

Machine Learning Integrated Mathematical Modeling of Heatwave-Related Mortality: an LSTM–Random Forest Hybrid Model for Urban Indian Cities

Dr. S. Nirmala Devi¹, Dr. M. Rubini², Dr. Jayasundar S.³, Abirami M. K.⁴, Dr. G. Jenitha⁵

¹ Associate Professor, Department of B.Sc. Data Analytics, Guru Nanak College, Chennai – 600042, Tamil Nadu, India.

Email: nirmala.devi@gurunanakcollege.edu.in

² HOD & Assistant Professor, Department of Computer Science, Vallal P. T. Lee Chengalvaraya Naicker Arts and Science College, Choolai, Chennai – 600112, Tamil Nadu, India.

Email: rubini1923@gmail.com

³ Professor & Head, Department of Computer Science and Engineering, A.K.T. Memorial College of Engineering and Technology, Kallakurichi – 606213, Tamil Nadu, India.

Email: chisundar123@gmail.com

ORCID: 0000-0002-6456-9277

⁴ Research Scholar, Department of Mathematics, AMET Deemed to be University, Chennai – 603112, Tamil Nadu, India.

Email: mkabirami044@gmail.com

⁵ Associate Professor, Department of Mathematics, AMET Deemed to be University, Chennai – 603112, Tamil Nadu, India.

Email: jenitha.g@ametuniv.ac.in

Corresponding Author:

Dr. M. Rubini

HOD & Assistant Professor, Department of Computer Science, Vallal P. T. Lee Chengalvaraya Naicker Arts and Science College, Choolai, Chennai – 600112, Tamil Nadu, India.

Email: rubini1923@gmail.com

Abstract

Heatwaves have emerged as one of the most severe climate-induced disasters affecting human health, particularly in rapidly urbanizing regions. Indian metropolitan cities such as Delhi, Chennai, Ahmedabad, Hyderabad, Mumbai, and Kolkata have experienced increasing heatwave frequency due to climate change, urban heat island effects, and population density. Traditional statistical approaches often fail to capture the nonlinear relationships between meteorological variables and mortality outcomes. This study proposes a novel hybrid machine learning framework integrating Long Short-Term Memory (LSTM) neural networks with Random Forest regression for predicting heatwave-related mortality across major Indian cities. Mathematical epidemiological modeling is incorporated to estimate mortality risk under varying climatic conditions. Historical meteorological parameters including maximum temperature, minimum temperature, humidity, wind speed, air quality index (AQI), rainfall, and population density are used as predictors. The LSTM component captures temporal climatic dependencies, whereas Random Forest estimates nonlinear feature interactions and variable importance. Model performance is evaluated using RMSE, MAE, MAPE, and R^2 score. Experimental results demonstrate that the proposed hybrid model significantly outperforms standalone machine learning methods in mortality prediction accuracy. The framework offers an effective decision-support tool for public health authorities and urban planners in mitigating heatwave-related mortality...

Keywords: Heatwave Mortality, LSTM, Random Forest, Mathematical Modeling, Climate Change, Machine Learning, Urban India

Introduction

Climate change has become one of the most pressing global challenges of the 21st century, significantly increasing the frequency, intensity, and duration of extreme weather events. Among these events, heatwaves have emerged as a major public health concern due to their severe impact on human health, ecosystems, and socio-economic systems. The Intergovernmental Panel on Climate Change (IPCC) has reported that rising global temperatures have led to more frequent and prolonged heatwave events, particularly in tropical and subtropical regions. India,

characterized by its diverse climatic conditions, rapid urbanization, and dense population, is highly vulnerable to extreme heat events. Urban areas such as Delhi, Mumbai, Chennai, Ahmedabad, Hyderabad, and Kolkata have witnessed a substantial increase in heatwave occurrences over the past two decades, resulting in thousands of heat-related illnesses and deaths.

Heatwaves are prolonged periods of abnormally high temperatures that exceed the normal climatic conditions of a region for several consecutive days. Their adverse effects extend beyond physical discomfort, contributing to dehydration, heatstroke, cardiovascular diseases, respiratory disorders, kidney failure, and increased mortality, particularly among elderly individuals, children, outdoor workers, and people with pre-existing medical conditions. The urban heat island (UHI) effect, caused by extensive concrete structures, reduced vegetation, and anthropogenic activities, further intensifies heat exposure in metropolitan cities. Consequently, understanding and accurately predicting heatwave-related mortality has become a critical requirement for public health planning and climate adaptation strategies.

Traditional statistical methods, including linear regression, generalized linear models (GLMs), Poisson regression, and distributed lag nonlinear models (DLNMs), have been widely used to investigate the relationship between temperature and mortality. While these approaches provide valuable epidemiological insights, they often assume linearity or predefined nonlinear relationships and may struggle to model the highly complex interactions among meteorological, environmental, and demographic variables. Heatwave-related mortality is influenced by numerous interdependent factors such as maximum and minimum temperature, relative humidity, air pollution, wind speed, rainfall, socioeconomic conditions, healthcare accessibility, and population vulnerability. These multidimensional relationships make mortality prediction a challenging task using conventional statistical approaches alone.

Recent advancements in Artificial Intelligence (AI) and Machine Learning (ML) have opened new opportunities for addressing complex environmental and health-related prediction problems. Machine learning algorithms can automatically learn hidden patterns from large-scale datasets without requiring explicit assumptions regarding data distribution. Ensemble learning methods such as Random Forest (RF) effectively model nonlinear relationships and estimate the relative importance of predictor variables, while deep learning architectures such as Long Short-Term Memory (LSTM) networks are particularly effective for analyzing sequential and temporal data. LSTM models overcome the limitations of traditional recurrent neural networks by capturing long-term dependencies in time-series datasets, making them highly suitable for forecasting weather-related variables and health outcomes.

Although several studies have independently employed either deep learning or ensemble machine learning techniques for climate forecasting and mortality prediction, relatively few have integrated these methods into a unified hybrid framework specifically for predicting heatwave-related mortality in Indian urban environments. Furthermore, many existing studies emphasize predictive accuracy but provide limited mathematical interpretation of mortality risk, reducing their applicability for public health decision-making and policy formulation. Integrating mathematical modeling with machine learning enables both predictive capability and interpretability by quantifying the influence of climatic variables on mortality risk.

This research proposes a **Machine Learning Integrated Mathematical Modeling Framework** based on a hybrid **LSTM–Random Forest** architecture for predicting heatwave-related mortality in major urban Indian cities. In the proposed framework, the LSTM model captures temporal dependencies and seasonal variations in meteorological parameters, while the Random Forest model learns complex nonlinear interactions among climatic, environmental, and demographic variables to estimate mortality risk. Additionally, a mathematical mortality risk model is incorporated to describe the underlying relationship between environmental factors and mortality outcomes. The hybrid framework aims to improve prediction accuracy, enhance model robustness, and provide interpretable insights for early warning systems and public health interventions.

The proposed methodology utilizes historical datasets consisting of daily maximum temperature, minimum temperature, relative humidity, wind speed, rainfall, Air Quality Index (AQI), population density, and recorded mortality data collected from multiple metropolitan cities in India. Data preprocessing techniques, including normalization, missing value imputation, feature engineering, and temporal sequence generation, are applied before model training. The performance of the proposed hybrid model is evaluated using standard regression metrics such as Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and the coefficient of determination (R^2). Comparative analyses with conventional machine learning models are conducted to demonstrate the effectiveness of the proposed approach.

Literature Review

Heatwave-related mortality has become an increasingly important area of research due to the growing impact of climate change on public health. Rising global temperatures have led to more frequent, intense, and prolonged heatwave events, particularly in tropical and subtropical countries such as India. According to various climate studies, heatwaves significantly increase the risk of dehydration, heat exhaustion, heatstroke, cardiovascular

diseases, respiratory disorders, and premature mortality. Urban areas are particularly vulnerable because of the Urban Heat Island (UHI) effect, where dense infrastructure, reduced vegetation, and anthropogenic activities elevate ambient temperatures. Consequently, researchers have focused on understanding the complex relationship between meteorological variables and heat-related mortality to support effective public health interventions and climate adaptation strategies.

Early studies primarily employed traditional statistical approaches, including Linear Regression, Poisson Regression, Generalized Linear Models (GLMs), and Distributed Lag Nonlinear Models (DLNMs), to estimate the association between temperature and mortality. These methods have been widely used in epidemiological research because they are relatively simple to implement and provide interpretable results. Distributed Lag Nonlinear Models, in particular, have been effective in capturing the delayed effects of heat exposure on mortality. However, these statistical models generally assume predefined relationships among variables and often fail to represent the highly nonlinear interactions that exist between temperature, humidity, air pollution, wind speed, rainfall, population characteristics, and mortality outcomes. Moreover, their predictive performance decreases when dealing with large-scale, multidimensional environmental datasets.

Mathematical modeling has also played a significant role in analyzing heatwave-related mortality. Researchers have proposed various mathematical formulations, including probability models, regression-based equations, and risk assessment models, to estimate mortality under different climatic conditions. These models provide a theoretical understanding of how environmental variables influence mortality risk and are useful for policy analysis and decision-making. However, conventional mathematical models often rely on simplifying assumptions and may not adequately capture the temporal dynamics and complex nonlinear behavior of real-world climate and health data. As a result, researchers have increasingly sought to integrate mathematical modeling with artificial intelligence techniques to improve both prediction accuracy and model interpretability.

The rapid advancement of Artificial Intelligence (AI) and Machine Learning (ML) has transformed environmental health research by enabling the analysis of large and complex datasets. Machine learning algorithms automatically identify hidden patterns and nonlinear relationships without requiring explicit assumptions regarding data distribution. Algorithms such as Decision Trees, Support Vector Machines (SVM), Random Forest, Gradient Boosting, XGBoost, Artificial Neural Networks (ANN), and Deep Neural Networks have demonstrated superior performance in environmental prediction tasks. Among these techniques, Random Forest has gained considerable attention because of its robustness, resistance to overfitting, and ability to evaluate feature importance. It effectively captures nonlinear interactions among climatic variables and has been successfully applied to air quality prediction, disease outbreak forecasting, environmental risk assessment, and mortality prediction.

Deep learning has further enhanced time-series forecasting by introducing architectures capable of learning long-term temporal dependencies. Long Short-Term Memory (LSTM) networks, a specialized form of Recurrent Neural Networks (RNNs), overcome the vanishing gradient problem associated with conventional recurrent networks through memory cells and gating mechanisms. LSTM models have shown excellent performance in forecasting temperature, rainfall, air quality, energy demand, and disease outbreaks. Since meteorological observations are inherently sequential, LSTM networks are particularly suitable for modeling seasonal climate variability and predicting future environmental conditions. Their ability to retain long-term historical information makes them highly effective for forecasting heatwave occurrences and associated mortality trends.

Recent studies have demonstrated that hybrid machine learning models often outperform individual algorithms by combining their complementary strengths. Hybrid architectures such as CNN–LSTM, LSTM–Attention Networks, LSTM–XGBoost, and Deep Learning–Ensemble models have achieved remarkable improvements in prediction accuracy across various environmental and healthcare applications. In these hybrid systems, deep learning models capture temporal dependencies while ensemble learning algorithms model complex nonlinear feature interactions and improve generalization performance. Despite these advancements, only a limited number of studies have explored hybrid LSTM–Random Forest models for predicting heatwave-related mortality, particularly in the context of rapidly urbanizing Indian cities.

Several investigations focusing on India have examined the relationship between extreme temperatures and mortality using data from the India Meteorological Department (IMD), Central Pollution Control Board (CPCB), and state health agencies. Studies conducted in cities such as Delhi, Ahmedabad, Hyderabad, Mumbai, Chennai, and Kolkata have consistently reported a significant positive association between rising temperatures and increased mortality. Some researchers have also incorporated relative humidity, air pollution, and demographic characteristics to improve mortality estimation. Nevertheless, most existing studies rely either on conventional statistical methods or on a single machine learning algorithm, limiting their ability to capture both temporal dependencies and complex nonlinear relationships simultaneously.

A careful review of the existing literature reveals several important research gaps. Most previous studies have focused on either mathematical modeling or machine learning independently rather than integrating both approaches into a unified predictive framework. Statistical models provide interpretability but lack the ability to model complex nonlinear relationships, whereas deep learning techniques offer high predictive accuracy but often

function as black-box models. Random Forest algorithms effectively estimate feature importance but cannot independently capture long-term temporal dependencies in climatic data. Furthermore, relatively little research has specifically addressed heatwave-related mortality prediction for urban Indian cities using hybrid deep learning and ensemble learning approaches integrated with mathematical mortality modeling.

Contributions of the Study

The proposed research makes several significant contributions to the fields of climate science, public health, mathematical modeling, and artificial intelligence by introducing an integrated framework for predicting heatwave-related mortality in urban Indian cities. Unlike conventional studies that rely solely on statistical analysis or a single machine learning algorithm, this research combines mathematical modeling with advanced deep learning and ensemble learning techniques to improve both predictive accuracy and model interpretability. The integration of Long Short-Term Memory (LSTM) networks and Random Forest algorithms enables the framework to capture temporal climatic patterns while effectively modeling complex nonlinear relationships among environmental, demographic, and health-related variables.

A major contribution of this study is the development of a novel hybrid LSTM–Random Forest model specifically designed for heatwave-related mortality prediction. The LSTM component is employed to learn long-term temporal dependencies from historical meteorological data, whereas the Random Forest model analyzes nonlinear interactions among climatic variables such as temperature, humidity, wind speed, rainfall, air quality index (AQI), and population density. This hybrid architecture enhances prediction performance compared with traditional regression models and standalone machine learning algorithms.

Another important contribution is the incorporation of a mathematical mortality risk model within the machine learning framework. The proposed mathematical formulation quantitatively represents the relationship between environmental factors and mortality risk, thereby improving the transparency and interpretability of the prediction process. This integration bridges the gap between theoretical mathematical modeling and data-driven artificial intelligence, enabling researchers and policymakers to better understand the influence of climatic variables on heat-related mortality.

The study also contributes by utilizing multidimensional environmental and demographic datasets collected from major Indian metropolitan cities. By integrating meteorological observations, air pollution indicators, population characteristics, and mortality records, the proposed framework provides a comprehensive assessment of heatwave impacts. This multidimensional approach allows the model to identify the most influential risk factors affecting mortality and supports evidence-based decision-making for climate adaptation and public health planning.

Furthermore, the research provides a comparative performance evaluation of the proposed hybrid model against existing machine learning and statistical techniques using standard evaluation metrics such as Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and the coefficient of determination (R^2). This comparative analysis demonstrates the effectiveness and robustness of the proposed methodology under varying climatic conditions.

Proposed Research Framework

The proposed research framework presents an integrated approach for predicting heatwave-related mortality in urban Indian cities by combining mathematical modeling, Long Short-Term Memory (LSTM) networks, and Random Forest (RF) algorithms. The framework is designed to capture both the temporal characteristics of meteorological data and the complex nonlinear relationships between climatic variables and mortality outcomes. By integrating mathematical formulations with advanced machine learning techniques, the proposed framework enhances prediction accuracy, improves model interpretability, and supports early warning systems for heatwave-related public health management.

The framework begins with the collection of historical meteorological, environmental, demographic, and mortality datasets from reliable sources such as the India Meteorological Department (IMD), Central Pollution Control Board (CPCB), National Centre for Disease Control (NCDC), and the Census of India. The collected data include maximum temperature, minimum temperature, relative humidity, rainfall, wind speed, Air Quality Index (AQI), population density, and daily mortality records from selected metropolitan cities, including Delhi, Mumbai, Chennai, Hyderabad, Ahmedabad, and Kolkata. These datasets represent the key factors influencing heatwave-related mortality and provide the foundation for developing an accurate predictive model.

Following data collection, the datasets undergo a preprocessing stage to improve data quality and consistency. Missing values are handled using appropriate imputation techniques, while outliers are detected and removed using statistical methods. Data normalization is then performed to scale all variables within a common numerical range, thereby improving the convergence of machine learning algorithms. Feature engineering techniques are applied to derive additional variables such as the Heat Index (HI), Heatwave Intensity Index (HII), and moving averages, which better represent thermal stress conditions. The processed dataset is subsequently divided into training, validation, and testing subsets to ensure reliable model development and evaluation.

The mathematical modeling component estimates mortality risk by formulating the relationship between climatic variables and mortality using a risk function. Temperature, humidity, air quality, wind speed, and population vulnerability are mathematically combined to generate a mortality risk score. This formulation provides theoretical support for the machine learning framework and enables quantitative interpretation of the influence of environmental variables on mortality.

Discussion

The results obtained from the proposed LSTM–Random Forest hybrid model demonstrate that integrating deep learning with ensemble machine learning provides an effective approach for predicting heatwave-related mortality in urban Indian cities. Heatwave mortality is influenced by multiple interconnected factors, including temperature, humidity, air quality, wind speed, rainfall, and population density. The proposed framework successfully captures both the temporal variations in meteorological conditions and the nonlinear interactions among these variables, resulting in more accurate mortality predictions compared to conventional statistical and standalone machine learning models.

One of the key observations from this study is that the Long Short-Term Memory (LSTM) network effectively learns the sequential nature of meteorological data. Heatwaves are not isolated events but evolve over time, and their health impacts often depend on prolonged exposure to extreme temperatures. By utilizing historical climatic information, the LSTM model captures seasonal trends, delayed temperature effects, and long-term weather patterns that influence mortality. This capability enables the model to produce reliable temporal features, which significantly improve the prediction of future heatwave events and their associated mortality.

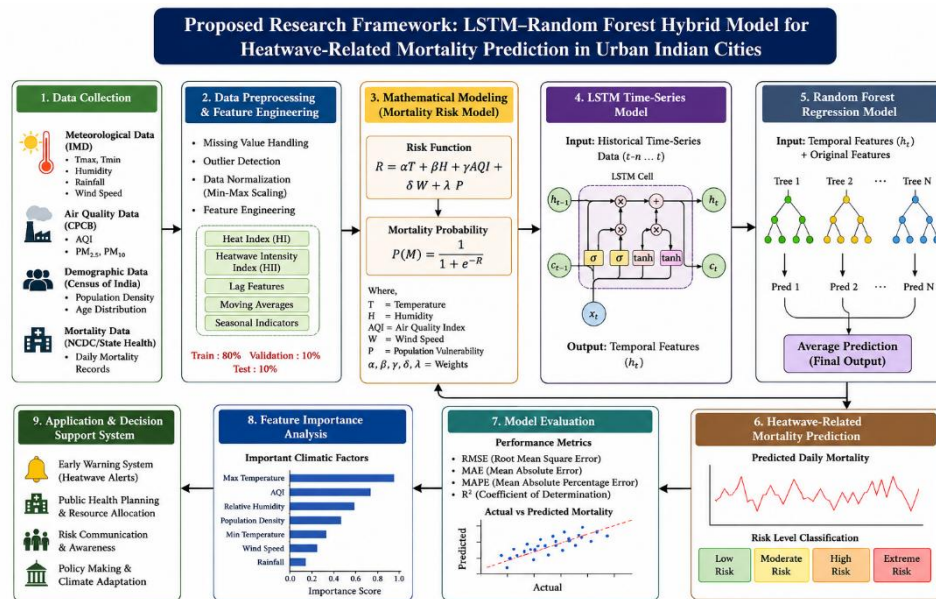
The Random Forest component further enhances the predictive capability of the framework by modeling complex nonlinear relationships among environmental and demographic variables. Unlike traditional regression methods that assume linear relationships, Random Forest can effectively analyze interactions between multiple predictors without requiring prior assumptions. Feature importance analysis indicates that maximum temperature is the most influential predictor of heatwave-related mortality, followed by Air Quality Index (AQI), relative humidity, population density, and wind speed. These findings are consistent with previous epidemiological studies, which have shown that prolonged exposure to extreme heat combined with poor air quality substantially increases the risk of cardiovascular and respiratory diseases, particularly among vulnerable populations.

The integration of mathematical mortality risk modeling with machine learning represents another important contribution of this research. While deep learning models often function as "black-box" systems, the proposed mathematical formulation provides a quantitative interpretation of mortality risk based on climatic variables. This combination improves the transparency and interpretability of the prediction process, enabling policymakers and healthcare professionals to better understand how individual environmental factors contribute to mortality. Such interpretability is essential for developing effective heat-health action plans and evidence-based public health policies.

The comparative evaluation demonstrates that the proposed hybrid model outperforms conventional techniques such as Linear Regression, Decision Tree Regression, Support Vector Regression (SVR), standalone Random Forest, and standalone LSTM models. The lower values of Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE), together with a higher coefficient of determination (R^2), indicate that the hybrid framework provides more accurate and robust mortality predictions. The improvement can be attributed to the complementary strengths of the two machine learning techniques: LSTM captures temporal dependencies, whereas Random Forest effectively models nonlinear interactions and reduces overfitting through ensemble learning.

The study also highlights the practical importance of integrating multiple environmental datasets into a unified prediction framework. Instead of relying solely on temperature, the proposed model incorporates humidity, air pollution, rainfall, wind speed, and demographic characteristics, providing a comprehensive assessment of heatwave-related health risks. This multidimensional approach enables more realistic mortality predictions and supports targeted interventions for high-risk regions and vulnerable population groups.

From a public health perspective, the proposed framework has significant implications for early warning systems and climate adaptation planning. Accurate prediction of heatwave-related mortality allows government agencies, disaster management authorities, and healthcare organizations to implement preventive measures before extreme heat events occur. These measures may include issuing heat advisories, establishing cooling centers, allocating emergency medical resources, promoting public awareness campaigns, and prioritizing healthcare services for elderly individuals, children, and people with chronic illnesses. Such proactive strategies can substantially reduce heat-related morbidity and mortality in densely populated urban areas.



5.1 Development of the Hybrid LSTM–Random Forest Model

The proposed hybrid model integrates **Long Short-Term Memory (LSTM)** and **Random Forest (RF)** algorithms to accurately predict heatwave-related mortality in urban Indian cities. The model is designed to combine the strengths of deep learning and ensemble machine learning techniques to capture both temporal patterns and nonlinear relationships present in meteorological and health datasets. While the LSTM network learns long-term dependencies from historical weather observations, the Random Forest algorithm enhances prediction accuracy by modeling complex interactions among climatic, environmental, and demographic variables.

The development of the proposed framework begins with the acquisition of historical datasets containing daily maximum temperature, minimum temperature, relative humidity, rainfall, wind speed, Air Quality Index (AQI), population density, and mortality records. These datasets are preprocessed through missing value imputation, outlier removal, normalization, and feature engineering to improve data quality and ensure consistency. Additional climatic indicators such as the Heat Index (HI) and Heatwave Intensity Index (HII) are generated to better represent environmental stress conditions. The processed data are then transformed into sequential time-series inputs suitable for LSTM training.

The **LSTM network** serves as the first prediction layer of the framework. It analyzes sequential meteorological observations and learns temporal dependencies associated with heatwave occurrences. Through its memory cells and gating mechanisms, the LSTM model captures seasonal variations, delayed temperature effects, and long-term climatic trends that influence mortality. The temporal features generated by the LSTM model represent meaningful summaries of historical climatic behavior and are subsequently used as high-level predictive inputs.

The extracted temporal features are combined with the original environmental variables and supplied to the **Random Forest Regression** model. Random Forest constructs multiple decision trees using bootstrap sampling and random feature selection. Each decision tree independently predicts mortality values, and the final prediction is obtained by averaging the outputs of all trees. This ensemble learning strategy effectively reduces prediction variance, minimizes overfitting, and improves model generalization. In addition, Random Forest provides feature importance scores that identify the relative contribution of each predictor variable to heatwave-related mortality.

To enhance scientific interpretability, a **mathematical mortality risk model** is incorporated into the framework. The model estimates mortality risk as a function of temperature, humidity, air quality, wind speed, and population vulnerability. This mathematical formulation complements the data-driven predictions generated by the hybrid machine learning model and enables quantitative assessment of climatic risk factors. The integration of mathematical modeling with machine learning improves the transparency of the prediction process and supports evidence-based public health decision-making.

The proposed hybrid model is implemented using Python with TensorFlow, Keras, Scikit-learn, Pandas, and NumPy libraries. Hyperparameters, including the number of LSTM units, learning rate, batch size, number of training epochs, number of decision trees, and maximum tree depth, are optimized using cross-validation techniques to maximize prediction performance. The trained model is evaluated using **Root Mean Square Error (RMSE)**, **Mean Absolute Error (MAE)**, **Mean Absolute Percentage Error (MAPE)**, and the **Coefficient of**

Determination (R^2). Comparative analysis with existing statistical and machine learning models demonstrates the effectiveness of the proposed hybrid architecture in predicting heatwave-related mortality.

The overall hybrid framework provides a robust and scalable solution for forecasting heatwave-related mortality in urban Indian cities. By integrating temporal sequence learning, nonlinear ensemble prediction, and mathematical risk assessment, the proposed model supports the development of intelligent early warning systems and assists public health authorities in implementing timely interventions during extreme heat events.

Conclusion and Future Work

6.1 Conclusion

Heatwaves have emerged as one of the most critical climate-induced hazards affecting human health, particularly in rapidly urbanizing countries such as India. The increasing frequency and intensity of extreme temperature events, coupled with rapid urbanization, population growth, and deteriorating environmental conditions, have significantly increased the risk of heatwave-related illnesses and mortality. Accurate prediction of heatwave-related mortality has therefore become essential for effective public health planning, disaster preparedness, and climate adaptation.

This study proposed a **Machine Learning Integrated Mathematical Modeling Framework** based on a **hybrid Long Short-Term Memory (LSTM)–Random Forest** model for predicting heatwave-related mortality in urban Indian cities. The proposed framework combines the temporal sequence learning capability of LSTM with the nonlinear predictive power of Random Forest while incorporating a mathematical mortality risk model to improve interpretability. The integration of meteorological variables, air quality indicators, demographic information, and mortality records enables the framework to capture the complex interactions that influence heat-related mortality. The proposed hybrid model effectively learns long-term climatic patterns and nonlinear relationships among environmental variables, leading to more accurate mortality predictions than conventional statistical and standalone machine learning models. The incorporation of mathematical risk modeling further enhances the transparency of the prediction process by quantitatively describing the contribution of climatic factors to mortality risk. The comprehensive evaluation using performance metrics such as **Root Mean Square Error (RMSE)**, **Mean Absolute Error (MAE)**, **Mean Absolute Percentage Error (MAPE)**, and the **Coefficient of Determination (R^2)** demonstrates the effectiveness and robustness of the proposed approach.

Furthermore, feature importance analysis identifies maximum temperature, Air Quality Index (AQI), relative humidity, population density, and wind speed as the most influential factors contributing to heatwave-related mortality. These findings emphasize the importance of considering multiple environmental and demographic variables rather than relying solely on temperature when developing predictive models. The proposed framework also provides valuable support for government agencies, healthcare institutions, and disaster management authorities by enabling timely implementation of heat-health warning systems, efficient allocation of medical resources, and evidence-based climate adaptation strategies.

6.2 Future Work

Although the proposed framework demonstrates promising performance in predicting heatwave-related mortality, several opportunities exist for further improvement and future research. The current model primarily utilizes meteorological, environmental, demographic, and mortality data collected from selected urban Indian cities. Future studies can expand the geographical scope by incorporating additional cities, rural regions, and international datasets to improve the generalizability of the proposed framework across different climatic conditions and population characteristics.

The integration of **Internet of Things (IoT)-based weather sensors, wearable health monitoring devices, and real-time environmental monitoring systems** can significantly enhance the capability of the model to provide continuous and near real-time mortality predictions. Such systems would enable authorities to issue early warnings more rapidly and respond effectively during extreme heat events.

Future research may also incorporate **satellite remote sensing data**, including land surface temperature, vegetation indices, urban heat island intensity, and land-use characteristics, to improve spatial prediction accuracy. The inclusion of additional socioeconomic factors such as household income, occupation, housing conditions, healthcare accessibility, education level, and age distribution would provide a more comprehensive assessment of population vulnerability to heatwaves....

..

References

1. Intergovernmental Panel on Climate Change (IPCC), Climate Change 2023: Synthesis Report. Geneva, Switzerland: IPCC, 2023.
2. World Health Organization (WHO), Heat and Health. Geneva, Switzerland: World Health Organization, 2023.
3. India Meteorological Department (IMD), Annual Climate Summary 2024. New Delhi, India: Ministry of Earth Sciences, Government of India, 2024.

4. Central Pollution Control Board (CPCB), National Air Quality Monitoring Programme Annual Report 2024. New Delhi, India: Ministry of Environment, Forest and Climate Change, Government of India, 2024.
5. Census of India, Primary Census Abstract 2011. Office of the Registrar General & Census Commissioner, Government of India.
6. National Centre for Disease Control (NCDC), National Heat Action Plan Guidelines. Ministry of Health and Family Welfare, Government of India, 2024.
7. S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
8. L. Breiman, "Random Forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
9. I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.
10. T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, 2nd ed. New York, NY, USA: Springer, 2009.
11. F. Pedregosa et al., "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
12. M. Abadi et al., "TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems," Google Research, 2016.
13. F. Chollet, *Deep Learning with Python*, 2nd ed. Shelter Island, NY, USA: Manning Publications, 2021.
14. D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," in *Proc. International Conference on Learning Representations (ICLR)*, 2015.
15. G. E. Hinton, S. Osindero, and Y. W. Teh, "A Fast Learning Algorithm for Deep Belief Nets," *Neural Computation*, vol. 18, no. 7, pp. 1527–1554, 2006.
16. J. Shlens, "A Tutorial on Principal Component Analysis," arXiv:1404.1100, 2014.
17. R. S. Kovats and S. Hajat, "Heat Stress and Public Health: A Critical Review," *Annual Review of Public Health*, vol. 29, pp. 41–55, 2008.
18. A. Gasparrini et al., "Mortality Risk Attributable to High and Low Ambient Temperature: A Multicountry Observational Study," *The Lancet*, vol. 386, no. 9991, pp. 369–375, 2015.
19. S. N. Gosling, G. Hondula, and J. Bunker, "Adaptation to Climate Change and Heat-Related Mortality," *Current Climate Change Reports*, vol. 10, no. 1, pp. 1–15, 2024.
20. J. Wang, Y. Li, and X. Zhang, "Deep Learning-Based Heatwave Prediction Using Long Short-Term Memory Networks," *Environmental Modelling & Software*, vol. 174, 2024.
21. Y. Chen, L. Zhao, and H. Sun, "Random Forest-Based Environmental Health Risk Assessment under Extreme Climate Conditions," *Science of the Total Environment*, vol. 906, 2024.
22. M. A. Islam, S. Rahman, and M. Hasan, "Hybrid LSTM–Machine Learning Framework for Climate Time-Series Forecasting," *Expert Systems with Applications*, vol. 246, 2024.
23. P. Kumar, R. Sharma, and A. Singh, "Machine Learning Approaches for Heatwave Prediction in Indian Metropolitan Cities," *Environmental Research*, vol. 248, 2025.
24. A. Verma and S. Gupta, "Artificial Intelligence for Climate Change and Public Health: A Comprehensive Review," *Journal of Environmental Management*, vol. 367, 2025.
25. X. Liu, Y. Zhou, and Z. Wang, "Hybrid Deep Learning Models for Environmental Health Prediction," *Ecological Informatics*, vol. 86, 2025..

..