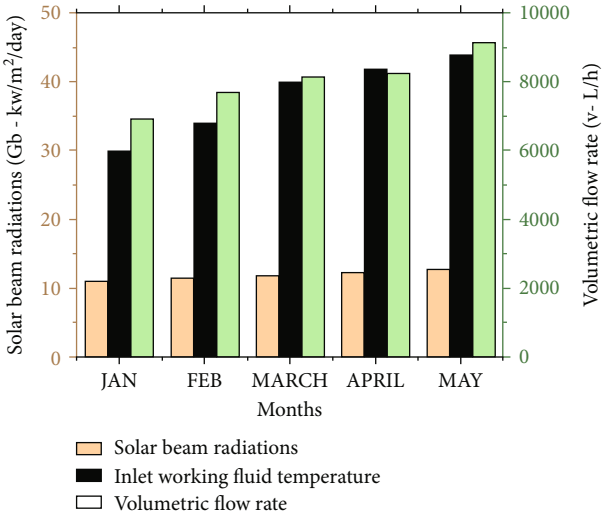


Figure 15: Measured data for solar beam radiations, inlet working fluid temperature, and volumetric flow rate from January 2020 to May 2020.

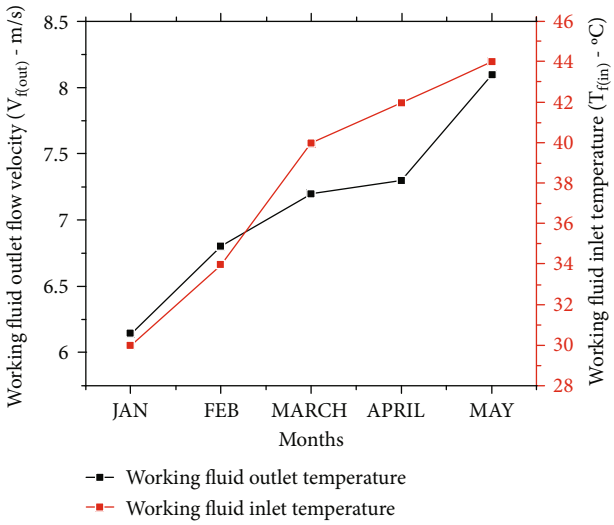


Figure 16: Behavior of working fluid outlet flow velocity and working fluid inlet temperature from January to May.

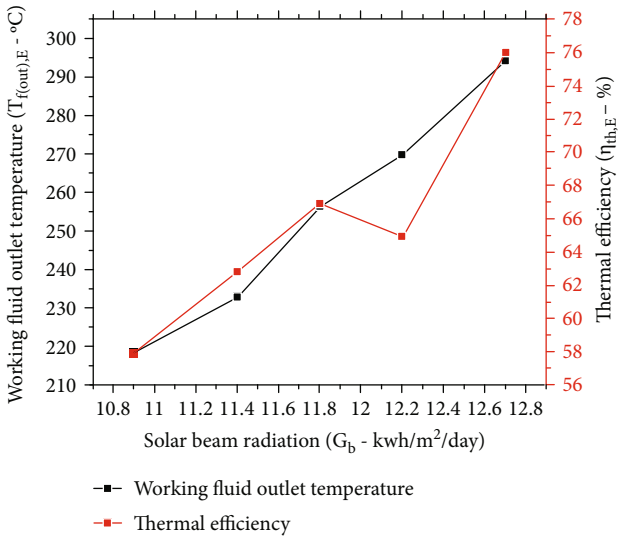


Figure 17: Behavior of working fluid outlet temperature and thermal efficiency with solar beam radiation.

4.2. Measuring Instruments. The main measuring instruments used during the experiments are as follows:

(1) Flow meter: the volumetric water flow rate (v) is measured with this flow meter of accuracy of 1%

(2) Pt100 thermometers: water inlet temperature (T_{in}), water outlet temperature (T_{out}), and ambient temperature (T_m) are measured with this thermometer with an accuracy of 0.1 K

(3) Pyranometers: the solar irradiation (global and diffuse) was measured with two pyranometers of 1% accuracy

5. Results and Discussion

Testing of modified spiral coil tubular absorber with solar thermal dish concentrator was performed from January 2020 to May 2020 at Energy Centre Lawn of Maulana Azad National Institute of Technology situated in the capital (i.e., Bhopal) of Madhya Pradesh state of India. The obtained output details, measured parameters, and deviation between experimental and numerical results are communicated in Table 4.

Figure 15 depicts the three measured parameters (solar beam radiations, inlet working fluid temperature, and volumetric flow rate) that increase continuously from January

Table 5: Thermal properties of working fluid at outlet temperature when $p = 150$ kPa.

$T_{f(out)}$ (°C)	ρ (kg/L)	m (kg/s)	C_p (kJ/kg K)	k_f (W/mK)	h_{ef} (kJ/kg)
218.6	0.0006653202	0.00133803	1.99379577	0.03522136135775	2910.2283315212
232.9	0.0006460501	0.00138003	1.99551452	0.036601677039559	2938.7499503425
256.4	0.0006167785	0.00139460	2.00132672	0.038944475744472	2985.7066147087
269.8	0.0006012796	0.00137960	2.00588169	0.040318305585457	3012.5540644076
294.2	0.0005750168	0.00146309	2.015843001	0.042884156799922	3061.615277978

to May. Solar beam radiation is the main responsible factor for this enhancement as the sun rays enter into earth surface; these are captured by solar PD for concentrating onto a modified absorber. The solar beam radiations reached to PD surface in April, and May has the highest values. Thus, the inlet working fluid temperature and volumetric flow rate also attain maximum values for these months.

A rapid increase in working fluid outlet flow velocity and working fluid inlet temperature was recorded from January to May, as shown in Figure 16. Performance of modified absorber for working fluid outlet flow velocity increases as the solar beam radiation values increases. And this outlet flow velocity affects the volumetric flow rate, which is measured through flow meter. Using the values of volumetric flow rate, mass flow rate was obtained for each reading.

The highest value of working fluid outlet temperature was obtained for May. It increases with the value of solar beam radiation, as shown in Figure 16. The performance of modified absorber during the whole period was optimum, but a noticeable change in the month of April was seen. At this particular point in time, the experimentally obtained thermal efficiency of the system falls to 64.94% from 66.94% in March. Except for this point of time, thermal efficiency increases as the solar beam radiation value increases as depicted in Figure 17.

This happened due to the inherent property of water (HTWF). At the working fluid temperature at the exit of the modified absorber for this month, the density of water decreased drastically and affected the mass flow rate of HTWF. It happened in May, but the mass flow rate for this point of time increased in order, as shown in Table 5 and Figure 18.

Figure 18 depicts that the numerically obtained thermal efficiency increases with an increase in mass flow rate every time. But experimentally obtained thermal efficiency of the system decreases in April and again increases for May.

Thermal efficiency changes in three stages as depicted by Figures 19 and 20. In the case of numerically obtained thermal efficiency, the first and last stages show that the values of the efficiency increase rapidly and the midstage between 11.8 kWh/m^2/day and 12.2 kWh/m^2/day (solar beam radiation) or between 8140 L/h and 8260 L/h (volumetric flow rate) thermal efficiency enhancement came slowly.

Experimentally obtained values of thermal efficiency increase rapidly in the first and the last stage. However, in the mid-mid-stage, between 11.8 kWh/m^2/day and 12.2 kWh/m^2/day (solar beam radiation) or between 8140 L/h and 8260 L/h

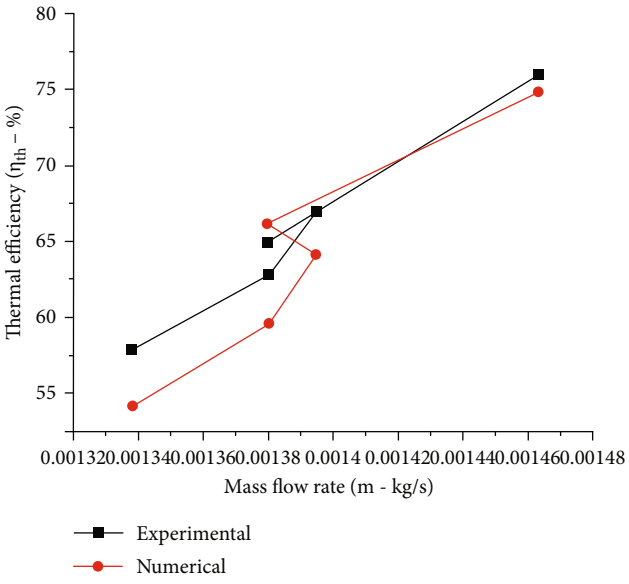


Figure 18: Comparative behavior of experimental and numerical thermal efficiency of the system with mass flow rate.

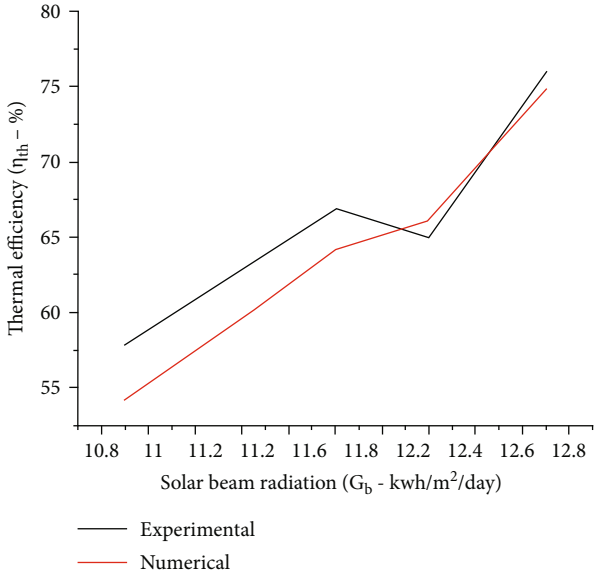


Figure 19: Changes in experimental and numerical values of thermal efficiency with solar beam radiation.

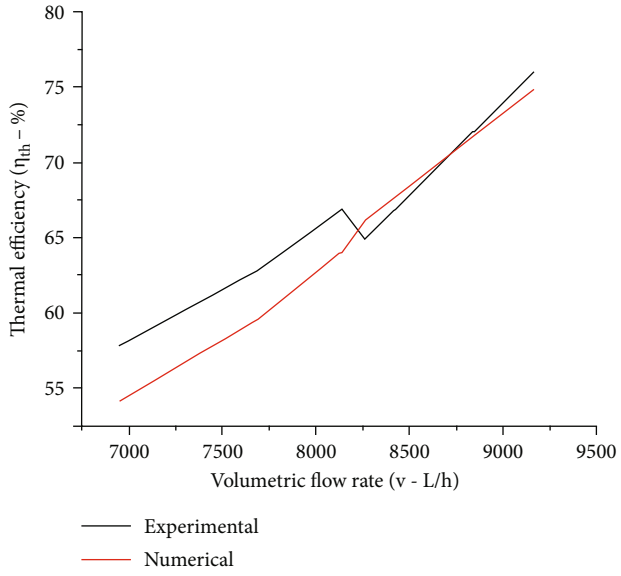


Figure 20: Changes in experimental and numerical values of thermal efficiency with volumetric flow rate.

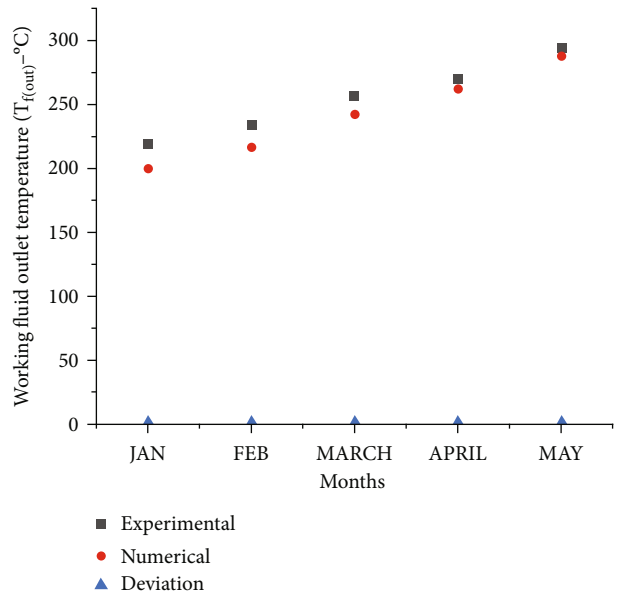


Figure 21: Deviation between experimental and numerical values of working fluid outlet temperature during January to May.

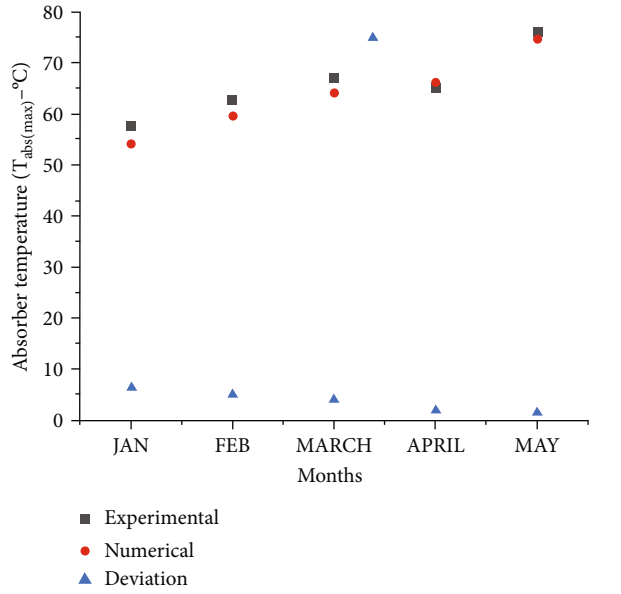


Figure 22: Deviation between experimental and numerical values of thermal efficiency during January to May.

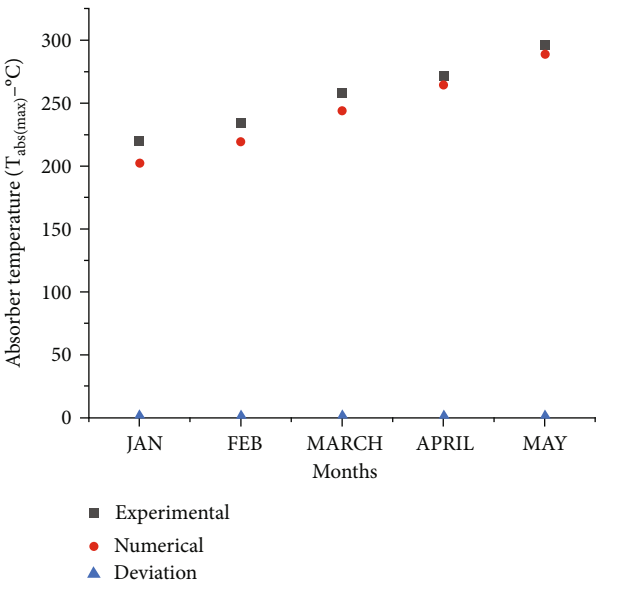


Figure 23: Deviation between experimental and numerical values of absorber temperature from January to May.

(volumetric flow rate), thermal efficiency reduces drastically due to inherent thermal properties of the working fluid.

As observed from Figures 21, 22, and 23, the deviation between experimental and numerical results of the SPDC system is that having a modified absorber was going to reduce from January to May. For May, it was very less. These less deviations' values validate the experimental results with the numerical results.

For the optimized performance of the SPDC system, absorber temperature and working fluid temperature at the exit of the absorber are the deciding factors. Maximum values of these two factors give maximum thermal efficiency of the system.

Comparative performance analysis of present and existing work is presented in Table 6. Researchers often focused on improving the thermal performance of the system by choosing different working fluid, absorber material, and improving absorber geometry. No one emphasizes on convolution effects on thermal performance. Present work emphasizes on improving thermal performance by reducing absorber size using convolution strategy. The superiority of this work is to use only 5 convolutions of spiral coil tubular

Table 6: Comparison with experimental results already published.

Performance parameter	Turrini et al.	Stefanovic et al.	Thirunavukkarasu and Cheralathan	Pavlovic et al.	Present work
Working fluid	Molten salt	Water	Water	Water	Water
Absorber material	Copper	Stainless steel AISI 304	Mild steel	Stainless steel	Copper
Absorber geometry	Spiraled tube cavity receiver	Spiral coil tubular receiver	Coiled spiral tube receiver	Spiral coil tubular flat receiver	Spiral coil tubular flat receiver with extended ends
Number of convolutions	09	13	06	13	05
Maximum η_{th} (%)	77%	49.83	56.21%	65%	75.98%

absorber with extended ends. Very less number of convolutions is found to achieve optimum performance of the modified absorber.

6. Conclusions and Recommendations

This work is a thermal and dynamic analysis of the utilization of a spiral coil tubular copper absorber on an SPDC. The spiral coil tubular copper absorber is appropriately designed to establish the most efficient substitute for the SPDC system. The system is first designed using a convolution strategy and then tested for different parameters with water as HTWF. The most valuable and vital concluding points of this innovative design of modified absorber are summarized as follows:

(i) Higher thermal efficiency for spiral coil tubular design leads to all the examined systems. It proves the need for the use of a spiral coil tubular copper absorber

(ii) Thermal efficiency was found the highest (i.e., η_{th} (max) = 75.98%) for May, and the mean thermal efficiency was found to be η_{th} (mean) = 65.72%

(iii) Numerical thermal efficiency for May and the mean value of this thermal efficiency are 74.86% and 63.80%, respectively, which validates with experimental results

(iv) The maximum absorber temperature and the maximum working fluid outlet temperature are about 296°C and 294.2°C, respectively, for May. This fact optimized the thermal efficiency of the modified spiral absorber

(v) The velocity of the working fluid is also increasing in order inside the absorber

(vi) Pumping pressure at the inlet of the absorber and the inlet of the steam inlet chamber coming out is 146 kPa and 142 kPa, respectively, and approximately equal to simulating one 150.1 kPa and 150 kPa, respectively

(vii) Improved working fluid flow characteristics resulting in less loss to turbulence

This optimized system geometry is to suit global location and season. The outputs from simulation and testing meet the solar thermal water pump system's demands for irrigation, heating, and drying applications and thermal process industries.

The recommendations for future work in this area are as follows:

(1) Optical analysis of this system in order to achieve the most optimal position of the absorber is recommended

(2) In order to reduce the cost of a concentrator, absorber material should be cheaper with efficient working. In place of copper, it could be other material or combination of materials

(3) The system is optimized for the use of water as the working fluid, although n-Pentane is the most suitable of those fluids studied. The flammability of n-Pentane could be hazardous if not handled correctly or if the system leaked near any source of ignition. This may make the system undesirable to certain markets. Water does however have many properties in its favour for this application. It is the researcher's opinion that this issue is only of minor concern and that it should not prevent continuing research into the system

(4) Preheating of working fluid enables it in performing efficiently, and it prevents absorber from hot-cold jerks which come from uneven heating at somehow extent

(5) Innovate the device or something for storing the solar energy so that system can be operated in rainy and cloudy days

Nomenclature

Alphabetical Symbols

A: Area, m^2
a: Constant
C_p: Specific heat of working fluid at constant pressure, kJ/kg K
d: Diameter, m
f: Friction factor
fr: Friction factor between wall of receiver and working fluid
G_b: Solar beam radiation, kwh/m^2/day
h: Coefficient of heat transfer, W/m^2K
h_e: Specific enthalpy, kJ/kg
I: Intensity of solar irradiation, W/m^2
k: Thermal conductivity, W/mK
L: Tube length from entry to exit
m: Mass flow rate, kg/s
M_x: Degree of momentum shifting
T: Temperature, °C
T_{fm}: Mean temperature of the working fluid flowing inside the receiver, °C
ΔP: Pressure difference between entry level and exit level, kPa
Pr: Prandtl number
Q_u: Useful energy generated, kW
Q_s: Available solar irradiation on dish aperture, kW/m^2
Q_{rad}: Radiation heat transfer losses, kW/m^2
$Q_{conv,in}$: Convective heat transfer losses inside the spiral coil tubular receiver, kW/m^2
$Q_{conv,ex}$: Convective heat transfer losses outside the spiral coil tubular receiver, kW/m^2
r: Radius, m
Re: Reynolds number

R: Radii of curvature of the spiral coil, m
V: Velocity, m/s
v_θ: Cross-sectional mean flow velocity of working fluid, m/s
x: Directional distance, m.

Subscripts

a: Aperture
abs: Absorber/receiver
cr: Critical
f: Working fluid
h: Hydraulic
i: Infinitesimal/inner
in: Inlet/entry
max: Maximum
o: Outer
opts: Optical solar
out: Outlet/exit
rec: Receiver/absorber
s: Surface
th: Thermal.

Abbreviations

CPC: Compound parabolic collector
CR: Concentration ratio
HTC: Heat transfer characteristics
HTWF: Heat transfer working fluid
PD: Parabolic dish
PTC: Parabolic trough collector
SPDC: Solar parabolic dish concentrator/collector.

Greek Symbols

θ: Total rotational angle
η: Efficiency
ρ: Fluid density, kg/L
v: Volumetric flow rate, L/h
γ: Intercept factor.

Data Availability

The data used to support the findings of this study are included within the article. These data were found experimentally, and no other specific data were used to support this study.

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.

Authors' Contributions

Rajkumar Malviya performed the methodology, wrote the original draft, and performed the investigation. Anil Kumar participated in the conceptualization, visualization, writing, reviewing, and editing of the manuscript, and supervision. Prashant V. Baredar did the supervision.

Acknowledgments

The first author would like to thank Maulana Azad National Institute of Technology, Bhopal (India), for awarding the fellowship for pursuing the Ph.D. in the Energy Centre and gratefully acknowledge the Centre for Energy and Environment, Delhi Technological University, Delhi, for the support to the present work.

References

[1] K. S. Reddy and N. Sendhil Kumar, "Combined laminar natural convection and surface radiation heat transfer in a modified cavity receiver of solar parabolic dish," *International Journal of Thermal Sciences*, vol. 47, no. 12, pp. 1647–1657, 2008.

[2] A. M. Daabo, S. Mahmoud, and R. K. Al-Dadah, "The optical efficiency of three different geometries of a small scale cavity receiver for concentrated solar applications," *Applied Energy*, vol. 179, pp. 1081–1096, 2016.

[3] S. Pavlovic, R. Loni, E. Bellos, D. Vasiljević, G. Najafi, and A. Kasaeian, "Comparative study of spiral and conical cavity receivers for a solar dish collector," *Energy Conversion and Management*, vol. 178, pp. 111–122, 2018.

[4] E. Bellos, E. Bousi, C. Tzivanidis, and S. Pavlovic, "Optical and thermal analysis of different cavity receiver designs for solar dish concentrators," *Energy Convers Manag X*, vol. 2, p. 100013, 2019.

[5] S. Pavlovic, A. M. Daabo, E. Bellos, V. Stefanovic, S. Mahmoud, and R. K. Al-Dadah, "Experimental and numerical investigation on the optical and thermal performance of solar parabolic dish and corrugated spiral cavity receiver," *Journal of Cleaner Production*, vol. 150, pp. 75–92, 2017.

[6] S. Chu, F. Bai, X. Zhang, B. Yang, Z. Cui, and F. Nie, "Experimental study and thermal analysis of a tubular pressurized air receiver," *Renewable Energy*, vol. 125, pp. 413–424, 2018.

[7] S. Turrini, M. Bettonte, M. Eccher, M. Grigiante, A. Miotello, and R. S. Brusa, "An innovative small-scale prototype plant integrating a solar dish concentrator with a molten salt storage system," *Renewable Energy*, vol. 123, pp. 150–161, 2018.

[8] V. P. Stefanovic, S. R. Pavlovic, E. Bellos, and C. Tzivanidis, "A detailed parametric analysis of a solar dish collector," *Sustainable Energy Technologies and Assessments*, vol. 25, pp. 99–110, 2018.

[9] H. Cherif, A. Ghomrassi, J. Sghaier, H. Mhiri, and P. Bournot, "A receiver geometrical details effect on a solar parabolic dish collector performance," *Energy Reports*, vol. 5, pp. 882–897, 2019.

[10] K. L. Lee, M. Jafarian, F. Ghanadi, M. Arjomandi, and G. J. Nathan, "An investigation into the effect of aspect ratio on the heat loss from a solar cavity receiver," *Solar Energy*, vol. 149, pp. 20–31, 2017.

[11] S. Yang, J. Wang, P. D. Lund, S. Wang, and C. Jiang, "Reducing convective heat losses in solar dish cavity receivers through a modified air-curtain system," *Solar Energy*, vol. 166, pp. 50–58, 2018.

[12] M. Tian, Y. Su, H. Zheng, G. Pei, G. Li, and S. Riffat, "A review on the recent research progress in the compound parabolic concentrator (CPC) for solar energy applications,"

Renewable and Sustainable Energy Reviews, vol. 82, pp. 1272–1296, 2018.

[13] R. Loni, E. Askari Asli-Ardeh, B. Ghobadian, E. Bellos, and W. G. le Roux, "Numerical comparison of a solar dish concentrator with different cavity receivers and working fluids," *Journal of Cleaner Production*, vol. 198, pp. 1013–1030, 2018.

[14] R. Loni, A. B. Kasaeian, E. Askari Asli-Ardeh, B. Ghobadian, and S. Gorjian, "Experimental and numerical study on dish concentrator with cubical and cylindrical cavity receivers using thermal oil," *Energy*, vol. 154, pp. 168–181, 2018.

[15] W. Fuqiang, C. Ziming, T. Jianyu, Y. Yuan, S. Yong, and L. Linhua, "Progress in concentrated solar power technology with parabolic trough collector system: a comprehensive review," *Renewable and Sustainable Energy Reviews*, vol. 79, pp. 1314–1328, 2017.

[16] E. Bellos, C. Tzivanidis, and K. A. Antonopoulos, "A detailed working fluid investigation for solar parabolic trough collectors," *Applied Thermal Engineering*, vol. 114, pp. 374–386, 2017.

[17] B. Zou, Y. Jiang, Y. Yao, and H. Yang, "Thermal performance improvement using unilateral spiral ribbed absorber tube for parabolic trough solar collector," *Solar Energy*, vol. 183, pp. 371–385, 2019.

[18] M. Bahrami, V. Madadi Avargani, and M. Bonyadi, "Comprehensive experimental and theoretical study of a novel still coupled to a solar dish concentrator," *Applied Thermal Engineering*, vol. 151, pp. 77–89, 2019.

[19] M. Ali, M. Rady, M. A. A. Attia, and E. M. Ewais, "Consistent coupled optical and thermal analysis of volumetric solar receivers with honeycomb absorbers," *Renewable Energy*, vol. 145, pp. 1849–1861, 2020.

[20] Q. Li, C. Zheng, S. Mesgari et al., "Experimental and numerical investigation of volumetric versus surface solar absorbers for a concentrated solar thermal collector," *Solar Energy*, vol. 136, pp. 349–364, 2016.

[21] A. B. Auti, D. R. Pangavane, T. P. Singh, and M. Sapre, "Estimation of time of tracking for designed absorber used in domestic solar desalination system using experimental and finite element analysis (ANSYS)," *Materials Today: Proceedings*, vol. 4, no. 2, pp. 2516–2524, 2017.

[22] J. Satpute, K. Chavan, and N. Gulhane, "Thermal performance investigation of concentrated solar collector using novel aluminum absorber," *Materials Today: Proceedings*, vol. 5, no. 2, pp. 4059–4065, 2018.

[23] W. Wang and B. Laumert, "Effect of cavity surface material on the concentrated solar flux distribution for an impinging receiver," *Solar Energy Materials & Solar Cells*, vol. 161, pp. 177–182, 2017.

[24] F. Wang, Y. Shuai, H. Tan, X. Zhang, and Q. Mao, "Heat transfer analyses of porous media receiver with multi-dish collector by coupling MCRT and FVM method," *Solar Energy*, vol. 93, pp. 158–168, 2013.

[25] S. Pavlovic R., E. Bellos, V. Stefanovic P., C. Tzivanidis, and Z. Stamenkovic M., "Design, simulation, and optimization of a solar dish collector with spiral-coil thermal absorber," *Thermal Science*, vol. 20, no. 4, pp. 1387–1397, 2016.

[26] R. Figaj, M. Szubel, E. Przenzak, and M. Filipowicz, "Feasibility of a small-scale hybrid dish/flat-plate solar collector system as a heat source for an absorption cooling unit," *Applied Thermal Engineering*, vol. 163, p. 114399, 2019.

[27] C. Zou, Y. Zhang, H. Feng et al., "Effects of geometric parameters on thermal performance for a cylindrical solar receiver using a 3D numerical model," *Energy Conversion and Management*, vol. 149, pp. 293–302, 2017.

[28] J. P. Bijarniya, K. Sudhakar, and P. Baredar, "Concentrated solar power technology in India: a review," *Renewable and Sustainable Energy Reviews*, vol. 63, pp. 593–603, 2016.

[29] V. Thirunavukkarasu and M. Cheralathan, "An experimental study on energy and exergy performance of a spiral tube receiver for solar parabolic dish concentrator," *Energy*, vol. 192, p. 116635, 2020.

[30] F. V. Barbosa, J. L. Afonso, F. B. Rodrigues, and J. C. F. Teixeira, "Development of a solar concentrator with tracking system," *The journal Mechanical Sciences*, vol. 7, no. 2, pp. 233–245, 2016.

[31] O. R. Sandoval, B. C. Caetano, M. U. Borges, J. J. García, and R. M. Valle, "Modelling, simulation and thermal analysis of a solar dish/Stirling system: a case study in Natal, Brazil," *Energy Conversion and Management*, vol. 181, pp. 189–201, 2019.

[32] S. Soltani, M. Bonyadi, and V. Madadi Avargani, "A novel optical-thermal modeling of a parabolic dish collector with a helically baffled cylindrical cavity receiver," *Energy*, vol. 168, pp. 88–98, 2019.

[33] R. Affandi, A. Ghani, C. K. Ghan, and L. G. Pheng, "The impact of the solar irradiation, collector and the receiver to the receiver losses in parabolic dish system," *Procedia - Social and Behavioral Sciences*, vol. 195, pp. 2382–2390, 2015.

[34] Y. O. O. Geun-jong, C. Hoon-ki, and D. Wa-ryong, "Fluid flow and heat transfer characteristics of spiral coiled tube : effects of Reynolds number and curvature ratio Fluid flow and heat transfer characteristics of spiral coiled tube : effects of Reynolds number and curvature ratio," 2014.

[35] R. Karimi, T. T. Gheinani, and V. Madadi Avargani, "A detailed mathematical model for thermal performance analysis of a cylindrical cavity receiver in a solar parabolic dish collector system," *Renewable Energy*, vol. 125, pp. 768–782, 2018.

[36] R. Loni, A. B. Kasaeian, O. Mahian, and A. Z. Sahin, "Thermodynamic analysis of an organic rankine cycle using a tubular solar cavity receiver," *Energy Conversion and Management*, vol. 127, pp. 494–503, 2016.

[37] S. A. Kalogirou, S. Karellas, K. Braimakis, C. Stanciu, and V. Badescu, "Exergy analysis of solar thermal collectors and processes," *Progress in Energy and Combustion Science*, vol. 56, pp. 106–137, 2016.

[38] X.-L. Li, X.-L. Xia, Z.-H. Li, and X. Chen, "Effects of double windows on optical and thermal performance of solar receivers under concentrated irradiation," *Solar Energy*, vol. 184, pp. 331–344, 2019.

[39] S. Skouri, S. Bouadila, M. Ben Salah, and S. Ben Nasrallah, "Comparative study of different means of concentrated solar flux measurement of solar parabolic dish," *Energy Conversion and Management*, vol. 76, pp. 1043–1052, 2013.

[40] T. Maatallah and W. B. Youssef, "Simulation and performance analysis of concentrating photovoltaic/thermal collector (CPV/T) with three-sided thermal insulation based on coupled optothermal model," *Solar Energy*, vol. 181, pp. 308–324, 2019.

[41] G. Barreto and P. Canhoto, "Modelling of a Stirling engine with parabolic dish for thermal to electric conversion of solar

energy," *Energy Conversion and Management*, vol. 132, pp. 119–135, 2017.

[42] C. Zou, Y. Zhang, Q. Falcoz et al., "Design and optimization of a high-temperature cavity receiver for a solar energy cascade utilization system," *Renewable Energy*, vol. 103, pp. 478–489, 2017.

[43] M. Hatami, J. Geng, and D. Jing, "Enhanced efficiency in concentrated parabolic solar collector (CPSC) with a porous absorber tube filled with metal nanoparticle suspension," *Green Energy & Environment*, vol. 3, no. 2, pp. 129–137, 2018.

Third Order Low Frequency Sinusoidal Oscillators Employing CFOAs and Grounded Capacitors

2021 7th International Conference on Signal Processing and Communication (ICSC) | 978-1-6654-2739-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICSC53193.2021.9673431

Ajishek Raj
Department of Electrical Engineering
Delhi Technological University
Delhi, India
ajishekaerotek@gmail.com

Pragati Kumar
Department of Electrical Engineering
Delhi Technological University
Delhi, India
pragatikumar@dtu.ac.in

D. R. Bhaskar
Department of Electrical Engineering
Delhi Technological University
Delhi, India
drbhaskar@dtu.ac.in

***Abstract*—In this communication, two new third order low frequency sinusoidal oscillators (TOLFSO) employing current feedback operational amplifiers (CFOAs) have been presented. The circuits employ four CFOAs, four/five resistors and three grounded capacitors.The proposed circuits offer fully decoupled/separate control of oscillation condition (OC) and oscillation frequency (OF). To corroborate the theoretical propositions, experimental results have been included using off-the-shelf available CFOA IC AD844.**

***Keywords* — Third Order Sinusoidal Oscillators, CFOA, Voltage Mode Oscillators, Low Frequency Oscillator**

I. Introduction

In biological and biomedical instrumentation, low frequency sinusoidal oscillator finds various applications [1]. These oscillators can also be employed to (i) measure an unknown capacitance, (ii) in frequency response display systems and (iii) in spectrum analysers [2] etc.

A large number of second order low frequency sinusoidal oscillators have been reported earlier employing operational amplifiers, operational transresistance amplifiers, current differencing buffered amplifier and CFOAs [3]. When compared with second order harmonic oscillators, third order sinusoidal oscillator circuits have low phase noise, better frequency accuracy, and high quality waveforms [4], and because of such advantageous features, third order sinusoidal oscillator (TOSO) circuits have been reported by analog circuit designers in open literature using various ABBs [5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34**?** , 35, 36, 37, 38, 39]. Since the proposed work deals with the realization of voltage mode (VM) third order sinusoidal oscillators, a brief description of previously reported VM TOSO circuits in which the output is available at low impedance nodes have been presented here, to put the proposed work in proper perspective.

In [11], two TOSOs using three OTRAs, three virtually grounded (VG) capacitors and five resistors have been presented. Out of two proposed circuits, only one circuit has seperate control of OC and OF. In [12], two different OTRA-based structures of TOSOs have been reported, in which, three OTRAs, three resistors and three floating capacitors (FC) have been utilized. The OC and OF of both the circuits can not be adjusted independently. One OTRA, three resistors, two VG capacitors and one grounded capacitor based TOSO was presented in [13]. The reported circuit does not have independent contol of OC and OF. Nagar and Paul reported a TOSO employing two OTRAs, four resistors and three FCs in [14]. In this circuit also, the OC and OF can not be adjusted independently without meeting additional condition(s). In [15], two OTRAs, five resistors and three VG capaciors based TOSO has been reported with seperate control of OC and OF. Three OTRAs, six resistors and three VG capacitors were used to realize TOSO in [16]. The proposed circuit did not possess the desirable property of seperate control of OC and OF. In [17], a TOSO circuit has been reported which utilizes three voltage differencing buffered amplifiers, two resistors and three grounded capacitors. The feature of seperate control of OC and OF is missing in this circuit also. Two voltage differencing buffered amplifiers, two grounded capacitors, one floating capacitor and one resistor was employed to realize a TOSO circuit with seperate control of OC and OF in [18]. In [19], voltage differencing inverting buffered amplifier (VDIBA)-based TOSO circuit has been reported which employs two VDIBAs, three grounded capacitors and one resistor having seperate control of OC and OF. Recently, in [39], twelve short circuit admittance matrices have been presented, out of which, four, low frequency third order sinusoidal oscillators derived having independent control of OC and OF.

A careful perusal of the available literature on TOSOs reveals that no TOSO circuit with fully decoupled control of OC and OF, capable of producing very low frequency oscillations, has been reported so far. Therefore, the main objective of this communication, is to present, two new TOSO circuits which can generate low frequency oscillations with fully decoupled control/seperate control of OC and OF. The presented TOSO circuits utilize grounded capacitors which are eminently acceptable for IC implementation. Experimental results have

978-1-6654-2739-5/21/$31.00 ©2021 IEEE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:34:35 UTC from IEEE Xplore. Restrictions apply.

been provided using off-the-shelf available CFOA IC AD844 to confirm the workability of proposed oscillators.

II. New Third Order Low Frequency Sinusoidal Oscillators

CFOA is a four terminal active block as shown in Fig. 1. Its terminal equations are given by:

$$v_x = v_y, \quad i_z = i_x, \quad v_y = 0, \quad and \quad v_w = v_z \tag{1}$$

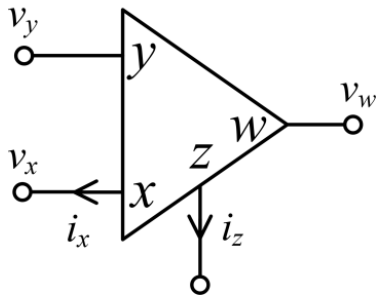
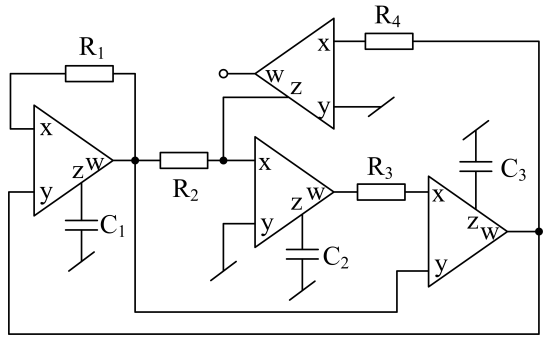
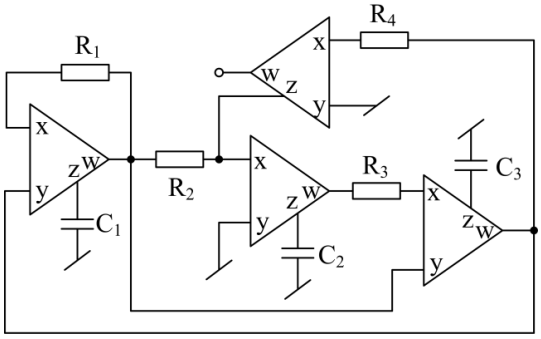


Fig. 1: Symbolic representation of CFOA

New third order low frequency sinusoidal oscillators employing CFOAs have been presented in Fig. 2.



(a)



(b)

Fig. 2: Proposed low frequency third order sinusoidal oscillators

Assuming ideal CFOAs, circuit analysis of the oscillator circuits shown in Fig. 2 (using equation(1)), gives the following characteristic equations (CE):
CE of Fig. 2a:

$$\begin{aligned} &s^3C_1C_2C_3R_1R_2R_3R_4R_5 + s^2C_2C_3R_2R_3R_4R_5 \\ &+sR_1R_2(C_1R_5 - C_2R_4) + (R_2R_5 - R_1R_4) = 0 \end{aligned} \tag{2}$$

CE of Fig. 2b:

$$\begin{aligned} &s^3C_1C_2C_3R_1R_2R_3R_4 + s^2C_2C_3R_2R_3R_4 + \\ &sR_2(C_1R_1 - C_2R_4) + (R_2 - R_4) = 0 \end{aligned} \tag{3}$$

Applying Routh-Hurwitz criterion on equations (2) and (3), OC and OF are found to be:
Fig. 2a

$$OC: \quad \frac{R_1}{R_2} = \frac{C_2}{C_1} \tag{4}$$

$$OF: \quad \omega_1 = \sqrt{\frac{1 - \left(\frac{R_4}{R_5}\right)}{C_2C_3R_3R_4}} \tag{5}$$

Fig. 2b

$$OC: \quad \frac{R_1}{R_2} = \frac{C_2}{C_1} \tag{6}$$

$$OF: \quad \omega_2 = \sqrt{\frac{1 - \left(\frac{R_4}{R_2}\right)}{C_2C_3R_3R_4}} \tag{7}$$

It has been noted that the OC and OF for the oscillator circuit given in Fig. 2a (as obtained from equations (4) and (5)), have fully uncoupled nature i.e., OC can be independently adjusted through R_1 and/or R_2 whereas, OF has three degrees of freedom (R_3, R_4 and R_5) to tune the frequency without disturbing OC. From the expression of OC and OF of Fig. 2b (given in equations (6) and (7)), it may be seen that the oscillator has seperate control of frequency by resistor R_3 and/or R_4 using which frequency can be controlled without disturbing the OC which can be adjusted through resistor R_1. From the equations (5) and (7), it may also be noticed that the proposed TOSO circuits are capable to generate low frequencies by taking the resistor ratio (R_4/R_5) and (R_4/R_2) very close to 1 under the consideration of $R_5 > R_4$ and $R_2 > R_4$ respectively.

III. Effects of CFOA parasitics on the proposed oscillators

Ideally, CFOA has a resistor of zero value between terminal 'x' and 'y', which shows in between 'x' and 'y' there is a voltage follower and at terminal 'z', it has a very high valued resistance in parallel with very low valued capacitor. But in actual, there is a voltage follower with a low valued resistor ($r_x = 50\Omega$) in between terminals 'x' and 'y' and, there is a resistor and capacitor of value $R_p = 3M\Omega$ and C_p = 4.5 pF at terminal 'z'.The non-ideal FOs including parasitic

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:34:35 UTC from IEEE Xplore. Restrictions apply.

components are given by:

$$\omega_{n1}^2 = \omega_{n1}^2 \left(\frac{1 - \frac{A}{C_1^n G_3^n G_4^n - C_2^n G_1^n G_3^n}}{1 + \frac{G_2+G_{p4}}{g_{x2}} + C_{p4} \frac{C_1^n C_2^n G_{p3} + C_2^n C_3^n (G_1^n + G_{p1})}{C_1^n C_2^n C_3^n g_{x2}}} \right) \tag{8}$$

$$\omega_{n2}^2 = \omega_{n2}^2 \left(\frac{1 - \frac{B}{C_1^n G_3^n G_4^n - C_2^n G_3^n G_5^n}}{1 + \frac{G_2+G_{p4}}{g_{x2}} + C_{p4} \frac{C_1^n C_2^n G_{p3} + C_2^n C_3^n G_1^n + C_1^n C_3^n G_{p1})}{C_1^n C_2^n C_3^n g_{x2}}} \right) \tag{9}$$

$$A = C_2^n G_1^n G_3^n \left(\frac{G_2^n + G_{p4}}{g_{x2}} \right) + \frac{C_{p4} G_{p2} (G_1^1 G_{p3} + G_{p1} G_{p3} - G_1^1 G_3^n)}{g_{x2}} + \frac{(g_{x2} + G_2^n + G_{p4}) \{C_1^n G_{p2} G_{p3} + (G_1^1 + G_{p1})(C_2^n G_{p3} + C_3^n G_{p2})\}}{g_{x2}} \tag{10}$$

$$B = C_2^n G_5^n G_3^n \left(\frac{G_2^n + G_{p4}}{g_{x2}} \right) + \frac{C_{p4} G_{p2} (G_1^n G_{p3} - G_5^n G_3^n)}{g_{x2}} + \left(1 + \frac{G_2^n + G_{p4}}{g_{x2}} \right) \{C_3^n G_{p2} G_1^n + G_{p3} (C_1^n G_{p2} + C_2^n G_1^n)\} \tag{11}$$

$$\omega_{n1} = \frac{C_1^n G_3^n G_4^n - C_2^n G_1^n G_3^n}{C_1^n C_2^n C_3^n}, \; \omega_{n2} = \frac{C_1^n G_3^n G_4^n - C_2^n G_5^n G_3^n}{C_1^n C_2^n C_3^n}$$

$$G_1^1 = \frac{1}{R_1 \left(1 + \frac{r_{x1}}{R_1}\right)}, \; G_1^n = \left(\frac{1}{R_1} \| \frac{1}{R_{p1}} \right), \; G_2^n = \frac{1}{R_2}$$

$$G_3^n = \frac{1}{R_3 \left(1 + \frac{r_{x3}}{R_3}\right)}, \; G_4^n = \frac{1}{R_4 \left(1 + \frac{r_{x4}}{R_4}\right)},$$

$$G_5^n = \frac{1}{R_5 \left(1 + \frac{r_{x5}}{R_5}\right)}, \; C_1^n = C_1 \left(1 + \frac{C_{p1}}{C_1}\right)$$

$$C_2^n = C_2 \left(1 + \frac{C_{p2}}{C_2}\right), \; C_3^n = C_3 \left(1 + \frac{C_{p3}}{C_3}\right)$$

and

$$G_{pi} = \frac{1}{R_{pi}}, \; i \to 2 - 4$$

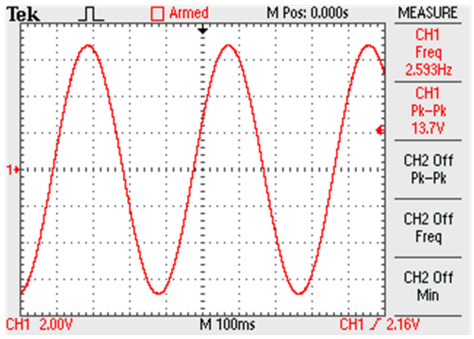
To reduce the effect of parasitic components on the FO of proposed oscillators, the passive components should be selected in such a manner:

$$C_i >> C_{pi}, \quad R_j >> r_{xi}, \quad R_1 << R_{p1} \quad and \quad R_{pi} >> 0$$
$$for \; i \to 1 - 3 \quad and \; j \to 3 - 5$$

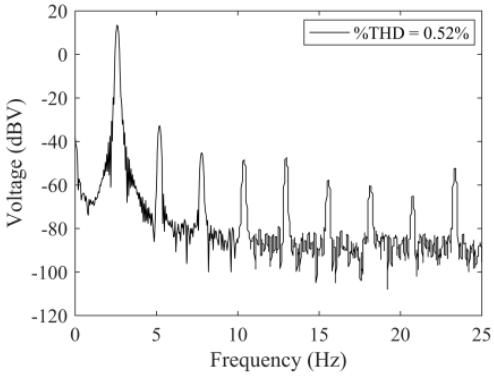
We have also evaluated the FOs quantitatively by substituting the parasitic component values in equations (8) and (9) for different design frequency of 2.81 Hz, 30 kHz and 1 MHz. The non-ideal oscillation frequencies were found as 2.87 Hz, 29.04 kHz and 978.45 kHz respectively. Thus, we can conclude that the parasitic resistances and capacitances of the CFOAs do not change the FO appreciably.

IV. Experimental Results

To validate the workability of theoretical propositions, both the low frequency TOSO circuits were bread-boarded with IC AD844, 2% tolerance resistors and 5% tolerance capacitors. The power supply voltages used were $\pm 10V$. The passive components used were: R_1 = (47+100) kΩ, R_2 = 100 kΩ, R_3 = 200 kΩ, R_4 = 80 kΩ, R_5 = 100 kΩ and identical valued capacitors $C_1 = C_2 = C_3$ = 200 nF for the designed frequency of 2.81 Hz. The experimental transient responses and frequency spectrum of the proposed TOSOs shown in Fig. 2a and Fig. 2b have been displayed in Fig. 3 and Fig. 4 respectively. The experimentally measured frequency for the oscillator circuit given of Fig. 2a was recorded as 2.59 Hz with peak to peak amplitude (V_{pp}) of 13.7 V, whereas for Fig. 2b, it was 2.92 Hz with 9.04 V_{pp}. The experimental total harmonic distortion (%THD) of the presented TOSO circuits were obtained through Keysight DSO X3034T and found to be 0.52% for Fig. 2a while for Fig. 2b it was 0.247%.



(a)



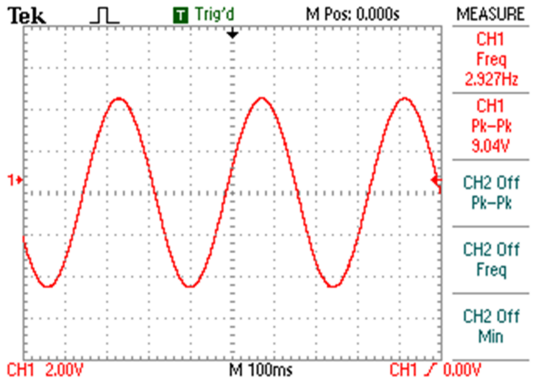
(b)

Fig. 3: Experimental results of circuit of Fig. 2a

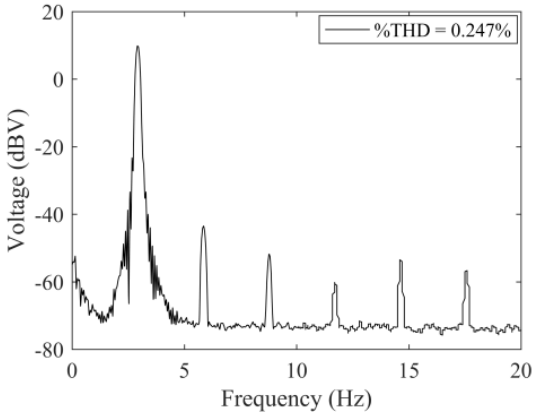
V. Conclusions

Two new third order sinusoidal oscillators employing CFOAs have been presented in this paper. The proposed oscillators utilize four/five resistors along with canonical numbers

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:34:35 UTC from IEEE Xplore. Restrictions apply.



(a)



(b)

Fig. 4: Experimental results of circuit of Fig. 2b

of grounded capacitors. Out of the two presented oscillator circuits, one has fully decoupled control of OC and OF and other possess non-interacting control of OC and OF. Both the proposed circuits have a feature to produce low frequency oscillations. Experimental results using CFOA IC AD844 have been presented to validate the theoretical propositions.

REFERENCES

[1] A. Manickam, A. Chevalier, M. McDermott, A. D. Ellington, and A. Hassibi, "A CMOS electrochemical impedance spectroscopy (EIS) biosensor array," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 4, no. 6, pp. 379–390, 2010.

[2] R. Helmer, "A test-signal generator for low-frequency instrumentation," *Behavior Research Methods, Instruments, & Computers*, vol. 18, no. 4, pp. 372–376, 1986.

[3] R. Senani, D. Bhaskar, V. K. Singh, and R. Sharma, *Sinusoidal oscillators and waveform generators using modern electronic circuit building blocks*. Springer, 2016, vol. 622.

[4] T. Hajder, "Higher order loops improve phase noise of feedback oscillators," *Applied Microwave and Wireless*, vol. 14, no. 10, pp. 24–31, 2002.

[5] S. Maheshwari and R. Verma, "Electronically tunable sinusoidal oscillator circuit," *Active and Passive Electronic Components*, vol. 2012, 2012.

[6] S. Maheshwari, "Quadrature oscillator using grounded components with current and voltage outputs," *IET circuits, devices & systems*, vol. 3, no. 4, pp. 153–160, 2009.

[7] J.-W. Horng*, C.-L. Hou, C.-M. Chang, W.-Y. Chung, H.-W. Tang, and Y.-H. Wen, "Quadrature oscillators us- ing CCIIs," *International journal of electronics*, vol. 92, no. 1, pp. 21–31, 2005.

[8] S. Maheshwari, "Current-mode third-order quadrature oscillator," *IET circuits, devices & systems*, vol. 4, no. 3, pp. 188–195, 2010.

[9] B. Chaturvedi and S. Maheshwari, "Third-order quadrature oscillator circuit with current and voltage outputs," *International Scholarly Research Notices*, vol. 2013, 2013.

[10] J.-W. Horng, H. Lee, and J.-Y. Wu, "Electronically tunable third-order quadrature oscillator using CDTAs," *Radioengineering*, vol. 19, no. 2, pp. 326–330, 2010.

[11] R. Pandey, N. Pandey, G. Komanapalli, and R. Anurag, "OTRA based voltage mode third order quadrature oscillator," *International Scholarly Research Notices*, vol. 2014, 2014.

[12] B. C. Nagar and S. K. Paul, "Voltage mode third order quadrature oscillators using OTRAs," *Analog Integrated Circuits and Signal Processing*, vol. 88, no. 3, pp. 517– 530, 2016.

[13] G. Komanapalli, N. Pandey, and R. Pandey, "New realization of third order sinusoidal oscillator using single OTRA," *AEU-International Journal of Electronics and Communications*, vol. 93, pp. 182–190, 2018.

[14] B. C. Nagar and S. K. Paul, "Realization of OTRA-based quadrature oscillator using third-order topology," in *Advances in Systems, Control and Automation*. Springer, 2018, pp. 375–386.

[15] G. Komanapalli and A. Tomar, "Realization of new third-order sinusoidal oscillator based on OTRA," in *Computing Algorithms with Applications in Engineering*. Springer, 2020, pp. 97–106.

[16] G. Komanapalli, N. Pandey, and R. Pandey, "OTRA based second and third order sinusoidal oscillators and their phase noise performance," in *AIP conference proceedings*, vol. 1859, no. 1. AIP Publishing LLC, 2017, p. 020017.

[17] C. Malhotra, V. Ahalawat, V. V. Kumar, R. Pandey, and N. Pandey, "Voltage differencing buffered amplifier based quadrature oscillator," in *2016 IEEE 1st International conference on power electronics, intelligent control and energy systems (ICPEICES)*. IEEE, 2016, pp. 1–4.

[18] K. L. Pushkar, "Voltage-mode third-order quadrature sinusoidal oscillator using VDBAs," *Circuits and systems*, vol. 8, no. 12, pp. 285–292, 2017.

[19] K. L. Pushkar and D. R. Bhaskar, "Voltage-mode third-order quadrature sinusoidal oscillator using VDIBAs," *Analog Integrated Circuits and Signal Processing*, vol. 98, no. 1, pp. 201–207, 2019.

[20] P. Prommee and K. Dejhan, "An integrable electronic-controlled quadrature sinusoidal oscillator using cmos operational transconductance amplifier," *International Journal of Electronics*, vol. 89, no. 5, pp. 365–379, 2002.

[21] J. Koton, N. Herencsar, K. Vrba, and B. Metin, "Current- and voltage-mode third-order quadrature oscillator," in *2012 13th International Conference on Optimization of Electrical and Electronic Equipment (OPTIM)*. IEEE, 2012, pp. 1203–1206.

[22] S. Lawanwisut and M. Siripruchyanun, "High output- impedance current-mode third-order quadrature oscillator based on CCCCTAs," in *TENCON 2009-2009 IEEE Region 10 Conference*. IEEE, 2009, pp. 1–4.

[23] N. Pandey and R. Pandey, "Approach for third order quadrature oscillator realisation," *IET Circuits, Devices & Systems*, vol. 9, no. 3, pp. 161–171, 2015.

[24] H.-P. Chen, Y.-S. Hwang, and Y.-T. Ku, "A new resistor- less and electronic tunable third-order quadrature oscillator with current and voltage outputs," *IETE Technical Review*, vol. 35, no. 4, pp. 426–438, 2018.

[25] H.-P. Chen, S.-F. Wang, Y.-N. Chen, and Q.-G. Huang, "Electronically tunable third-order quadrature oscillator using VDTAs," *Journal of Circuits, Systems and Computers*, vol. 28, no. 04, p. 1950066, 2019.

[26] H.-P. Chen, Y.-S. Hwang, and Y.-T. Ku, "Voltage-mode and current-mode resistorless third-order quadrature oscillator," *Applied Sciences*, vol. 7, no. 2, p. 179, 2017.

[27] K. Phanruttanachai and W. Jaikla, "Third order current- mode quadrature sinusoidal oscillator with high output impedances," *International Journal of Electronics and Communication Engineering*, vol. 7, no. 3, pp. 300–303, 2013.

[28] H.-P. Chen, Y.-S. Hwang, and Y.-T. Ku, "A systematic realization of

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:34:35 UTC from IEEE Xplore. Restrictions apply.

third-order quadrature oscillator with controllable amplitude," *AEU-International Journal of Electronics and Communications*, vol. 79, pp. 64–73, 2017.

[29] K. Khaw-Ngam, M. Kumngern, and F. Khateb, "Mixed- mode third-order quadrature oscillator based on single MCCFTA," *Radioengineering*, vol. 26, no. 2, pp. 522–535, 2017.

[30] A. M. Soliman, "Generation of third-order quadrature oscillator circuits using NAM expansion," *Journal of Circuits, Systems and Computers*, vol. 22, no. 07, p. 1350060, 2013.

[31] M. Kumngern and I. Kansiri, "Single-element control third-order quadrature oscillator using OTRAs," in *2014 Twelfth International Conference on ICT and Knowledge Engineering*. IEEE, 2014, pp. 24–27.

[32] M. Kumngern and U. Torteanchai, "Third order quadrature sinusoidal oscillator using single CDCTA," in *2017 2nd International conferences on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*. IEEE, 2017, pp. 440–443.

[33] M. Kumngern and J. Chanwutitum, "Single MCCCCTA-based mixed-mode third-order quadrature oscillator," in *2012 Fourth International Conference on Communications and Electronics (ICCE)*. IEEE, 2012, pp. 426–429.

[34] O. Channumsin and A. Jantakun, "Third-order sinusoidal oscillator using VDTAs and grounded capacitors with amplitude controllability," in *The 4th Joint International Conference on Information and Communication Technology, Electronic and Electrical Engineering (JICTEE)*. IEEE, 2014, pp. 1–4.

[35] Komal, K. L. Pushkar, and R. Kumar, "Electronically controllable third-order quadrature sinusoidal oscillator employing CMOS-OTAs," *Analog Integrated Circuits and Signal Processing*, pp. 1–7, 2020.

[36] S. Roy and R. R. Pal, "Electronically tunable third- order dual-mode quadrature sinusoidal oscillators employing VDCCs and all grounded components," *Integra- tion*, vol. 76, pp. 99–112, 2021.

[37] D. R. Bhaskar, A. Raj, and P. Kumar, "New resistorless third-order quadrature sinusoidal oscillators," *Journal of Circuits, Systems and Computers*, p. 2150194, 2021.

[38] A. Raj, D. R. Bhaskar, and P. Kumar, "Two new third- order quadrature sinusoidal oscillators," *IETE Journal of Research*, pp. 1–14, 2021.

[39] A. Raj, P. Kumar, and D. R. Bhaskar, "Systematic realization of low-frequency third-order sinusoidal oscillators," *International Journal of Circuit Theory and Applications*, 2021.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 25,2022 at 08:34:35 UTC from IEEE Xplore. Restrictions apply.

IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 11, No. 1, March 2022, pp. 129~136
ISSN: 2252-8938, DOI: 10.11591/ijai.v11.i1.pp129-136

Transfer learning for cancer diagnosis in histopathological images

Sandhya Aneja[1], Nagender Aneja[2], Pg Emeroylariffion Abas[1], Abdul Ghani Naim[2]
[1]Faculty of Integrated Technologies, Universiti Brunei Darussalam, Bandar Seri Begawan, Brunei Darussalam
[2]School of Digital Science, Universiti Brunei Darussalam, Bandar Seri Begawan, Brunei Darussalam

Article Info

Article history:

Received Jul 19, 2021
Revised Nov 30, 2021
Accepted Dec 12, 2021

Keywords:

Cancer detection
CNN
Feature extraction
Fine tuning
Transfer learning

ABSTRACT

Transfer learning allows us to exploit knowledge gained from one task to assist in solving another but relevant task. In modern computer vision research, the question is which architecture performs better for a given dataset. In this paper, we compare the performance of 14 pre-trained ImageNet models on the histopathologic cancer detection dataset, where each model has been configured as naive model, feature extractor model, or fine-tuned model. Densenet161 has been shown to have high precision whilst Resnet101 has a high recall. A high precision model is suitable to be used when follow-up examination cost is high, whilst low precision but a high recall/sensitivity model can be used when the cost of follow-up examination is low. Results also show that transfer learning helps to converge a model faster.

This is an open access article under the CC BY-SA license.

Corresponding Author:

Nagender Aneja
School of Digital Science, Universiti Brunei Darussalam
Bandar Seri Begawan, Brunei Darussalam
Email: nagender.aneja@ubd.edu.bn

1. INTRODUCTION

Transfer learning focuses on transferring latent knowledge from a source domain to a target domain, to solve the issue of insufficient data on the target domain. In the case of deep transfer learning, knowledge comprises architecture and trained parameters, wherein a nonlinear function is learned for the task in the target domain. Transfer learning has been used in computer vision, natural language processing, speech recognition, and deep reinforcement learning and has been successfully deployed in industries, including health care, autonomous driving, robotics, spam filtering, search, and recommendation. Applications of transfer learning in convolutional neural network (CNN) provide the benefits of starting the training process with a better initial model, and thus, faster learning and better performance. Transfer learning also is a useful approach for data and resource-constrained applications.

Although transfer learning is advantageous in transferring features from a source domain to a target domain, it may negatively impact the performance in the target domain, especially if the source domain is not related to the target domain. This is also known as negative transfer [1]. Recently, a number of CNN architectures have been proposed, with varying width i.e., number of feature maps, and depth i.e., number of layers. The architectures also vary in other parameters such as connections between hidden layers, normalization, and non-linear function. Most of the architectures have been shown to perform well on the ImageNet dataset. As such, for a specific application, there is a need to study which of the existing network architectures generalize well on that particular application.

Transfer learning is an extension of supervised learning such that the training set is from a source domain while the test set is from a target domain [2]. In traditional supervised learning, it is assumed that the test set comes from the same joint probability distribution of features and labels, and the posterior probability $P(X \cap Y) = P(Y|X).P(X)$ can be learned. However, transfer learning relaxes the condition that the test is from the same distribution and may not have labels or the same labels. A number of studies have been published on machine learning on cancer data, as it is the second-leading cause of death in the world. Mahmood *et al.* [3] surveyed neural network techniques for the classification of different cancers, where it has been observed that on many occasions, practitioners' opinion on the detection of a disease may vary according to clinical facts. A computer-based analysis would certainly help practitioners, by providing a second expert-opinion since neural network mimics the brain process. A carefully selected training data would be needed to accurately predict the disease.

Kornblith *et al.* [4] presented performance of 16 networks on non-health datasets. The authors found a correlation between ImageNet Accuracy and transfer accuracy when used as a fixed feature extractor or fine-tuned, however, the correlation has been found to be sensitive, in the case of a fixed feature extractor. The authors also discovered that in the case of Stanford Cars and fine-grained visual classification (FGVC) Aircraft dataset, pretraining using ImageNet provides minimal benefits. Tajbakhsh *et al.* [5] presented work on transfer learning in classification, detection, and segmentation between natural and medical images from radiology, cardiology, and gastroenterology. The authors demonstrated that pretraining with fine-tuning outperforms CNN that has been trained from scratch. Layer-wise tuning has also been shown to offer the best performance.

Guo *et al.* [6] presented an algorithm to find an optimal strategy for transfer learning for each instance, whereby the authors have used a policy network to decide whether to pass the image through fine-tuned layers or pre-trained layers. Opbroek *et al.* [7] have shown that support vector machine (SVM) based classifiers can outperform and minimize classification error by 60%, with a small amount of training data by using transfer learning. Raghu *et al.* [8] explored properties of transfer learning for medical imaging datasets, pertaining to Retina and CheXpert (chest x-ray images) data, and have shown that transfer learning offers little benefit to performance, with lightweight models giving comparable performance to the more complicated ImageNet architectures. However, the authors have only used two architectures: ResNet50 and Inception-v3, in their analysis. Cheplygina *et al.* [9] discussed around 140 papers on semi-supervised, multi-instance, and transfer learning, and have found transfer learning as the most popular. The usefulness of transfer learning has been demonstrated, but also mentioned the need for future research to examine connections between learning scenarios and generalization of results.

Rustam *et al.* [10] applied regression logistics and random forest for pancreatic cancer detection and have found random forest to perform better. Al-Khowarizmi and Suherman [11] applied a simple evolving connectionist system (SECoS) on a balanced cancer dataset, whilst Dabeer *et al.* [12] applied CNN for breast cancer detection using histopathology images from a biopsy. Pratiwi *et al.* [13] ensembled three architectures: inception V3, inception ResNet V2, and DenseNet 201, with the ensembled model showing sensitivity of 90%, specificity of 97%, precision of 82%, and recall of 85%. Thus, a comprehensive study of major architectures will be helpful for researchers to select the architecture for other datasets also.

2. RESEARCH METHOD

Transfer learning is an optimization technique to apply learning of one task to another task. This is important since deep learning needs lot of data to train and collecting data for each task is a challenging problem. Transfer learning can be applied in various ways. The section below provides a few definitions to formally define transfer learning, as well as presents various transfer learning methods.

- Definition 2.1 (Feature Space). Feature space, *X*, is set of all possible images of a particular size where each pixel value ranges 0-255. Training set is assumed to be subset of feature space, $X = \{x^i, x^i \in X, 1 \leq i \leq n\}$.
- Definition 2.2 (Label Space). Label space, *Y*, includes all possible labels for samples, and $Y = \{y^i, y^i \in Y, 1 \leq i \leq n\}$. In this paper, $Y = \{0, 1\}$, where 1 indicates presence of cancer tissue and 0 otherwise.
- Definition 2.3 (Prediction Function). Our objective is to learn prediction function $\hat{y} = f(x) or\ P(\hat{y}|x)$.
- Definition 2.4 (Domain). Domain, *D*, is set of feature space and probability distribution of training set $D = \{X, P(X)\}$.
- Definition 2.5 (Task). Task, *T*, is set of label space to predict and prediction function $T = \{Y, P(Y|X)\}$.
- Definition 2.6 (Transfer Learning). Given source domain $D_S = \{X_S,\ P(X_S)\}$ and source learning task $T_S = \{Y_S,\ P(Y_S|X_S)\}$, and target domain $D_T = \{X_T,\ P(X_T)\}$ and target learning task

$T_T = \{Y_T,\ P(Y_T|X_T)\}$, transfer learning deals with learning prediction function of target $P(Y_T|X_T)$ using knowledge of source domain D_S and source task T_S such that $D_S \neq D_T, T_S \neq T_T$, or no or limited labels in target domain [14].

2.1. Classification of deep transfer learning approaches

Transfer learning can help to provide a better initial model that can be used to faster training fro a different task. Neural information processing systems (NIPS) 1995 workshop learning to learn: knowledge consolidation and transfer in inductive systems provided initial motivations for this. Transfer learning approaches can be classified based on Domain and Task; Feature space and label space; Source and Target label space; and knowledge transfer process as described below:

2.1.1. Domain and task

Transfer of knowledge is possible in a wide variety of domains. Transfer Learning aims to improve learning of target task on target domain using knowledge from source task on source domain [15]. Transfer Learning is sometimes referred as domain adaption. We can classify transfer learning techniques, based on domain and task:

a. Same Domain and Different Task. Since the task is defined by $T = \{Y, P(Y|X)\}$, either *Y* is different or $P(Y|X)$ is different for source and target domain (or both are different).

– $Y_S \neq Y_T$, e.g., the source data set is CIFAR-10 whilst the target data set is CIFAR-100,

– $P(Y_S|X_S) \neq P(Y_T|X_T)$ i.e., the same image may be labeled differently in the source and the target domain, and it happens when the class balance is different (prior *P* (*Y*) is different in source and target.)

b. Different domain and same task. Since the domain is defined by $D = \{X, P(X)\}$, either *X* is different or *P* (*X*) is different for source and target domain (or both are different)

– $X_S \neq X_T$, different feature space, e.g., source data set is gray-scale images whilst target data set is color images or documents in different languages, $P(X_S) \neq P(Y_T)$, different distributions, while feature space may be same, e.g., the source domain consists of hand-drawn images while the target domain consists of photographs or documents in the same language but different topics. This is the most common scenario, commonly referred to as domain adaption, e.g. An autonomous model is trained for day driving but the trained model applied in all other scenarios.

2.1.2. Feature space and label space

Zhuang *et al.* [16] proposed a classification of transfer learning as homogeneous and heterogeneous transfer learning, based on feature space *X* and label space *Y*. Homogeneous transfer learning means X_S=X_T and Y_S=Y_T. Heterogeneous transfer learning means $X_S \neq X_T$ or/and $Y_S \neq Y_T$.

2.1.3. Source and target label space

Pan and Yang [14] classified transfer learning as Transductive, Inductive, and Unsupervised transfer learning, based on the label space *Y*. Transductive transfer learning means Y_S exists but Y_T does not exist. Inductive transfer learning means both Y_S and Y_T exist. \Unsupervised transfer learning means both Y_S and Y_T do not exist.

2.1.4. Knowledge transfer process

Tan *et al.* [17] classified deep transfer learning in four categories based on knowledge transfer process, namely, instances based, mapping based, network-based, and adversarial based. On the other hand, Li *et al.* [18] classified deep transfer learning into three approaches: instance transfer, model transfer, and feature representation transfer.

a. Instances-based or instance transfer: selects supplementary data from the source domain, whereby the selected source instances are similar to instances in the target domain. This is different from generative data augmentation, which generates synthetic data using generative adversarial network (GAN). Instance based transfer learning avoids negative transfer by selecting similar instances.

b. Mapping-based: both source and domain instances are mapped to a third domain space, whereby both domains are similar to each other. A union deep neural network is trained on modified source and target domain.

c. Network-based or model/feature representation transfer: this is the most common approach, and it focuses on network architecture and/or trained parameters. Source network architecture may be used as a feature extractor wherein convolution layers are frozen to extract the features, or may be used to fine-tune the target model, whereby the target model is initialized with trained parameters from the source domain, and then few/all of the layers are subsequently trained for the target domain. In other cases,

only the source network architecture is used with random weights for the target domain. This approach can further be classified in four cases:

- Target data set is large but different from the source: in this case, the entire model is trained for the target domain using sequential training for the source domain or joint training for both source and target domain
- Target data set is large and similar to the source: in this case, some layers are trained while other layers are frozen
- Target data set is small but different from source: in this case, some layers are trained while leaving other layers frozen and/or the use of domain adaption to find common feature space, which may also be determined using GANs [19], and
- Target data set is small and similar to the source: in this case, convolution base is frozen (no adaption to target).

d. Adversarial-based: In this case, transferable representations are determined using GAN. These transformations are applicable to both source and target domains. The features from both domains are extracted using front layers and sent to adversarial layers. The adversarial layer is trained to discriminate origin of features. Thus, the network is trained to learn general features. The representations are called transferable when the features are discriminative for the source domain, and indiscriminate with the shift between source and target domains [20].

In this paper, network-based or model/feature representation transfer for cancer detection has been analyzed. In particular, PatchCamelyon (PCam) benchmark data set for histopathologic cancer detection competition organized by Kaggle [21], has been used. The objective of the competition was to create an algorithm to identify metastatic cancer in small image patches taken from larger digital pathology scans.

The dataset [21] is a modified version of the PatchCamelyon (PCam) benchmark dataset [22], [23], with duplicate images removed. It has a total of 220,025 images, with 130,908 (60%) benign images and 89,117 (40%) images, where the center 32×32 px region of a patch contains at least one pixel of tumor tissue. Thus, the dataset is not balanced. Eight sample images from the dataset, containing tissue images of healthy and patient with tumor, are shown in Figure 1. Digitization of microscopic glass slide images has encouraged deep learning researchers for pathology diagnosis, where manual results differ considerably. In particular, trained deep convolutional neural network has been shown to perform better than manual cancer detection by different Pathologists. In this paper, performance of transfer learning from 14 CNN architectures has been presented. research methodology comprises three model training approaches for each architecture. The three training approaches are naive, feature extraction, and fine-tuning. These approaches are based on neural network architecture that comprises of convolutional layers and fully-connected layer(s).

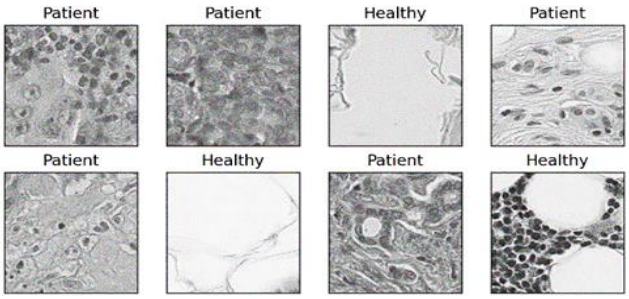


Figure 1. Sample images of healthy and patient with tumor, with center 32×32 px containing tumor tissue

In the Naive training experiment, neither the convolutional layers nor the fully connected layer(s) are trained, with only the pre-trained (trained on ImageNet dataset) layers used for testing. On the other hand, in the feature extractor experiment, the fully connected layer is trained for five epochs, whilst freezing the convolutional layers. Finally, in the fine-tune experiment, all layers from the convolutional layers and fully connected layer are trained. The fine-tune experiment is conducted for five epochs, only after conducting the second feature extractor experiment for five epochs. It has been found training for five epochs are sufficient for the model to converge, for both the convolutional and fully connected layers. At every epoch, we saved the model based on error rate i.e., after five epochs, we get the final model with the lowest error rate. The final saved model with the lowest error rate has been used to evaluate and compare different architectures.

3. RESULTS AND DISCUSSION

14 modern deep convolutional neural networks: Alexnet, Vgg16 bn, Vgg19 bn, Resnet18, Resnet34, Resnet50, Resnet101, Resnet152, Squeezenet1 0, Squeezenet1 1, Densenet121, Densenet169, Densenet201, and Densenet161, pretrained on ImageNet have been analyzed. We have considered batch size 32, input size 196×196, 80% training data, and 20% validation data, and learning rate of 0.01. Although the recommended input size is 224×224 for pre-trained models, we have reduced the input size slightly to 196×196 so that all models can fit onto our Nvidia 1080 GPU memory. Increasing input size is expected to increase Area Under the receiver operating characteristic curve (AUC) Score. Three experiments with 14 models, thus, a total of 42 sub-experiments, have been performed for this study.

3.1. Naive model

In the naive model, we have used weights from the pre-trained network. Since the model was not trained for the target task on the target domain, the model didn't generalize well. This will happen in particular when the target domain is different from source somain as is the case in this research. Results as in Figure 2 indicate that most models do not generalize well.

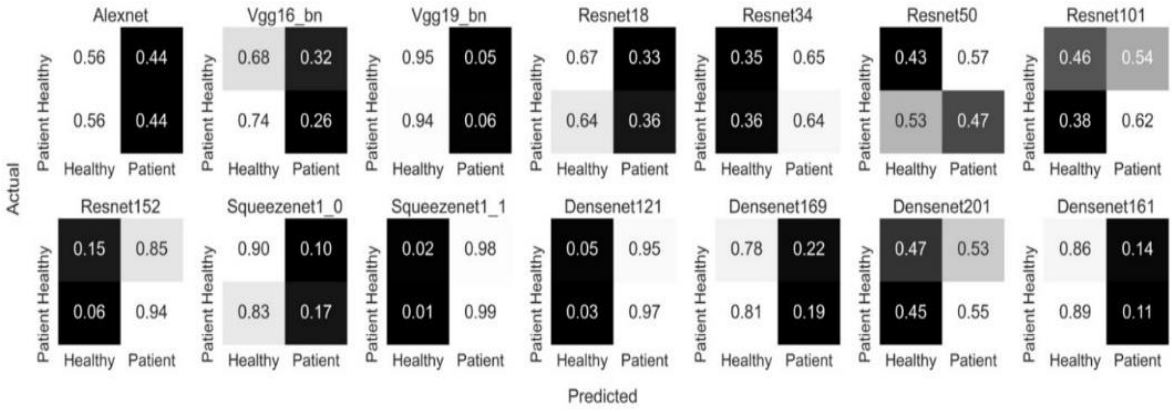


Figure 2. Naive model: confusion matrix (normalized)

3.2. Feature extractor model

In the feature extractor model, convolution layers are frozen, and the fully connected layer is trained for five epochs. Figure 3 displays the normalized confusion matrix after five epochs using the model that gives the lowest error rate during five epochs for every architecture. Figure 3 shows confusion matrix and Table 1 shows the performance of every architecture using the feature extractor model, in terms of AUC Score, F1 Score, Precision, and Recall. It can be seen that Densenet161 represents a high precision model, with precision value of 0.9710, while Resnet101 represents a high recall model, with recall value of 0.9569.

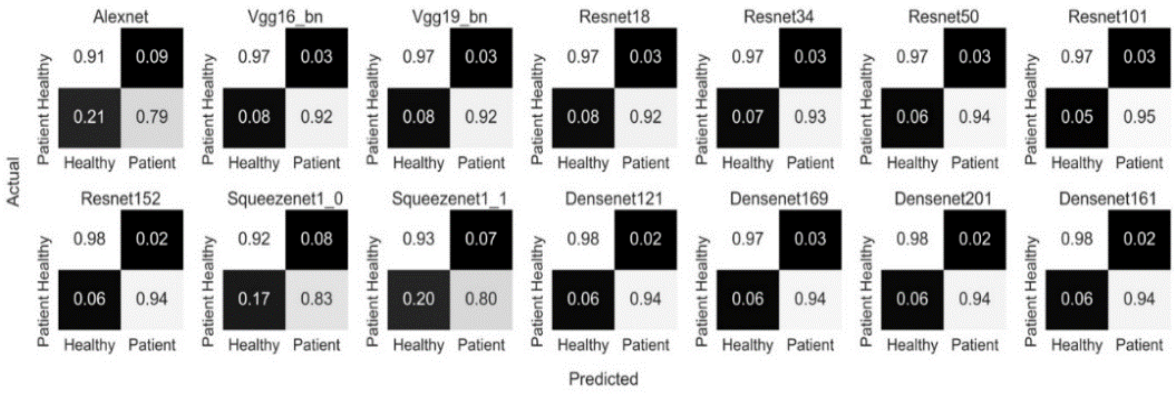


Figure 3. Feature extractor model: confusion matrix (normalized)

Table 1. Feature extractor model: AUC score, F1, precision, and recall for the 14 model

	AUC Score	F1	Precision	Recall
Alexnet	0.93872	0.82750	0.87121	0.78796
Vgg16 bn	0.98606	0.93751	0.95044	0.92492
Vgg19 bn	0.98710	0.93987	0.95831	0.92212
Resnet18	0.98640	0.93700	0.95482	0.91983
Resnet34	0.98823	0.94438	0.95614	0.93292
Resnet50	0.99046	0.94923	0.95990	0.93879
Resnet101	0.99131	0.95106	0.95685	0.94534
Resnet152	0.99131	0.95256	0.96525	0.94019
Squeezenet1 0	0.94847	0.85222	0.87606	0.82964
Squeezenet1 1	0.94659	0.84265	0.89129	0.79904
Densenet121	0.99065	0.95115	0.96355	0.93907
Densenet169	0.99151	0.95181	0.96173	0.94209
Densenet201	0.99196	0.95468	0.96566	0.94394
Densenet161	0.99240	0.95656	0.97101	0.94254

3.2. Fine tune model

In the fine tune model, we run five epochs to train the full network in addition to training a fully connected layer, with the fine-tuning experiment done only after feature extraction. Figure 4 displays normalized confusion matrix after five epochs using the models that give the lowest error rate. For this experiment, surprisingly, Squeezenet1 1 gave the worst performance, predicting most subjects as patients with cancer tissue as shown in the confusion matrix. Squeezenet1 1 has smaller number of parameters than the number of parameters in Squeezenet1 0 and other models, due to model's compression technique, as given in Iandola *et al.* [24], [25]. This trade-off has resulted in reduced performance of the model, as observed in the Figure 4. Table 2 shows that Densenet161 has a high precision value of 0.9703, whilst Densenet201 has a high recall value of 0.9458.

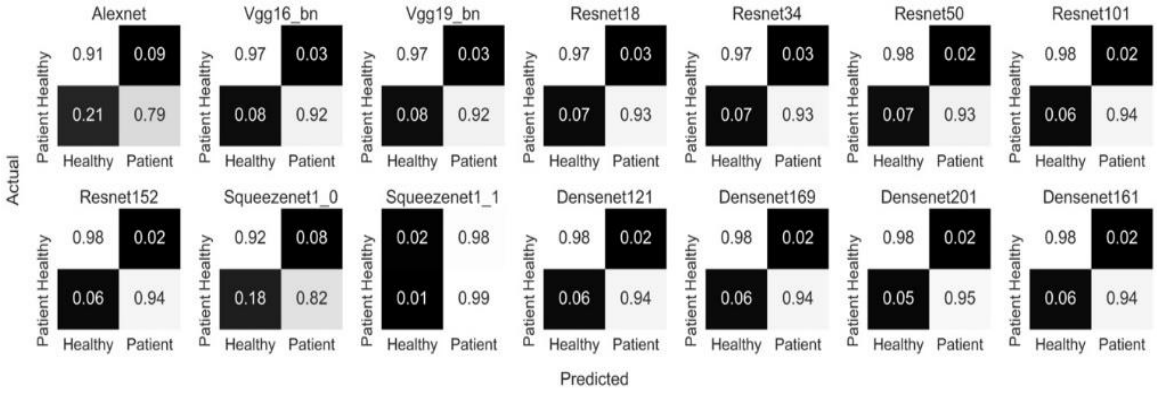


Figure 4. Fine tune model: confusion matrix (normalized)

Table 2. Fine tune model: AUC score, F1, precision, and recall for the 14 models

	AUC Score	F1	Precision	Recall
Alexnet	0.93362	0.82175	0.85449	0.79143
Vgg16 bn	0.98600	0.93684	0.95306	0.92117
Vgg19 bn	0.98707	0.93853	0.96015	0.91787
Resnet18	0.98630	0.93874	0.95239	0.92548
Resnet34	0.98844	0.94409	0.96111	0.92766
Resnet50	0.99097	0.94946	0.96852	0.93113
Resnet101	0.99149	0.95133	0.96646	0.93667
Resnet152	0.99080	0.95192	0.96612	0.93812
Squeezenet1 0	0.94676	0.85020	0.87813	0.82399
Squeezenet1 1	0.52655	0.57968	0.40978	0.99027
Densenet121	0.99092	0.95146	0.96277	0.94042
Densenet169	0.99186	0.95321	0.96653	0.94025
Densenet201	0.99235	0.95490	0.96418	0.94579
Densenet161	0.99225	0.95520	0.97028	0.94058

4. CONCLUSION

As humans learn new things by applying knowledge from previous learnings, transfer learning can help the progress of artificial general intelligence further. This research has shown that Densenet161 outperforms other architectures, with an AUC score of 0.9924 and F1 Sore 0.95 in the feature extractor model. Densenet161 has been shown to have high precision whilst Resnet101 is a high recall model. No significant performance improvement has been observed through fine-tuning, as the model has been first trained using feature extraction, such that there is limited scope for further improvement. A high precision model may be used where follow-up examination cost is high, whilst low precision but a high recall/sensitivity model can be used when the cost of follow-up examination is low. Future work may include applying pre-trained models on other categories of health datasets.

REFERENCES

[1] Z. Wang, Z. Dai, B. Poczos, and J. Carbonell, "Characterizing and avoiding negative transfer," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2019, pp. 11285–11294, doi: 10.1109/CVPR.2019.01155.

[2] N. Aneja and S. Aneja, "Transfer learning using CNN for handwritten devanagari character recognition," in *2019 1st International Conference on Advances in Information Technology (ICAIT)*, Jul. 2019, pp. 293–296, doi: 10.1109/ICAIT47043.2019.8987286.

[3] M. Mahmood, B. Al-Khateeb, and W. Makki Alwash, "A review on neural networks approach on classifying cancers," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 9, no. 2, pp. 317–326, Jun. 2020, doi: 10.11591/ijai.v9.i2.pp317-326.

[4] S. Kornblith, J. Shlens, and Q. V. Le, "Do better imagenet models transfer better?," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2019, pp. 2656–2666, doi: 10.1109/CVPR.2019.00277.

[5] N. Tajbakhsh *et al.*, "Convolutional neural networks for medical image analysis: full training or fine tuning?," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1299–1312, May 2016, doi: 10.1109/TMI.2016.2535302.

[6] Y. Guo, H. Shi, A. Kumar, K. Grauman, T. Rosing, and R. Feris, "Spottune: transfer learning through adaptive fine-tuning," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2019, pp. 4800–4809, doi: 10.1109/CVPR.2019.00494.

[7] A. van Opbroek, M. A. Ikram, M. W. Vernooij, and M. de Bruijne, "Transfer learning improves supervised image segmentation across imaging protocols," *IEEE Transactions on Medical Imaging*, vol. 34, no. 5, pp. 1018–1030, May 2015, doi: 10.1109/TMI.2014.2366792.

[8] M. Raghu, C. Zhang, J. Kleinberg, and S. Bengio, "Transfusion: understanding transfer learning for medical imaging," *Advances in Neural Information Processing Systems*, vol. 32, pp. 3342–3352, 2019.

[9] V. Cheplygina, M. de Bruijne, and J. P. W. Pluim, "Not-so-supervised: a survey of semi-supervised, multi-instance, and transfer learning in medical image analysis," *Medical Image Analysis*, vol. 54, pp. 280–296, May 2019, doi: 10.1016/j.media.2019.03.009.

[10] Z. Rustam, F. Zhafarina, G. S. Saragih, and S. Hartini, "Pancreatic cancer classification using logistic regression and random forest," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 10, no. 2, pp. 476–481, Jun. 2021, doi: 10.11591/ijai.v10.i2.pp476-481.

[11] A.-K. Al-Khowarizmi and S. Suherman, "Classification of skin cancer images by applying simple evolving connectionist system," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 10, no. 2, pp. 421–429, Jun. 2021, doi: 10.11591/ijai.v10.i2.pp421-429.

[12] S. Dabeer, M. M. Khan, and S. Islam, "Cancer diagnosis in histopathological image: CNN based approach," *Informatics in Medicine Unlocked*, vol. 16, 2019, doi: 10.1016/j.imu.2019.100231.

[13] R. A. Pratiwi, S. Nurmaini, D. P. Rini, M. N. Rachmatullah, and A. Darmawahyuni, "Deep ensemble learning for skin lesions classification with convolutional neural network," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 10, no. 3, pp. 563–570, Sep. 2021, doi: 10.11591/ijai.v10.i3.pp563-570.

[14] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010, doi: 10.1109/TKDE.2009.191.

[15] C. Tan, F. Sun, T. Kong, W. Zhang, C. Yang, and C. Liu, "A survey on deep transfer learning," *International conference on artificial neural networks*, pp. 270-279. Springer, Cham, 2018, doi: 10.1007/978-3-030-01424-7_27.

[16] F. Zhuang *et al.*, "A comprehensive survey on transfer learning," *Proceedings of the IEEE*, vol. 109, no. 1, pp. 43–76, Jan. 2021, doi: 10.1109/JPROC.2020.3004555.

[17] C. Tan, F. Sun, T. Kong, W. Zhang, C. Yang, and C. Liu, "A survey on deep transfer learning," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11141, Springer International Publishing, 2018, pp. 270–279.

[18] Y. Li, Y. Yang, S. Zhou, J. Qiao, and B. Long, "Deep transfer learning for search and recommendation," in *Companion Proceedings of the Web Conference 2020*, Apr. 2020, pp. 313–314, doi: 10.1145/3366424.3383115.

[19] E. Tzeng, J. Hoffman, K. Saenko, and T. Darrell, "Adversarial discriminative domain adaptation," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jul. 2017, pp. 2962–2971, doi: 10.1109/CVPR.2017.316.

[20] Y. Ganin *et al.*, "Domain-adversarial training of neural networks," in *Advances in Computer Vision and Pattern Recognition*, Springer International Publishing, 2017, pp. 189–209.

[21] B. Veeling, B. E. Bejnordi, G. Litjens, and J. V. D. Laak, "Histopathologic cancer detection," *Kaggle*. 2019.

[22] B. S. Veeling, J. Linmans, J. Winkens, T. Cohen, and M. Welling, "Rotation equivariant CNNs for digital pathology," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11071, Springer International Publishing, 2018, pp. 210–218.

[23] B. Ehteshami Bejnordi *et al.*, "Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer," *JAMA*, vol. 318, no. 22, pp. 2199–2210, Dec. 2017, doi: 10.1001/jama.2017.14585.

[24] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size," Feb. 2016, [Online]. Available: http://arxiv.org/abs/1602.07360.

[25] B. Koonce, "SqueezeNet," in *Convolutional Neural Networks with Swift for Tensorflow*, Berkeley, CA: Apress, 2021, pp. 73–85.

BIOGRAPHIES OF AUTHORS



Dr Sandhya Aneja is working as Assistant Professor of Information and Communication System Engineering at the Faculty of Integrated Technologies, Universiti Brunei Darussalam. Her primary areas of research interest include wireless networks, high-performance computing, internet of things, artificial intelligence technologies - machine learning, machine translation, deep learning, data science, and data analytics. Further info on her homepage http://expert.ubd.edu.bn/sandhya.aneja. She can be contacted at sandhya.aneja@ubd.edu.bn.



Dr Nagender Aneja is working as Assistant Professor at School of Digital Science, Universiti Brunei Darussalam. He did his Ph.D. in Computer Engineering from J.C. Bose University of Science and Technology YMCA, and M.E. Computer Technology and Applications from Delhi College of Engineering. He is currently working in the area of deep learning, computer vision, and natural language processing. He is also founder of ResearchID.co, further info on his homepage http://expert.ubd.edu.bn/nagender.aneja. He can be contacted at nagender.aneja@ubd.edu.bn



Pg Dr Emeroylariffion Abas received his B.Eng. Information Systems Engineering from Imperial College, London in 2001, before obtaining his Ph.D. Communication Systems in 2005 from the same institution. He is now working as an Assistant Professor in General Engineering, Faculty of Integrated Technologies, Universiti Brunei Darussalam. His present research interest are data analytic, energy systems and photonics. Further info on his homepage https://expert.ubd.edu.bn/emeroylariffion.abas. He can be contacted at emeroylariffion.abas@ubd.edu.bn.



Dr Abdul Ghani Naim is a Senior Assistant Professor at School of Digital Science, Universiti Brunei Darussalam. He did his Ph.D. in Information Security from the Royal Holloway College, London. His present research interests include computer security, cryptography, high performance computing, and machine learning. Further info on his homepage https://expert.ubd.edu.bn/ghani.naim. He can be contacted at ghani.naim@ubd.edu.bn.

Multimedia Tools and Applications
https://doi.org/10.1007/s11042-021-11492-w

1205: EMERGING TECHNOLOGIES FOR INFORMATION HIDING AND FORENSICS IN MULTIMEDIA SYSTEMS



Watermarking of ECG signals compressed using Fourier decomposition method

Prashant Mani Tripathi[1] · Ashish Kumar[2] · Rama Komaragiri[1] · Manjeet Kumar[3]

Received: 27 November 2020 / Revised: 12 July 2021 / Accepted: 19 August 2021
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract
Secure and efficient transmission of biomedical signals containing delicate and vital health information of a subject over the network is challenging. Digital watermarking is an attractive choice for the secure and reliable transmission of biomedical signals using smart healthcare devices. This paper proposes a method that integrates watermarking and compression of an electrocardiogram (ECG) of a subject. ECG watermarking secures the cardiac information of a subject, while signal compression reduces the amount of data. In this study, a sequence of binary ones and zeros are used as a watermark to secure the ECG signal. The watermarked ECG signal is analyzed and compressed by the Fourier decomposition method by removing the redundant component present in the ECG signal. The proposed method is validated using the MIT-BIH arrhythmia database. Percentage root-mean-square difference (PRD), the output signal to noise ratio (SNR), compression ratio, and computation time are used to evaluate the performance of the proposed method. Results show that the proposed method reduces the PRD by a factor of three, and the compression ratio increases by a factor of two. The proposed algorithm is evaluated at different input SNR values. The improved output SNR demonstrates the denoising capabilities of the watermarked signal.

Keywords Electrocardiogram · Watermarking · Compression · Fourier decomposition method

1 Introduction

In the past years, the advancement in the internet and communication networks created many avenues for information exchange. The internet is a ubiquitous information infrastructure and became a necessity of everyday life. The communication between people, exchange of information and data became more accessible and convenient with

✉ Manjeet Kumar
manjeetchhillar@gmail.com

Extended author information available on the last page of the article

Published online: 02 January 2022

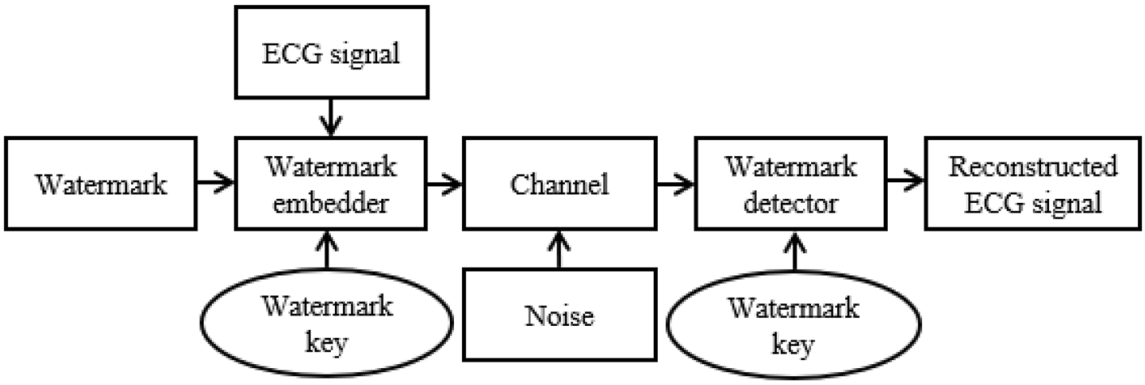


Fig. 1 Block diagram of the ECG watermarking system

the internet. However, the internet is an insecure channel for the transmission of information [36]. While using the internet, information security and privacy have become a primary concern for researchers [42]. The emerging technologies and the internet allows the exchange of subjects' bio-medical information with medical practitioners to seek remote medical assistance [37]. In modern-day technology, biomedical signals like ECG and electroencephalogram (EEG) provide vital information on health conditions and have become popular in biometric identity and security applications [15, 29, 32, 38]. Hence the modern health care systems demand a secure and reliable transmission of these signals.

It is legally required to securely keep the medical and personal data of a subject [1]. Watermarking is an important technique to secure health information in an ECG signal of a subject [37]. ECG watermarking can provide secure transmission of an ECG signal and enable it to be restored to its original state. In ECG watermarking, the ECG signal is the cover signal. The details of the subject are considered the watermark [9, 37]. Embedding watermarking data into a cover signal (ECG signal) degrades the information contained by an ECG signal, thus affecting the diagnosis. Deterioration of the ECG signal may result in the loss of critical health information, which is the primary disadvantage of ECG signal watermarking [37]. In an ECG signal, the QRS complex, R-peak, P, and T waves provide critical diagnostic information. Hence, ECG watermarking should minimally affect the signal to preserve the information for accurate diagnosis and still provide sufficient security.

In ECG watermarking, a watermark is embedded using two approaches. First is the frequency domain, where the frequency is modified. The second method uses the spatial domain to modify the intensities of pixels [1]. Figure 1 shows the block diagram of ECG watermarking architecture.

Apart from security, watermarking is used to add additional information to the signal. Watermarking differs from other techniques of associating the information, such as encoding information in an image in the form of visible bar code and keeping the information in the header of a digital file [9]. The virtues of watermarking technique that make it valuable to apply to ECG signal transmission are (i) watermarking is imperceptible, (ii) embedded watermarks are inseparable from the embedded work (iii) watermarked signals undergo the same transformation as the signal [9]. Every watermarking system has some desirable properties, as listed below:

Effectiveness is defined as the probability of the detection of messages just after embedding. Ideally, *100%* effectiveness is required.

Fidelity represents the similarity between the watermarked and the original signal.

Data payload is the number of bits encoded within a signal or within a unit of time by a watermark encoder. Data payload encodes within a signal or within a unit of time.

The false-positive rate is the number of signals identified as watermark embedded, but they have no embedded watermark. The watermarking system should possess a low false-positive rate.

Robustness represents the ability to detect watermarks after signal processing. A robust watermark system should be able to withstand rotation, scaling, cropping, compression, additive Gaussian noise, and many more operations [5]. The robustness of the watermark depends on the application [22, 44].

Security implies the ability of the watermark to protect the information against hostile attacks. Any process which spoils the purpose of a watermark is known as a hostile attack. The types of attacks are categorized as Unauthorized embedding, unauthorized detection, and unauthorized removal. Unauthorized embedding and removal are classified as active attacks as they modify the cover signal. Unauthorized detection is considered a passive attack because they do not modify the cover work. In general, a secret watermark is required and must be undetectable by unauthorized users.

The organization of this paper is as follows: Section 2 introduces literature and background. The proposed method and evaluation parameters are discussed in Section 3. Results are discussed in Sections 4 and 5 provides conclusions and future scope.

2 Literature and background

At present, the demand for secure and reliable transmission of ECG signals is ever increasing. Watermarking is a possible solution to protect and secure the ECG signal without affecting the original information contained in an ECG signal [42]. The protection of the ECG signal using watermarking is still in its infancy. In literature, various groups of researchers have proposed many algorithms for ECG watermarking. Sanivarapu et al. [37] proposed a wavelet-based ECG watermarking approach to hide the patient information in ECG signal as a QR image. In this work, a 1D-ECG signal is converted into a 2D-ECG image. The resultant 2D-ECG image is decomposed using DWT. The QR image and the detail coefficients of the wavelet are decomposed using QR decomposition. In the next step, the watermarking data is embedded into a 2D-ECG image. The embedded patient information is easily extractable without distortion. Jero et al. [20] proposed a curvelet-based ECG signal watermarking. In [20], the 2D-ECG image is obtained from a 1D-ECG signal. Jero et al. [20] use curvelet transform, adaptive selection of watermark position, and a new threshold selection algorithm to obtain a high-quality watermarked signal. Kaur et al. [21] proposed a simple embed watermarking technique using a low-frequency chirp signal. This method allows complete removal of the embedded watermark at the receiver due to blind recovery.

Jero et al. [19] used the DWT and singular value decomposition (SVD) to watermark an ECG signal. DWT decomposes the ECG signal, and SVD embeds the secret data in the decomposed signal. Results show that the proposed method in [19] achieved a very low signal degradation. The available methods in the literature use a 2D-ECG image to watermark an ECG signal.

In this work, to protect the ECG signal, a 1D-ECG watermarking technique using frequency decomposition method (FDM) based compression is proposed. The proposed

watermarking scheme keeps the watermarked signal quality at a satisfying level without damaging the characteristics of the ECG signal. This ability of the watermarking process makes it difficult to notice or perceive the embedded information in the watermarked signal.

In modern health care systems, sometimes it is required to continuously observe the ECG signal of a subject over extended periods of durations, which generates a large volume of data. This data requires to be transmitted remotely to a hospital or a health supervisor. This demands a large storage capacity and channel bandwidth for the storage and transmission of generated ECG data. Compression of the ECG signal is a key technology for efficient transmission and storage of extensive ECG data. With advancements in communication and digital storage technology, ECG signal compression is getting more attention from researchers. ECG signal compression techniques are classified into three categories: (i) time-domain techniques, (ii) frequency (transform) domain techniques, and (iii) hybrid techniques [31]. Time-domain techniques are based on direct data compression and parameter extraction, while frequency domain techniques are transformation-based [18, 31]. Compression techniques, like differential pulse code modulation (DPCM), reduce signal redundancy by considering the correlation between inter-sample [18]. DPCM based compression technique has advantages of simplicity and easy implementation, but it suffers from a low compression rate [42]. Data compression by entropy coding is based on assigning the variable-length codewords to the signal according to the frequency of occurrence. Huffman and Shannon Fano coding methods are commonly used for assigning variable length codeword to the signals. The frequent occurring signals are assigned a shorter codeword length, while less occurring signals have longer codeword length. However, in the variable-length coding, decoding errors may occur due to transmission errors [18].

In the transform compression technique, the input signal is pre-processed by a linear orthogonal transform. The transformed output is appropriately encoded with a reduced amount of data needed to characterize the input signal. Many orthogonal transforms, such as Karhunen–Loeve (KL) transform [35], Fourier transform [3], discrete cosine transform (DCT) [6], are to used to compress the ECG signal. Olmos et al. [35] proposed an ECG data compression technique based on KL transform. With KL transform, they achieved a mean compression ratio (CR) of 12.1 and an average mean squared error (MSE) of 0.44%. Although the MSE obtained in [35] is good, but the compression ratio is low. Al-Nashash in [3] proposed a new ECG data compression technique that models the quasi-periodic ECG signal by a dynamic Fourier series. The proposed method in [3] achieved a CR of up to 16 and a PRD of about 3%. However, the CR in [3] varies with the heart rate. Batista et al. [6] utilized DCT for compression of the ECG signal. In [6], the ECG compression is achieved by optimized quantization of the DCT coefficient. This DCT based method has a trade-off between CR and processing time.

New techniques based on discrete wavelet transform (DWT) [10] and neural networks [43] are also used to compress an ECG signal. Charri et al. [10] proposed an ECG compression technique that utilizes the DWT and run-length encoding. The use of DWT provides a better localization of different morphological features of the ECG signal. In neural network-based compression, the data is compressed by extracting the features of the ECG signal through self-learning [43]. Zhang et al. [43] obtained a CR of 1:19 with PRD of 12% using neural network-based ECG compression. Although the neural network-based approach has strong adaptability, anti-interference, and parallel processing, the optimization of hidden layers is difficult [43]. Kumar et al. [27] obtained an average CR of 31.41 using the biorthogonal wavelet transform based compression technique.

In this paper, the ECG signal compression is achieved by the Fourier decomposition method (FDM). FDM being free from mode mixing, end effect artifacts, and detrend uncertainty preserves the valuable information embedded in the ECG signal. FDM is a promising method to analyze nonlinear and nonstationary signals [39]. The FDM is a generalized Fourier expansion of a time series with variable amplitude and variable frequency [39]. FDM is based on a Fourier theory, representing a multicomponent signal with a sum of average value and nonstationary monocomponent signals, known as Fourier intrinsic band function (FIBF). A FIBFs is a zero-mean orthogonal function and allows analytic representation. A FIBF is the real part of an analytic FIBF [39]. FDM preserves the energy of the signal by decomposing the signal into a set of orthogonal or linearly independent non-orthogonal yet energy preserving FIBFs in desired frequency bands [40]. Each FIBFs is an amplitude and frequency modulated signal, thus providing local analysis of nonstationary and nonlinear signals [12]. FDM makes the individual component analysis easy by providing the time–frequency-energy representation of the signal. The energy preservation property of FDM helps to retain the characteristic of the ECG signal. FDM decomposes the ECG signal into different FIBFs. Depending on the frequency component, each FIBF corresponds to a noise like a baseline wander, power line interference, or a useful signal. Once the FIBFs corresponding to the unwanted signal is rejected, the remaining FIBFs represent a clean ECG signal.

Singh et al. [39] have shown that FDM is a powerful tool for analyzing earthquake data. FDM provides an enhanced time–frequency-energy(TFE) distribution of earthquake data to understand better the effect of an earthquake recording station from the epical center. FDM is also advantageous in speech signal analysis. In [39], Singh et al. demonstrated that the FIBFs of a quasi-periodic voice signal provide detailed information regarding noise, signal components, and fundamental frequency. FDM emerges as a powerful tool for analyzing the nonlinear and nonstationary signals and analyzing the biomedical signals. Singhal et al. [40] show that the FDM is an efficient tool for denoising the ECG signal. Fatimah [12] et al. proposed a sleep Apnea detection algorithm from the ECG signal based on FDM.

The proposed approach safeguards the ECG signal and compresses it to minimize the storage and transmission bandwidth requirement. In literature, many databases are used by various researchers to evaluate the performance of proposed ECG signal detection and compression techniques. Some frequently used ECG databases are the MIT-BIH arrhythmia database [13, 24, 33], PTB database [7, 16], European ST-T database [41], QT database [8, 28], Fantasia database [14, 17, 25], American Heart Association (AHA) database [4, 23]. Brief information on these databases is provided in Table 1.

Table 1 Features of the standard ECG database

Database	Number of records	Duration (minf)	Sampling frequency (Hz)	Availability
MIT-BIH Arrhythmia [24]	48	30	360	Open access
PTB [33]	549	1.55	1000	Open access
European ST-T [16]	90	120	250	Open access
QT [41]	105	15	250	Open access
Fantasia [8]	40	120	250	Open access
AHA [4]	155	150	250	Open access

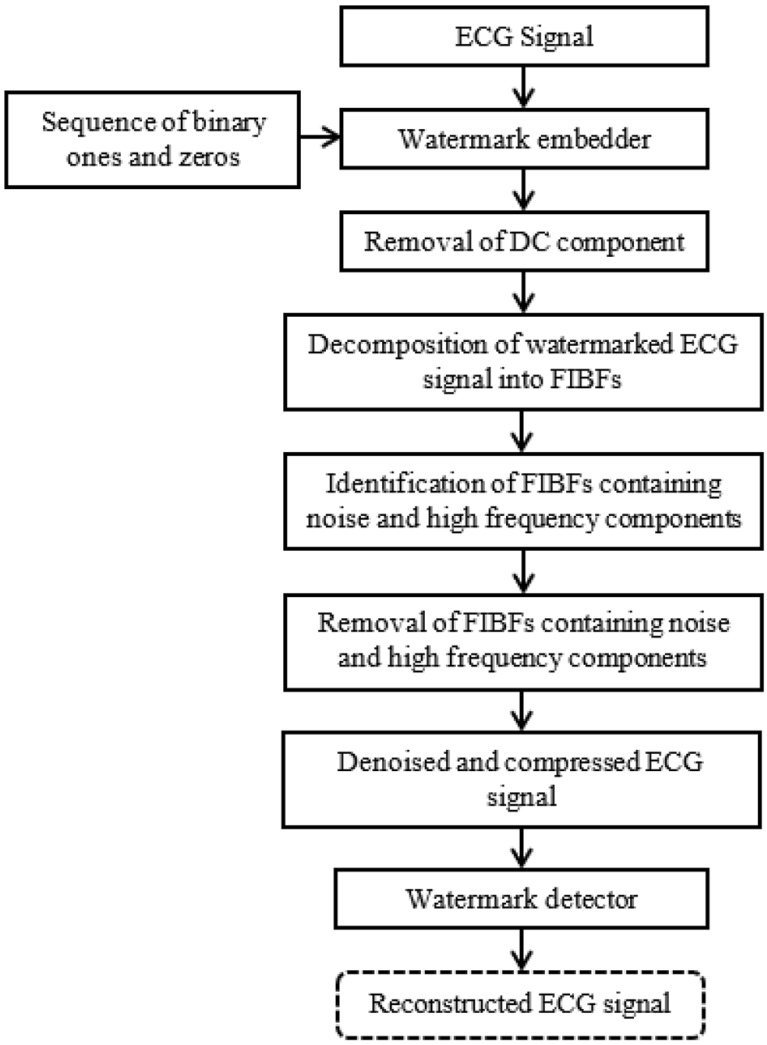


Fig. 2 Process flow diagram of the proposed method

The sampling frequency directly influences the accuracy of the QRS detection algorithm [2]. The American Heart Association Task Force [30] suggested that a low sampling frequency may produce jitter in detecting R peak, altering the spectrum. So, the selection of sampling frequency must be proper. Ajdaraga and Gusev [2] show that a sampling frequency of 120 Hz or higher yields sufficient QRS detection accuracy.

3 Proposed method

The MIT-BIH arrhythmia database is used for the validation of the proposed method. A record from the MIT-BIH arrhythmia database is used as a representative measure of the heart's electrical activity. With the help of digital watermarking, the ECG signal is protected from intruders. Figure 2 shows the process flow diagram of the proposed watermarking method using FDM. The watermarks in the proposed method are embedded in the ECG signal. A sequence of binary *ones* and *zeros* are added at even and odd-numbered samples of the ECG signal to produce a watermarked ECG signal. This watermarking does not alter the shape of the ECG signal. Hence the diagnosability is not influenced by watermarking. The watermarked ECG signal is further processed with the FDM. The steps used to generate analytic FIBFs are detailed in an algorithm as below.

Algorithm 1: The algorithm to decompose a signal into AFIBFs for $i = 1, 2, \cdots M$ and $N_0 = 0$, $N_M = \left(\frac{N}{2} - 1\right)$

Input: $w[n]$ a watermarked ECG signal with a length N

1 Obtain $W[k] = FFT\{w[n]\}$;

2 Set Analytic FIBFs, $\text{AFIBF}_i = b_i[n]\exp(j\theta_i[n]) = \sum_{k=N_{i-1}+1}^{N_i} W[k]e^{\left(\frac{j2\pi kn}{N}\right)}$;

3 For minimum number of FIBFs obtain the maximum value of N_i such that $(N_{i-1} + 1) \le N_i \le \left(\frac{N}{2} - 1\right)$ and phase $\theta_i[n]$ is a monotonically increasing function;

4 Obtain instantaneous frequency $\omega_i[n] = \left(\frac{\theta_i[n+1]-\theta_i[n-1]}{2}\right) \ge 0, \forall n$ and $b_i[n] \ge 0, \forall n \text{ and } i = 1, 2....., M$

Output: $\{AFIBF_i[n], \omega_i[n]\}_{i=1}^{M}$

The FDM technique using discrete Fourier transform (DFT) based zero-phase filter bank (ZPF) is shown in Fig. 3. The FDM decomposes the watermarked ECG signal $w[n]$ into a set of orthogonal FIBFs $w_i[n]$ with desirable frequency bands as in Eq. (1) [40]:

$$w[n] = d_0 + \sum_{i=1}^{M} w_i[n] \tag{1}$$

Here d_0 is the mean value of watermarked ECG signal $w[n]$, $w_i[n]$ represents M orthogonal FIBFs with $i = 1, 2, 3, \cdots M$. The frequency response of the i^{th} filter in a zero-phase filter bank is defined in Eq. (2) [40].

$$H_i[k] = \begin{cases} 1, \; k_{i-1} + 1 \le k \le k_i, \; \left(N - k_i\right) \le k \le \left(N - k_{i-1} - 1\right) \\ 0, \quad otherwise \end{cases} \tag{2}$$

Here N represents the length of the signal, $K_0 = 0$, and $K_M = N/2$. $w_i[n]$ are found by taking inverse DFT as defined in Eq. (3) [40].

$$w_i[n] = \sum_{k=0}^{N-1} \left[H_i[k]W[k]\exp^{\left(\frac{j2\pi kn}{N}\right)}\right] \tag{3}$$

In the proposed method, FDM decomposes the watermarked ECG signal into three FIBFs. The second FIBF corresponding to the high-frequency components of the signal is discarded. The combination of the remaining two FIBFs provides the watermarked reconstructed ECG signal. The FDM removes noises and artifacts present in the ECG signal. The FDM improves the output signal-to-noise ratio (SNR_{out}) and provides data compression by removing the redundant information present in the original signal. In the last step, the reconstructed original ECG signal is obtained from the compressed ECG signal by removing the watermarked sequence embedded in the original ECG signal.

The fast Fourier transform is used to compute the DFT of the signal, and the computational complexity of the proposed method is equal to $N\left(log_2 N\right)$. Hence the proposed method has the advantage of low computation complexity.

4 Results and discussion

The proposed method is implemented in MATLAB® R2020. Output signal to noise ratio (SNR_{out}), percentage root mean square difference (PRD), compression ratio (CR), and computational time (CT) is used to evaluate the performance of the proposed method. The evaluation parameters are defined in Eq. (4) through Eq. (6) [24, 26, 40].

$$SNR_{out} = 10\log_{10}\left(\frac{\sum_{n=1}^{N} Y^2[n]}{\sum_{n=1}^{N} (X[n]-Y[n])^2}\right) \tag{4}$$

$$PRD = \sqrt{\frac{\sum_{n=1}^{N} (X[n]-Y[n])^2}{\sum_{n=1}^{N} X^2[n]}} \times 100 \tag{5}$$

$$CR = \frac{(Data\ size\ of\ ECG\ signal\ before\ compression)}{(Data\ size\ of\ ECG\ signal\ after\ compression)} \tag{6}$$

Here, *X(n)* and *Y(n)* represent the original ECG signal and denoised ECG signal. The SNR_{out} is a measure of the quality of the output signal. A large value of SNR_{out} reflects the better quality of the output signal. PRD quantifies the quality of the reconstructed signal by computing the difference between the original and reconstructed signal. A lower value of PRD is required to show the efficacy of the method in signal reconstruction. The CR is the measure of relative reduction in the data size.

All the evaluation parameters described above are evaluated using a machine with configuration: Intel(R) Core i3-4030U CPU @ 1.90 GHz, 1.90 GHz, and 12 GB RAM. Records from the MIT-BIH arrhythmia database are used to validate the efficacy of the proposed method. Figure 4 shows the simulation results obtained from the proposed

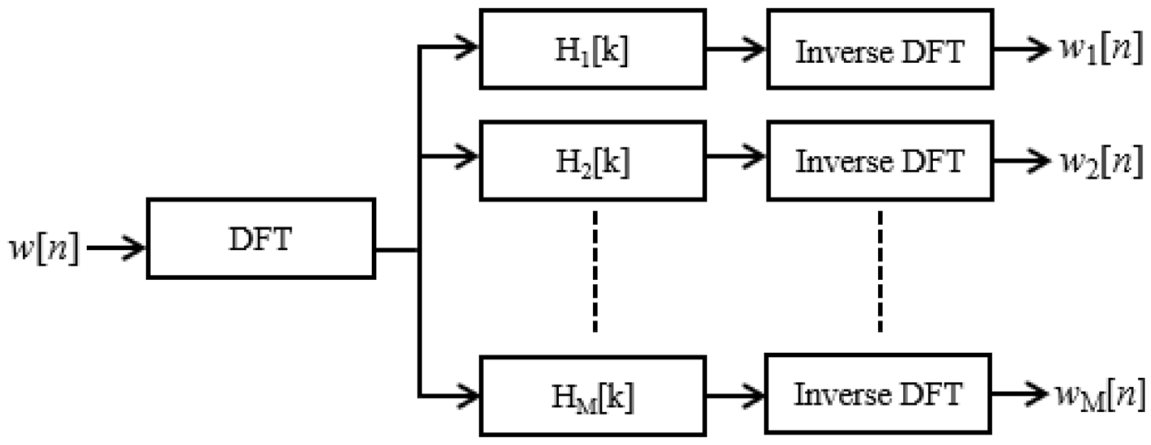


Fig. 3 A schematic representation of DFT based FDM

method. Record number 109 from the MIT-BIH arrhythmia database, selected as an ECG input, is shown in Fig. 4a. The watermark and compressed version of record number 109 is shown in Fig. 4b and c, respectively. Figure 4d depicts the reconstructed ECG signal as a replica of the input ECG signal. It is clear from Fig. 4d that the proposed method provides lossless compression.

Table 2. shows the value of output PRD (PRD_{out}) obtained at different values of input PRD (PRD_{in}) for some ECG signal records of the MIT-BIH arrhythmia database. From

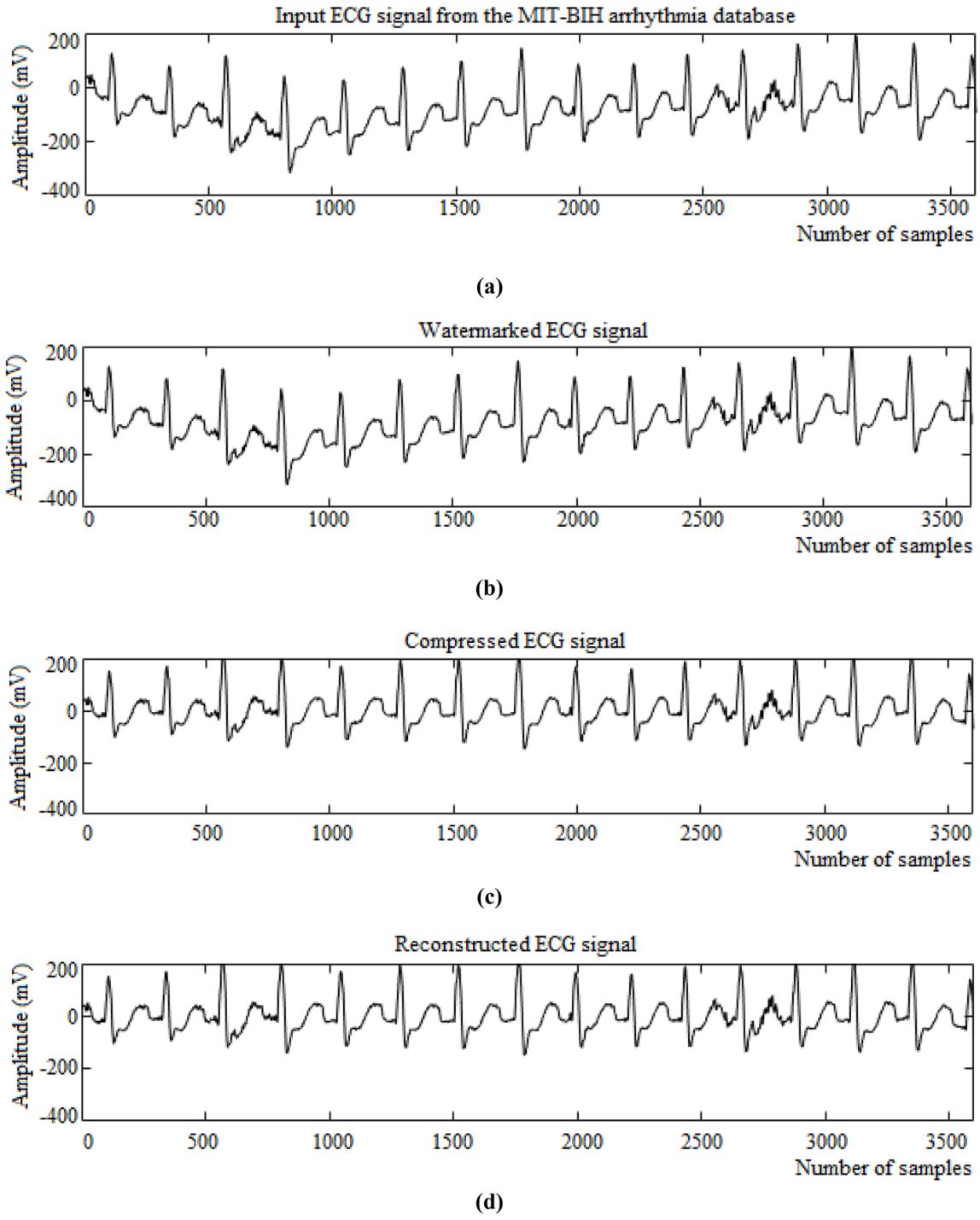


Fig. 4 **a** Input ECG signal (MIT-BIH– 109) **b** ECG signal with watermark **c** compressed ECG signal **d** reconstructed output ECG signal

Table 2 PRD_{out} of various MIT-BIH ECG records as a function of SNR values

ECG record	PRD_{out} (%)				
	− 10 dB	− 5 dB	0 dB	5 dB	10 dB
100	0.531	0.507	0.431	0.328	0.3
101	0.668	0.618	0.457	0.349	0.311
103	0.471	0.448	0.369	0.279	0.254
104	0.338	0.311	0.237	0.21	0.187
105	0.421	0.405	0.371	0.298	0.258
109	0.452	0.419	0.372	0.301	0.271
112	0.515	0.498	0.364	0.286	0.257
116	0.437	0.413	0.375	0.294	0.264
117	0.446	0.42	0.373	0.319	0.292
118	0.512	0.503	0.468	0.403	0.384
205	0.627	0.612	0.563	0.521	0.5
213	0.635	0.613	0.584	0.536	0.504
221	0.548	0.536	0.445	0.401	0.37
231	0.689	0.614	0.541	0.5	0.468
234	0.624	0.592	0.423	0.386	0.342

Table 2, the proposed method provides low PRD_{out} values at different SNR_{in}. For record 104, the value of PRD_{out} is 0.338% and 0.187% at − 10 dB and 10 dB SNR_{in}, respectively. Record 116 has output PRD values of 0.375% and 0.294% at 0 dB and 5 dB, respectively.

Similarly, the output signal-to-noise ratio (SNR_{out}) for various ECG records at different SNR_{in} is tabulated in Table 3. The proposed method showed good signal-to-noise ratio improvement for various ECG records at different SNR_{in} (Table 4). For record 221, SNR_{out} is 19.1 dB and 30.8 dB at − 5 dB and 10 dB SNR_{in}. Record 117 has a 19.4 dB output signal-to-noise ratio at − 5 dB SNR_{in}.

Table 3 Output SNR obtained at different values of SNR_{in}

ECG record	SNR_{out} (dB)				
	− 10 dB	− 5 dB	0 dB	5 dB	10 dB
100	13.15	18.2	23.3	27.5	30.4
101	10.7	18.3	24.5	27.1	29.8
103	12.8	17.9	23.1	27.9	30.8
104	11.6	19	23.9	27.6	29.9
105	10.6	18.7	24.2	27.4	30.4
109	12.2	18.3	23.7	27.5	30.7
112	13.5	19.2	24.1	27.7	29.9
116	13.1	18.8	23.6	28.1	30.5
117	9.6	19.4	23.1	26.9	30.6
118	12.9	17.8	23.9	27.8	30.9
205	10.4	18.7	24.3	28	30.4
213	11.8	18.5	23.8	27.7	29.8
221	9.7	19.1	23.4	27.2	30.8
231	10.9	18.8	23.7	28.1	30.7
234	12	19.3	22.9	27.9	30.4

Table 4 Output SNR obtained at different values of SNR_{in}

ECG record	SNR_{out} (dB)				
	− 10 dB	− 5 dB	0 dB	5 dB	10 dB
Muscle contraction corrupted ECG data	11.7	18.9	23.3	28.7	31
Motion artifact corrupted ECG data	10.4	17.9	24.2	26.8	30.4
Electromyographic noise corrupted ECG data	12	16.4	22.3	27.6	30.8

The compression ratio is used to measure the compression ability of the proposed method. The compression ratio (CR) for various ECG records is shown in Table 5.

The proposed method provides an excellent compression ratio. For record 101, the compression ratio is 35.2, while 33.6 is the compression ratio for record 118. The computation time is an indicative measurement of the complexity of the proposed method. Table 5 depicts the computation time for the ECG record. For record 205, the computation time is 12.935 s, while record 221 has a computation time of 12.033 s.

Furthermore, a comparison of PRD and CR of the proposed method with other techniques is presented in Table 6. Table 6 shows that the proposed method provides a high compression ratio along with low PRD values. For record 100 Fathi et al. [11] achieved a CR of 31.42 with a PRD of 3.21%, while Pandey et al. [34] obtained a CR of 31.42 with a PRD of 1%. The proposed method provides a CR of 34.1 with a PRD of only 0.3%.

Similarly, for record 101, the proposed method provides a CR of 35.2 with a PRD of 0.311%. In comparison, the methods in [11] and [34] provide CR of 28.47 (PRD-3.41%) and 19.21 (PRD-1.19%), respectively. The proposed method provides a compression ratio ranging from 30.8 to 35.2, while the PRD value ranges from 0.254 to 0.311%. The other existing method proposed by Fathi et al. [11] provides a compression ratio ranging from 26.56 to 31.42 with a PRD value ranging from 3.07 to 5.34%. The method suggested by Pandey et al. [34] offers a compression ratio ranging from 15.6 to 22.08 while the PRD value ranges from 1 to 1.19%.

Table 5 Compression ratio and computation time for ECG records

ECG record	Compression ratio (CR)	Computation time (s)
100	34.1	13.271
101	35.2	13.660
103	30.8	13.044
104	33.7	13.480
105	33.7	13.560
109	32.8	13.465
112	31.4	13.727
116	30.9	13.384
117	32	13.308
118	33.6	13.499
205	34	12.935
213	31.5	12.342
221	31	12.033
231	32.7	12.167
234	31.8	12.166

Table 6 Comparing the results of the proposed method and some existing methods

Record	Fathi et al. [11]		Pandey et al. [34]		Proposed method	
	CR	PRD (%)	CR	PRD (%)	CR	PRD (%)
100	31.42	3.21	16.87	1	34.1	0.3
101	28.47	3.41	19.21	1.19	35.2	0.311
103	26.56	5.34	15.6	1.01	30.8	0.254
117	29.34	3.07	22.08	1.01	32	0.292

The above results show the efficacy of the proposed method. The high compression ratio of the proposed method shows the high compression ability of the proposed method, and a lower value of PRD reflects the lossless nature of the proposed method.

5 Conclusions

This paper proposes an FDM-based technique to analyze ECG watermarking and compression. The proposed method can secure the transmission of ECG signals and reduce the transmission bandwidth while preserving the shape of the ECG signal waveform. The proposed method effectively removes the noise components and improves the quality of the reconstructed ECG signal. The efficacy of the proposed method is validated by using ECG records from the popular and widely used MIT-BIH arrhythmia database. The proposed method provides low PRD, high SNR_{out}, and high compression ratio compared to existing techniques. Hence, the proposed watermarking and compression technique using FDM provides an efficient, reliable, and secure way to transfer the ECG signals by minimizing the losses in the original ECG signal.

In the future, studies on the proposed work involve comparing the computational complexity with other methods and watermarking and compression of morphological features of an ECG signal. Further, it is required to explore the compression method's efficiency on the long-term and real-time ECG signals. The studies can be extended to other biomedical signals like EEG, photoplethysmogram, electroneurogram.

References

1. Act A (1996) Health insurance portability and accountability act of 1996. Public law 104:191
2. Ajdaraga E, Gusev M (2017) Analysis of sampling frequency and resolution in ECG signals. In: 2017 25th Telecommunication Forum (TELFOR), p. 1–4. IEEE, 2017
3. Al-Nashash HAM (1995) A dynamic Fourier series for the compression of ECG using FFT and adaptive coefficient estimation. Med Eng Phys 17(3):197–203
4. American Heart Association, AHA Database, ECRI, Butler Pike, Plymouth Meeting, PA, USA
5. Averkiou M (2015) Digital watermarking. Department of Computer Science University of Cyprus
6. Batista LV, Melcher EUK, Carvalho LC (2001) Compression of ECG signals by optimized quantization of discrete cosine transform coefficients. Med Eng Phys 23(2):127–134
7. Bousseljot R, Kreiseler D, Schnabel A (1995) Nutzung der EKG-Signaldatenbank CARDIODAT der PTB über das Internet. Biomed Tech 1995:317–318
8. Chin W-L, Chang C-C, Tseng C-L, Huang Y-Z, Jiang T (2019) Bayesian real-time QRS complex detector for healthcare system. IEEE Internet Things J 6(3):5540–5549

9. Cox I, Miller M, Bloom J, Fridrich J, Kalker T (2007) Digital watermarking and steganography. Morgan Kaufmann
10. El B'charri O, Latif R, Jenkal W, Abenaou A (2016) The ECG signal compression using an efficient algorithm based on the DWT. Int J Adv Comput Sci Appl 1(7):181–187
11. Fathi A, Faraji-Kheirabadi F (2016) ECG compression method based on adaptive quantization of main wavelet packet subbands. SIViP 10(8):1433–1440
12. Fatimah B, Singh P, Singhal A, Pachori RB (2020) Detection of apnea events from ECG segments using Fourier decomposition method. Biomed Signal Process Control 61:102005
13. Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PC, Mark RG, Mietus JE, Moody GB, Peng CK, Stanley HE (2000) PhysioBank, PhysioToolkit, and PhysioNet – components of a new research resource for complex physiologic signals. Circulation 101:E215–E220
14. Gutiérrez-Rivas R, García JJ, Marnane WP, Hernández A (2015) Novel real-time low-complexity QRS complex detector based on adaptive thresholding. IEEE Sens J 15(10):6036–6043
15. Hammad M, Maher A, Wang K, Jiang F, Amrani M (2018) Detection of abnormal heart conditions based on characteristics of ECG signals. Measurement 125:634–644
16. Hao H, Wang H, Rehman N, Chen L (2017) An improved multivariate wavelet denoising method using subspace projection. IEICE Trans Fund Electron Commun Comput Sci 100(3):769–775
17. Iyengar N, Peng CK, Morin R, Goldberger AL, Lipsitz LA (1996) Age-related alterations in the fractal scaling of cardiac interbeat interval dynamics. Am J Physiology 271(4):R1078–R1084
18. Jalaleddine SMS, Hutchens CG, Strattan RD, Coberly WA (1990) ECG data compression techniques-a unified approach. IEEE Trans Biomed Eng 37(4):329–343
19. Jero SE, Ramu P, Ramakrishnan S (2014) Discrete wavelet transform and singular value decomposition-based ECG steganography for secured patient information transmission. J R Med Sys 38(10):1–11
20. Jero SE, Ramu P, Ramakrishnan S (2015) ECG steganography using curvelet transform. Biomed Signal Process Control 22:161–169
21. Kaur, S, R Singhal, O Farooq, and BS Ahuja (2010) Digital watermarking of ECG data for secure wireless communication. In: 2010 international conference on recent trends in information, telecommunication and computing, pp. 140–144. IEEE, 2010
22. Kirovski D, Malvar HS (2003) Spread-spectrum watermarking of audio signals. IEEE Trans Signal Process 51(4):1020–1033
23. Kumar A, Komaragiri R, Kumar M (2018) From pacemaker to wearable: techniques for ECG detection systems. J Med Syst 42(2):1–17
24. Kumar A, Komaragiri R, Kumar M (2019) Time-frequency localization using three-tap biorthogonal wavelet filter bank for electrocardiogram compressions. Biomed Eng Lett 9(3):407–411
25. Kumar A, Komaragiri R, Kumar M (2019) Design of efficient fractional operator for ECG signal detection in implantable cardiac pacemaker systems. Int J Circuit Theory Appl 47(9):1459–1476
26. Kumar A, Tomar H, Mehla VK, Komaragiri R, Kumar M (2020) Stationary wavelet transform based ECG signal denoising method. ISA Ttrans 114:251
27. Kumar A, Ranganatham R, Singh S, Komaragiri R, Kumar M (2021) A robust digital ECG signal watermarking and compression using biorthogonal wavelet transform. Res Biomed Eng 37(1):79–85
28. Laguna P, Mark RG, Goldberg A, Moody GB (1997) A database for evaluation of algorithms for measurement of QT and other waveform intervals in the ECG. In: Computers in cardiology, pp 673–676. IEEE, 1997
29. Lee W, Kim S, Kim D (2018) Individual biometric identification using multi-cycle electrocardiographic waveform patterns. Sensors 18(4):1005
30. Malik M (1996) Task force of the European society of cardiology and the North American society of pacing and electrophysiology. Heart rate variability. Standards of measurement, physiological interpretation, and clinical use. Eur Heart J 17:354–381
31. Manikandan MS, Dandapat S (2014) Wavelet-based electrocardiogram signal compression methods and their performances: a prospective review. Biomed Signal Process Control 14:73–107
32. Marcel S, del José R, Millán. (2007) Person authentication using brainwaves (EEG) and maximum a posteriori model adaptation. IEEE Trans Pattern Anal Machine Intell 29(4):743–752
33. Moody GB, Mark RG (2001) The impact of the MIT-BIH arrhythmia database. IEEE Eng Med Biol Mag 20(3):45–50
34. Pandey A, Singh B, Saini BS, Sood N (2016) A joint application of optimal threshold based discrete cosine transform and ASCII encoding for ECG data compression with its inherent encryption. Australas Phys Eng Sci Med 39(4):833–855
35. Olmos S, MillAn M, Garcia J, Laguna P (1996) ECG data compression with the Karhunen-Loeve transform. In: Computers in cardiology 1996, pp. 253–256. IEEE, 1996

36. Rhee MY (2003) Internet security: cryptographic principles, algorithms, and protocols. Wiley, New York
37. Sanivarapu PV, Rajesh KNVPS, Rajasekhar Reddy NV, Sekhar Reddy N (2020) Patient data hiding into ECG signal using watermarking in transform domain. Phys Eng Sci Med 43(1):213–226
38. Sellami A, Zouaghi A, Daamouche A (2017) ECG as a biometric for individual's identification. In: 2017 5th International Conference on Electrical Engineering-Boumerdes (ICEE-B), pp. 1–6. IEEE, 2017
39. Singh P, Joshi SD, Patney RK, Saha K (2017) The Fourier decomposition method for nonlinear and nonstationary time series analysis. Proc R Soc A 473(2199):20160871
40. Singhal A, Singh P, Fatimah B, Pachori RB (2020) An efficient removal of power-line interference and baseline wander from ECG signals by employing Fourier decomposition technique. Biomed Signal Process Control 57:101741
41. Taddei A, Distante G, Emdin M, Pisani P, Moody GB, Zeelenberg C, Marchesi C (1992) The European ST-T database: standard for evaluating systems for the analysis of ST-T changes in ambulatory electrocardiography. Eur Heart J 13(9):1164–1172
42. Tseng K-K, He X, Kung W-M, Chen S-T, Liao M, Huang H-N (2014) Wavelet-based watermarking and compression for ECG signals with verification evaluation. Sensors 14(2):3721–3736
43. Zhang B, Zhao J, Chen X, Wu J (2017) ECG data compression using a neural network model based on multi-objective optimization. PLoS ONE 12(10):e0182500
44. Zou D, Shi YQ, Ni Z, Su W (2006) A semi-fragile lossless digital watermarking scheme based on integer wavelet transform. IEEE Trans Circuits Syst Video Technol 16(10):1294–1300

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Prashant Mani Tripathi[1] · Ashish Kumar[2] · Rama Komaragiri[1] · Manjeet Kumar[3]

Prashant Mani Tripathi
pmanitripathi@gmail.com

Ashish Kumar
akumar.1june@gmail.com

Rama Komaragiri
rama.komaragiri@gmail.com

[1] Department of Electronics and Communication Engineering, Bennett University, Greater Noida, India

[2] School of Electronics Engineering, Vellore Institute of Technology, Chennai, Tamil Nadu, India

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