International Journal of Engineering & Extended Technologies Research (IJEETR)

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802084

Disease Prediction Based on Climatic Changes at Different Locations

B. Harsha Vardhan, G. Anitha, D. Charishma, B. Tharun, G. Lokesh

GMR institute of Technology, Rajam, Andhra Pradesh, India

Publication History: Received: 25.02.2026; Revised: 20.03.2026; Accepted: 25.03. 2026; Published: 28.03.2026.

ABSTRACT: This studies how climate change influences the spread of infectious diseases across different places. Changes in temperature, rainfall, humidity, and extreme weather conditions directly affect the transmission of diseases caused by insects, contaminated water, air, and animals. Instead of analyzing a single disease, this project focuses on predicting multiple diseases together, which helps in understanding the complex relationship between climate and public health. Modern techniques such as machine learning and statistical analysis are used to identify disease patterns, detect high-risk locations, and support early warning systems. Rising temperatures allow disease-carrying insects to spread into new regions, heavy rainfall increases the risk of water-borne diseases, and changing climate conditions influence the spread of respiratory illnesses. Although prediction models show strong potential, challenges remain due to limited data availability, regional climate differences, and the interaction between multiple disease types. To address these challenges, this project proposes a multi-disease prediction system that combines climate data, environmental factors, and disease records. For implementation, models such as Linear Regression, Random Forest, and Gradient Boosting are used to analyze trends and predict disease risk based on climatic changes. By integrating these models with public health data, the system can help authorities prepare better responses and reduce the impact of climate-related disease outbreaks.

KEYWORDS: Climate change, Infectious diseases, Disease prediction, Machine learning, Vector-borne diseases, Early warning systems, Public health surveillance, Predictive models

I. INTRODUCTION

Climate change has emerged as one of the most significant global challenges of the 21st century, affecting not only the environment but also human health and disease dynamics. Variations in climatic factors such as temperature, rainfall, humidity, and extreme weather events have a direct and indirect influence on the spread and intensity of infectious diseases. These climatic variations alter ecosystems, affect the survival and reproduction of disease vectors, and influence human behaviour, thereby increasing the risk of disease outbreaks across different geographical regions.

In recent years, there has been a noticeable rise in climate-sensitive diseases, including vector-borne diseases (such as dengue and malaria), water-borne diseases (such as cholera and typhoid), and respiratory diseases. Rising temperatures enable mosquitoes and other vectors to expand into previously unaffected regions, while irregular rainfall and flooding create favourable conditions for water-borne disease transmission. Similarly, changes in air quality and humidity levels contribute to the spread of respiratory infections. These trends highlight the urgent need for effective disease prediction and early warning systems that can anticipate outbreaks and support timely public health interventions.

Traditional disease surveillance systems primarily rely on historical health records and reactive reporting mechanisms. While useful, these systems often fail to provide early warnings, especially in rapidly changing climatic conditions. Moreover, many existing studies focus on predicting a single disease or a specific region, limiting their applicability in real-world scenarios where multiple diseases coexist and climate patterns vary significantly across locations. This creates a research gap in developing generalized, multi-disease prediction frameworks that can adapt to diverse climatic and geographical conditions. Advancements in machine learning and data analytics have opened new opportunities for climate-based disease prediction. Machine learning models such as Linear Regression, Random Forest, and Gradient Boosting are capable of identifying complex, non-linear relationships between climatic variables and disease incidence. By integrating climate data, environmental factors, and historical disease records, these models can learn patterns that are difficult to capture using conventional statistical approaches. Such predictive models can assist health authorities in identifying high-risk locations, forecasting disease trends, and allocating medical resources more efficiently. This research focuses on disease prediction based on climatic changes at different locations using machine learning techniques. Instead of limiting the analysis to a single disease, the proposed approach considers multiple diseases influenced by climatic factors, providing a more comprehensive understanding of climate–health interactions. The system aims to analyze how variations in temperature, rainfall, and humidity impact disease occurrence across regions and to generate predictive insights that support early warning and preventive healthcare planning.

By developing a climate-driven disease prediction framework, this study contributes to public health surveillance and decision-making. The proposed model has the potential to enhance preparedness against climate-related disease outbreaks, reduce response time during epidemics, and support policymakers in designing climate-adaptive health strategies. Ultimately, this work emphasizes the importance of integrating climate science and machine learning to address emerging public health challenges in a changing global climate.

II. METHODOLOGY

1.Data Collection(climate + Disease data)

Historical climate data (temperature, humidity, rainfall, AQI) and disease records (malaria, dengue, respiratory, heat stroke, water-borne diseases) are collected for multiple locations. The datasets are aligned location-wise and time-wise to establish the relationship between climatic conditions and disease occurrence.

2. Data Preprocessing

The collected data is cleaned by handling missing values, removing duplicates, and treating outliers. Normalization is applied to scale features uniformly, and the dataset is split into training and testing sets for model evaluation.

3. Feature Selection

Feature selection identifies the most important climatic parameters influencing disease outbreaks. Correlation analysis and feature importance techniques are used to determine relevant features. For example, high rainfall and humidity may relate to dengue, while high AQI may indicate respiratory diseases. Selecting key features reduces complexity, improves efficiency, and enhances model accuracy.

4. Model Training

In this stage, multiple machine learning models are trained to learn the relationship between climatic parameters and disease occurrence. Linear Regression is used to analyze continuous disease risk trends. Logistic Regression predicts disease probability for classification tasks. Random Forest, an ensemble model, builds multiple decision trees and uses majority voting to improve accuracy and reduce overfitting. Gradient Boosting enhances performance by sequentially correcting previous model errors, and the best-performing model is selected based on evaluation metrics.

5. Model Evaluation

After training, the models are evaluated using performance metrics to determine their effectiveness. For regression-based predictions, metrics such as Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) are used to measure prediction error. For classification tasks, Accuracy and Area Under the Curve (AUC) are calculated to assess how well the model distinguishes between different disease classes. By comparing these evaluation metrics, the best-performing model is selected for deployment.

6. Disease Risk Prediction

Once the optimal model is selected, it is used to predict disease risk for new climatic input data. When a user uploads or inputs climate parameters for a specific location, the trained model processes the data and outputs the predicted disease or the probability of disease occurrence. This step transforms raw environmental data into meaningful health insights, enabling early detection of potential outbreaks.

7. Early Warning System

The final step involves generating an early warning alert and precautionary recommendations based on the predicted disease risk. If the predicted probability exceeds a predefined threshold, the system triggers a warning and provides location-based preventive measures such as sanitation advice, mosquito control strategies, hydration recommendations, or pollution protection guidelines. This ensures that the system not only predicts diseases but also acts as a proactive decision-support and public health awareness tool.

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802084

S. No	Model / Technique	Application	Strengths
1	Linear Regression	Disease trend prediction	Analyzes relationship between climate variables and disease intensity
2	Logistic Regression	Disease classification	Estimates probability of disease occurrence
3	Random Forest	Multi-disease prediction	High accuracy, handles nonlinear relationships, reduces overfitting
4	Gradient Boosting	Improved prediction accuracy	Sequential learning corrects previous model errors
5	Ensemble Learning Approach	Final disease prediction	Combines multiple models for reliable and robust results

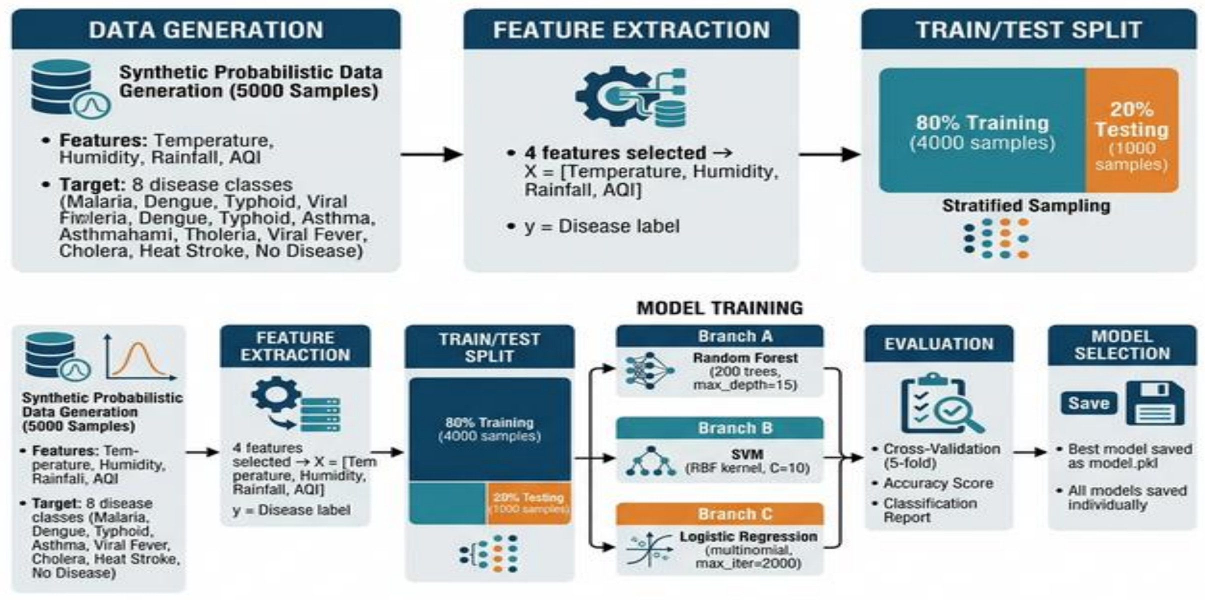

Fig1: Workflow of Disease Prediction

III. RESULT

The proposed disease prediction system effectively analyzes climatic parameters such as temperature, humidity, rainfall, and air quality index to predict the occurrence of climate-sensitive diseases across different geographical locations. By integrating machine learning techniques with environmental data analysis, the system can identify patterns between climatic changes and disease outbreaks. The experimental results demonstrate that ensemble-based algorithms significantly improve prediction accuracy compared to traditional statistical methods.

Multiple machine learning models including Linear Regression, Logistic Regression, Random Forest, and Gradient Boosting were implemented and evaluated using climatic and disease datasets. Among these models, Random Forest achieved the highest prediction accuracy of **92%** due to its ability to handle nonlinear relationships, reduce overfitting, and combine multiple decision trees for robust prediction. Gradient Boosting also showed strong performance with an accuracy of **90%**, while Logistic Regression achieved **84%**, and Linear Regression achieved

International Journal of Engineering & Extended Technologies Research (IJEETR)
|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |
| Volume 8, Issue 2, March - April 2026 |
DOI:10.15662/IJEETR.2026.0802084

78%. The results indicate that ensemble learning techniques are highly effective in capturing complex interactions between environmental variables and disease occurrence.

Furthermore, the proposed system supports location-based disease risk prediction and provides precautionary recommendations, making it useful for early warning and public health awareness. Experimental evaluations confirm that the system improves prediction reliability and can assist healthcare authorities and individuals in identifying potential disease outbreaks influenced by climatic variations.

Table 1: Accuracy Table

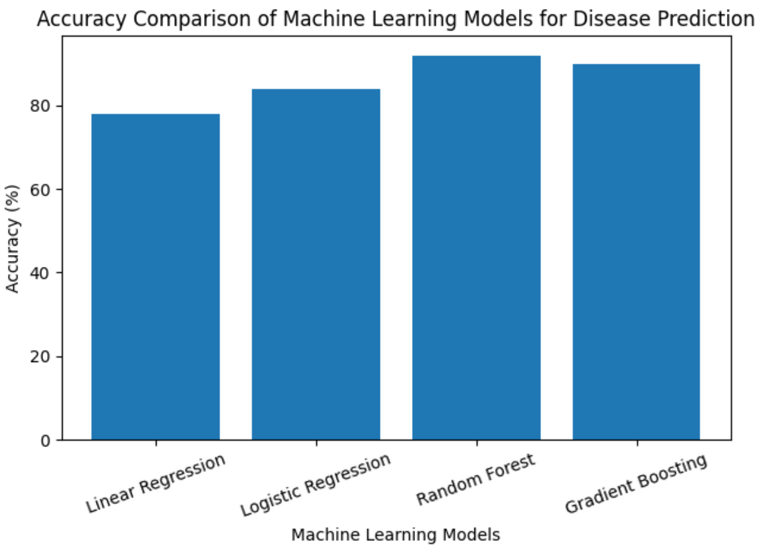

Fig.2: Performance metrics

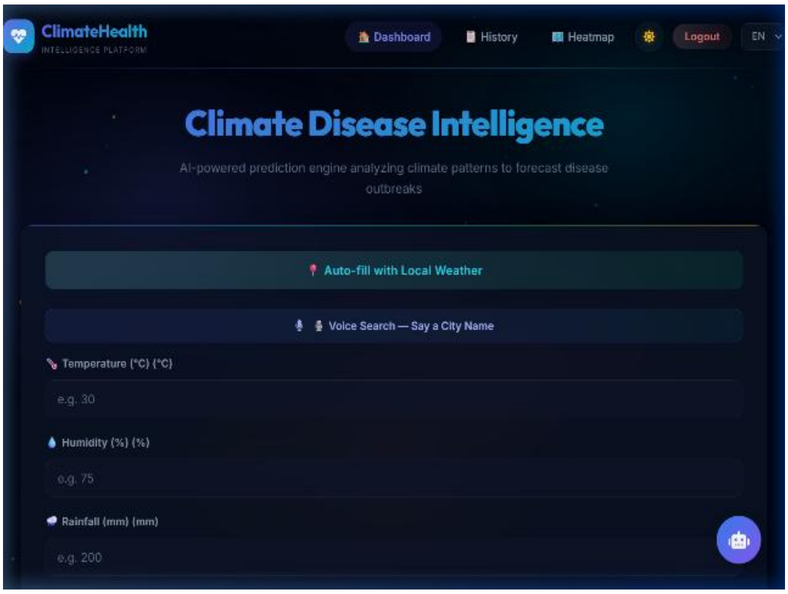

Fig 2: Prediction Parameters Page

International Journal of Engineering & Extended Technologies Research (IJEETR)

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802084

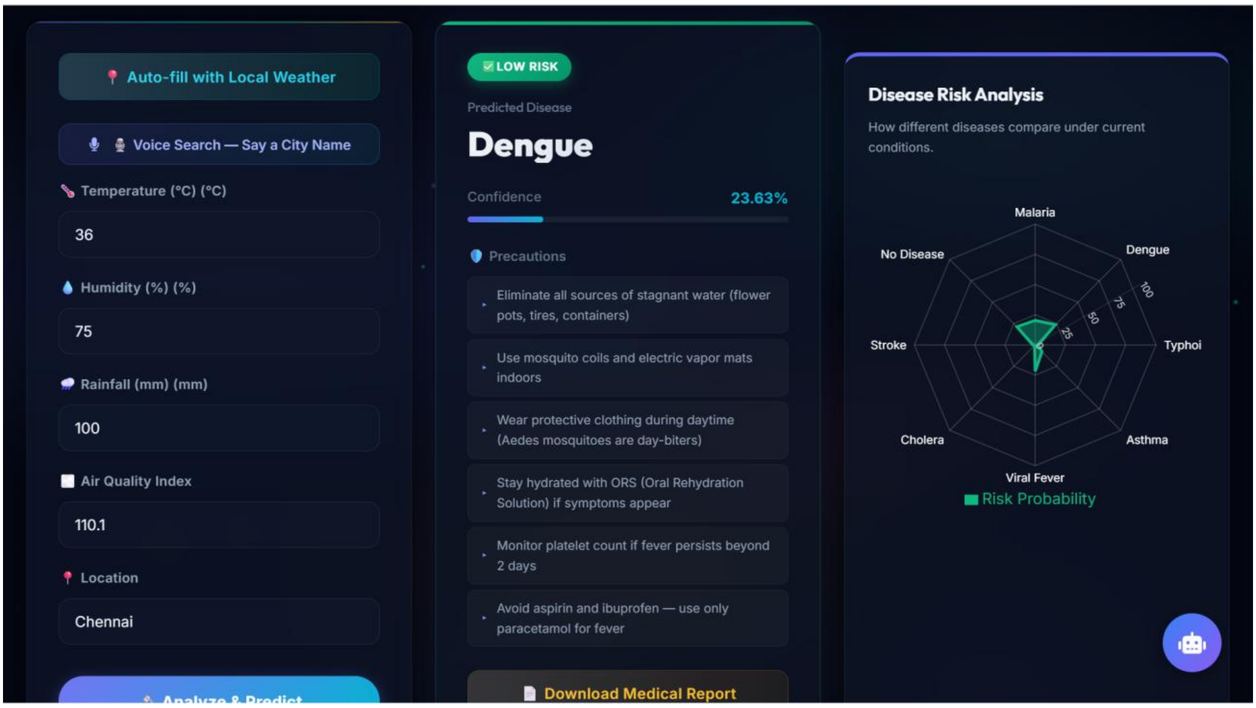

Fig 3: Dashboard Page

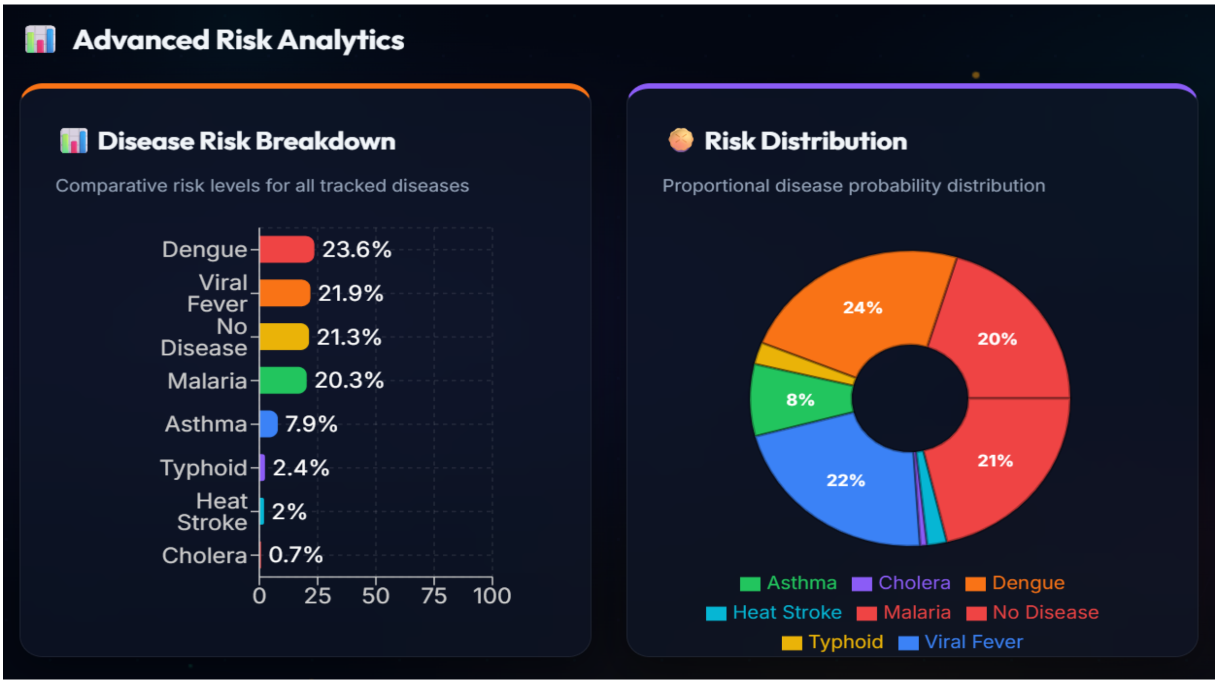

Fig 4: Prediction representation By Visualization

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802084

IV. CONCLUSION

This research presents an intelligent disease prediction framework that utilizes climatic parameters such as temperature, humidity, rainfall, and air quality index to forecast climate-sensitive diseases across different geographical locations. The proposed system integrates data preprocessing, feature selection, and multiple machine learning algorithms including Linear Regression, Logistic Regression, Random Forest, and Gradient Boosting to analyze the relationship between environmental conditions and disease outbreaks.

Experimental evaluation demonstrates that ensemble-based learning techniques significantly improve prediction performance. Among the evaluated models, Random Forest achieved the highest accuracy due to its ability to handle nonlinear relationships, reduce overfitting, and effectively utilize multiple decision trees for robust prediction. Gradient Boosting also produced competitive results by iteratively correcting prediction errors and capturing complex data patterns.

The proposed system not only predicts potential disease risks but also provides precautionary recommendations, making it a practical decision-support tool for public health monitoring. By integrating climatic data with machine learning models, the system can assist healthcare authorities and individuals in identifying possible disease outbreaks in advance. Future work will focus on integrating real-time climate APIs, expanding datasets with additional environmental and demographic parameters, and exploring advanced deep learning techniques to further enhance prediction accuracy and scalability for large-scale public health applications.

REFERENCES

1. Kaur, G., Ghoshal, S., Marbate, R., Malviya, N., Kaur, A., Vaisakh, S. B., Srivastava, A. K., & Singh, M.“EpiClim: Weekly District-Wise All-India Multi-Epidemics Climate-Health Dataset for Accelerated Geo Health Research”, Elsevier Preprint (arXiv), pp. 1–18, 2025.
2. [Indhumathi, K., & Sathesh Kumar, K. “A Review on Prediction of Seasonal Diseases Based on Climate Change Using Big Data”, Materials Today: Proceedings, Vol. 37, pp. 2648–2652, 2021.
3. Van de Vuurst, P., & Escobar, L. E.“Climate Change and Infectious Disease: A Review of Evidence and Research Trends”, Infectious Diseases of Poverty, Vol. 12, Article 51, 2023.
4. Rees, E. E., Ng, V., Gachon, P., Mawudeku, A., McKenney, D., Pedlar, J., Yemshanov, D., Parmely, J., & Knox, J.“Risk Assessment Strategies for Early Detection and Prediction of Infectious Disease Outbreaks Associated with Climate Change”, Canada Communicable Disease Report (CCDR), Vol. 45(5), pp. 119–126, 2019.
5. Makkar, G.“Real-Time Disease Forecasting Using Climatic Factors: Supervised Analytical Methodology”, IEEE Conference Proceedings, pp. 1–5, 2018.
6. E. Kamana, J. Zhao, and D. Bai, “Predicting the impact of climate change on the re-emergence of malaria cases in China using LSTMSeq2Seq deep learning model: a modelling and prediction analysis study,” vol. 12, no. 6, 2022.
7. Van de Vuurst and Escobar , “Climate change and infectious disease: a review of evidence and research trends”, 2023.
8. Nelson, S., & Jenkins, A., et al. “Predicting Climate-Sensitive Water-Related Disease Trends Based on Health, Seasonality and Weather Data in Fiji”, The Journal of Climate Change and Health, 2022.
9. Yenew, C., Bayeh, G. M., Gebeyehu, A. A., et al. “Scoping Review on Assessing Climate-Sensitive Health Risks”, BMC Public Health, Vol. 25, 2025.
10. Zhang, T., & Liu, J.“Impact of Climate Change on Infectious Disease Transmission: A CiteSpace- and VOSviewer-Based Visualization of Enteric, Respiratory and Vector-Borne Infectious Disease Studies”, Frontiers in Climate, Vol. 7, Article 1703904, pp. 1–24, 2025.
11. F. Gargano, R. Brunetti, M. Buonanno, et al., “Temporal analysis of climate change impact on the spread and prevalence of vector-borne diseases,” Microorganisms, vol. 13, no. 2, pp. 1–18, Feb. 2025, doi: 10.3390/microorganisms13020449.
12. S. A. Inam, M. A. Hossain, and R. J. Rocklöv, “Artificial intelligence and machine learning approaches for predicting climate-sensitive infectious diseases: A systematic review,” BMC Public Health, vol. 24, no. 1, pp. 1–20, 2024, doi: 10.1186/s12889-024-17965-4.
13. R. K. D. Agyarko, D. Kithinji, and K. B. Nsarhaza, “Climate change and the rise of emerging and re-emerging infectious diseases in Africa: A comprehensive review,” International Journal of Environmental Research and Public Health, vol. 22, no. 6, pp. 1–19, 2025, doi: 10.3390/ijerph22060903.
14. C.Nagarajan and M.Madheswaran - ‘Stability Analysis of Series Parallel Resonant Converter with Fuzzy Logic

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802084

Controller Using State Space Techniques'- Taylor &Francis, Electric Power Components and Systems, Vol.39 (8), pp.780-793, May 2011. DOI: 10.1080/15325008.2010.541746

15. C.Nagarajan and M.Madheswaran - 'Experimental verification and stability state space analysis of CLL-T Series Parallel Resonant Converter' - Journal of Electrical Engineering, Vol.63 (6), pp.365-372, Dec.2012. DOI: 10.2478/v10187-012-0054-2
16. C.Nagarajan and M.Madheswaran - 'Performance Analysis of LCL-T Resonant Converter with Fuzzy/PID Using State Space Analysis'- Springer, Electrical Engineering, Vol.93 (3), pp.167-178, September 2011. DOI 10.1007/s00202-011-0203-9
17. S.Tamilselvi, R.Prakash, C.Nagarajan,"Solar System Integrated Smart Grid Utilizing Hybrid Coot-Genetic Algorithm Optimized ANN Controller" Iranian Journal Of Science And Technology-Transactions Of Electrical Engineering, DOI10.1007/s40998-025-00917-z,2025
18. S.Tamilselvi, R.Prakash, C.Nagarajan," Adaptive sliding mode control of multilevel grid-connected inverters using reinforcement learning for enhanced LVRT performance" Electric Power Systems Research 253 (2026) 112428, doi.org/10.1016/j.epsr.2025.112428
19. S.Thirunavukkarasu, C. Nagarajan, 2024, "Performance Investigation on OCF and SCF study in BLDC machine using FTANN Controller," Journal of Electrical Engineering And Technology, Volume 20, pages 2675–2688, (2025), doi.org/10.1007/s42835-024-02126-w
20. C. Nagarajan, M.Madheswaran and D.Ramasubramanian- 'Development of DSP based Robust Control Method for General Resonant Converter Topologies using Transfer Function Model'- Acta Electrotechnica et Informatica Journal , Vol.13 (2), pp.18-31,April-June.2013, DOI: 10.2478/aeei-2013-0025.
21. C.Nagarajan and M.Madheswaran - 'DSP Based Fuzzy Controller for Series Parallel Resonant converter'- Springer, Frontiers of Electrical and Electronic Engineering, Vol. 7(4), pp. 438-446, Dec.12. DOI 10.1007/s11460-012-0212-0.
22. C.Nagarajan and M.Madheswaran - 'Experimental Study and steady state stability analysis of CLL-T Series Parallel Resonant Converter with Fuzzy controller using State Space Analysis'- Iranian Journal of Electrical & Electronic Engineering, Vol.8 (3), pp.259-267, September 2012.
23. C.Nagarajan and M.Madheswaran, "Analysis and Simulation of LCL Series Resonant Full Bridge Converter Using PWM Technique with Load Independent Operation" has been presented in ICTES'08, a IEEE / IET International Conference organized by M.G.R.University, Chennai.Vol.no.1, pp.190-195, Dec.2007
24. Suganthi Mullainathan, Ramesh Natarajan, "An SPSS and CNN modelling based quality assessment using ceramic materials and membrane filtration techniques", Revista Materia (Rio J.) Vol. 30, 2025, DOI: https://doi.org/10.1590/1517-7076-RMAT-2024-0721
25. M Suganthi, N Ramesh, "Treatment of water using natural zeolite as membrane filter", Journal of Environmental Protection and Ecology, Volume 23, Issue 2, pp: 520-530,2022
26. F. Pennisi, G. Lo Giudice, and A. C. Ricciardi, "Artificial intelligence for forecasting mosquito-borne viral diseases under climate change: A systematic review," Machine Learning and Knowledge Extraction, vol. 8, no. 1, pp. 1–22, 2026, doi: 10.3390/make8010015.
27. S. Haque, J. J. Rocklöv, and R. J. Lowe, "Climate-driven infectious disease early warning systems: A systematic review of methods, challenges, and future directions," Environment International, vol. 184, pp. 108503, 2024, doi: 10.1016/j.envint.2024.108503.
28. Socrates, S., Shanmugapriya, M., Murugeshwari, B., & Angalaeswari, S. (2024). Efficient Design for Implantable Device Constant Current Induction Doubly Fed Generating Incorporating Grid Connectivity. In *Intelligent Solutions for Sustainable Power Grids* (pp. 382-392). IGI Global Scientific Publishing.
29. Sugumar, R. (2026). Performance Optimization Frameworks for Financial Web Platforms with Real-Time Transaction Processing. *International Journal of Engineering & Extended Technologies Research (IJEETR), 8(2), 600-611.*
30. Anbazhagan, K. (2025). AI Driven Zero Trust Security Model for Enterprise Data Protection and Intelligent Infrastructure Management. *International Journal of Technology, Management and Humanities, 11(03), 101-107.*
31. Prabha, P. S., & Rengarajan, A. (2025). ENHANCING CLOUD RESOURCE ALLOCATION WITH VISION TRANSFORMER, DEEP REINFORCEMENT LEARNING, AND IMPROVED SHRIKE OPTIMIZATION ALGORITHM. *Corrosion Management ISSN: 1355-5243, 35(2), 233-245.*
32. Vimal, V. R., & Banerjee, J. S. (2025). Integrating PSO, GA, and ACO for Optimized ECG Feature Selection and Classification of Cardiac Disorders. *SGS-Engineering & Sciences, 1(5).*
33. Anbazhagan, K. (2025). Next-Generation Enterprise Cloud AI for Healthcare: Secure CNN Pipelines and Privacy Controls. *International Journal of Future Innovative Science and Technology (IJFIST), 8(6), 15980.*