International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

A Review Paper on Machine Learning Models for Weather Forecasting: Techniques, Trends, and Challenges

Monika Jakhar[1] Dr. Vinod Kumar Srivastava[2]

[1]Research Scholar Department of Computer Science and Applications Baba Mastnath University Rohtak, Haryana, India
[2]Professor Department of Computer Science and Applications Baba Mastnath University Rohtak, Haryana, India

Abstract-
Weather forecasting is a crucial point of discussion in agriculture, aviation, disaster management, and energy decision-making. Conventional numerical weather prediction (NWP) models are efficient but need a lot of computer resources and can be inaccurate in the beginning. Machine learning (ML) has become a powerful alternative and addition to traditional forecasting approaches, giving data-driven methodologies that can simulate complex meteorological patterns more efficiently and adaptably. This review covers weather forecasting machine learning methods, including supervised learning models like Support Vector Machines (SVM), Random Forests, and Gradient Boosting and deep learning frameworks like CNNs, RNNs, and LSTM networks. This review focuses on using these models to anticipate temperature, precipitation, wind speed, and humidity across various temporal and spatial scales. Modern model designs, hybrid models, ensemble learning, and satellite and sensor data to improve forecast accuracy are examined. It addresses crucial issues such data quality and accessibility, model interpretability, overfitting, and real-time forecasting. Explainable AI and uncertainty quantification are crucial to trusting machine learning-based weather systems, according to the review. This paper reviews the current and future state of machine learning in meteorological forecasting.

Keywords: Weather Forecasting, Machine Learning, Climate Change, Trends and challenges etc.

1. INTRODUCTION

Weather forecasting is an essential component in contemporary life, impacting decisions in agriculture, transportation, disaster planning, energy management, and public safety. Precise weather forecasts assist in alleviating the effects of natural disasters such as storms, floods, and droughts, while also enhancing daily operations in industries such as aviation and logistics. Historically, weather forecasting has depended on numerical weather prediction (NWP) models, which employ mathematical equations to replicate atmospheric dynamics based on beginning conditions and scientific principles. Although these models have progressed much over the years, they are computationally demanding and frequently encounter difficulties with uncertainty, particularly in long-term or localized predictions.

In recent years, the proliferation of data from satellites, meteorological stations, and IoT-based sensors has created new opportunities for data-driven forecasting. Machine learning (ML), capable of discerning intricate patterns from extensive datasets, has surfaced as a viable alternative or enhancement to conventional approaches. In contrast to NWP models, ML methodologies do not necessitate explicit physical modeling; rather, they deduce correlations directly from historical data. This renders them especially proficient in identifying nonlinear dependencies and anomalies in atmospheric behavior, which are frequently challenging to model with traditional methods (Breiman, 2001 [1]).

Diverse machine learning techniques have been used for weather forecasting, encompassing traditional models such as Decision Trees and Support Vector Machines, as well as sophisticated deep learning frameworks like Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks. These models have demonstrated encouraging outcomes in forecasting temperature, precipitation, wind velocity, and other meteorological parameters, frequently with diminished computing demands and enhanced adaptability. Furthermore, hybrid models that integrate physical simulations with machine learning-based adjustments are increasingly popular due to their capacity to utilize the advantages of both approaches (Wilks, 2007 [2]).

Notwithstanding these achievements, numerous difficulties persist. Concerns persist regarding the quality and granularity of input data, the interpretability of intricate models, and the necessity for real-time processing. Moreover, guaranteeing the robustness and generalizability of machine learning models across diverse geographic regions and climatic circumstances is essential for their extensive adoption. This paper intends to examine the present state of machine learning in weather forecasting, emphasizing essential methodologies, recent developments, and the obstacles that must be overcome to fully harness the capabilities of AI-driven meteorology (Akaike, 1974 [3]).

2. RELATED WORK

Ejike et al. (2025) [4] conducted an analysis to evaluate the effectiveness of machine learning algorithms in predicting rainfall across tropical and temperate climatic zones. Historical meteorological data were collected and preprocessed to train models including Support Vector Machines (SVM), Random Forest, and Artificial Neural Networks (ANN). Performance was assessed using metrics such as Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and R^2 score. Results indicated that tropical regions exhibited higher prediction variability due to complex atmospheric dynamics, while temperate zones yielded more stable outputs. Among the models, Random Forest consistently demonstrated superior accuracy across both climates. The study highlighted the importance of climate-specific model tuning and feature selection in enhancing predictive reliability.

Jankauskas et al. (2024) [5] sought to conduct a thorough analysis of five weather forecasting models sourced from the Open-Meteo historical data collection, focusing on their efficacy in predicting wind power output. With the growing emphasis on renewable energy, particularly wind power, precise weather forecasting is essential for optimizing energy production and maintaining the integrity of the power grid. This study's analysis includes various models: ICO sahedral Nonhydrostatic (ICON), the Global Environmental Multiscale Model (GEM Global), Meteo France, the Global Forecast System (GFS Global), and the Best Match approach. This strategy was validated by juxtaposing the model predictions with empirical data on wind power generation. The ICON model achieved a root mean squared error of 1.7565, surpassing Best Match, which recorded a root mean squared error of 1.7604, representing a minor yet significant enhancement. GEM Global and GSF Global exhibited significant alterations, with root mean squared errors (RMSEs) of 2.0086 and 2.0242, respectively, reflecting a decline in predictive accuracy of around 24% to 31% relative to ICON. The findings indicated substantial discrepancies in the accuracy of the different models employed, with some models demonstrating markedly superior predicted precision.

Guo et al. (2024) [6] created a rainfall forecasting platform utilizing the GNSS-assisted weather research and forecasting (WRF) model and quantitatively evaluated the impact of GNSS precipitable water vapor (PWV) on the accuracy of WRF model predictions for light rain (LR), moderate rain (MR), heavy rain (HR), and torrential rain (TR). In 2021, three strategies were developed and evaluated utilizing data from seven ground meteorological stations in Xi'an City, China. The findings indicated that using GNSS PWV markedly enhanced the forecasting precision of the WRF model across various rainfall intensities, with root mean square error (RMSE) improvement rates of 8%, 15%, 19%, and 25% for LR, MR, HR, and TR, respectively. The results validated the platform's high precision, visualization capabilities, and robustness in rainfall forecasting.

Teixeira et al. (2024) [7] utilized regression models to estimate absent historical data across three distinct time horizons, integrating long short-term memory (LSTM) to predict short- to medium-term weather conditions at Quinta de Santa Bárbara in the Douro region. A genetic algorithm (GA) was employed to optimize the hyperparameters of the LSTM. The findings indicated that the proposed improved LSTM significantly diminished the assessment measures across various time horizons. The results emphasized the significance of precise weather forecasting in informing critical choices across all industries.

Zhang et al. (2024) [8] enhanced the precision of atmospheric temperature and humidity profile retrieval and examined the impacts of cloud data (cloud-base height and cloud thickness) on these retrievals. The observational data from the ground-based multichannel microwave radiometer (GMR) and the millimeter-wave cloud radar (MWCR) were integrated into the retrieval procedure of atmospheric temperature and relative humidity profiles. The retrieval was executed with the backpropagation neural network (BPNN). The retrieval outcomes were evaluated using the mean absolute error (MAE) and root mean square error (RMSE). The statistical results indicated that the temperature profiles were less influenced by the cloud data than the relative humidity profiles. In comparison to the retrieval profiles devoid of cloud information, the MAE and RMSE values for most height levels exhibited varying degrees of reduction following the incorporation of cloud data, with the relative humidity (RH) errors in certain altitude layers diminishing by almost 50%. The greatest decrease in RMSE and MAE values for temperature profile retrieval using cloud data was approximately 1.0 ◦C at 7.75 km, whereas the maximum reduction in RMSE and MAE values for relative humidity profiles was roughly 10%, achieved at 2 km.

Sim et al. (2024) [9] introduced a sophisticated ocean fog prediction model for the Yellow Sea region, utilizing satellite-based detection and high-performance data-driven techniques. The study utilized Himawari-8 satellite data to acquire extensive spatiotemporal ocean fog references and implemented Auto ML to integrate numerical weather prediction (NWP) outputs with sea surface temperature (SST)-related variables. The model exhibited enhanced performance relative to conventional NWP-based approaches, with impressive metrics in both quantitative—probability of detection at 81.6%, false alarm ratio at 24.4%, F1 score at 75%, and proportion correct at 79.8%and qualitative assessments over lead periods ranging from 1 to 6 hours. Significant contributing factors comprised relative humidity, cumulative shortwave radiation, and atmospheric pressure, underscoring the necessity of using varied data sources. The study highlighted the potential of satellite-derived data to enhance ocean fog prediction, while also addressing the issues of overfitting and the necessity for more complete reference data.

Shiferaw et al. (2024) [10] sought to determine an optimal configuration of the Weather, Research, and Forecasting (WRF) model for Ethiopia. Thirty-five WRF simulations employing various combinations of parameterization approaches for cumulus (CU), planetary boundary layer (PBL), cloud microphysics (MP), longwave (LW), and shortwave (SW) radiation were evaluated throughout the summer season (June to August, JJA) of 2002. The WRF simulations employed a two-domain design featuring a 12 km nested domain encompassing Ethiopia. The starting and boundary forcing data for WRF were sourced from the Climate Forecast System Reanalysis (CFSR). The simulations were assessed against station and gridded observations to determine their efficacy in replicating various characteristics of JJA rainfall. An objective ranking methodology employing an aggregate score of multiple statistics was utilized to identify the optimal model configuration. All models accurately represented the regional distribution of JJA rainfall, with the pattern correlation coefficient (PCC) varying between 0.89 and 0.94.

Ahmadgourabi et al. (2024) [11] sought to develop a dependable water-demand forecasting system utilizing Long Short-Term Memory networks. The model incorporated hourly water needs from ten District Metered Areas within a Water Distribution Network in northeastern Italy, alongside weather data, addressing missing values through LSTM-based data imputation. It took into account temporal factors such as time, weekdays, holidays, and weekend routines, utilizing sine and cosine transforms to represent daily cycles. The model's robustness was ensured by conducting testing in the final week of the dataset, especially week 81, while making iterative tweaks to the LSTM's hyperparameters to enhance prediction accuracy. The tuning efforts concentrated on the learning rate, layer count, and batch size, optimized to enhance the system's performance.

Ayyash et al. (2024) [12] proposed the application of a neural network model for short-term water demand forecasting. Their study demonstrated how artificial intelligence techniques can effectively capture non-linear patterns in water consumption, thereby improving prediction accuracy compared to traditional statistical methods. The research emphasized the importance of machine learning in resource management and highlighted the adaptability of neural networks in handling dynamic demand fluctuations.

Xian et al. (2024) [13] examined the impact of assimilating geostationary interferometric infrared sounder observations from long- and middle-wave bands into a locally cloud-resolving global model for weather forecasting. Their findings showed that integrating these advanced satellite observations significantly enhanced the accuracy of weather predictions, particularly in cloud-dense regions. The study underscored the role of remote sensing technologies in strengthening global forecasting systems.

Díaz-Ramírez and Badilla-Torrico (2024) [14] presented a comparative analysis of machine learning techniques for weather forecasting through a case study. They evaluated multiple algorithms to determine their relative strengths and weaknesses in predicting weather conditions. The study concluded that certain machine learning models outperform others depending on the dataset and forecasting horizon, thereby providing valuable insights into model selection for meteorological applications.

Bajad et al. (2024) [15] proposed the use of machine learning approaches for weather prediction, focusing on practical implementation and performance evaluation. Their research highlighted how algorithms such as decision trees and neural networks can process large volumes of meteorological data to generate reliable forecasts. The study contributed to the growing body of evidence supporting machine learning as a robust alternative to conventional forecasting techniques.

Gupta and Sharma (2024) [16] examined advancements in weather forecasting achieved through machine learning algorithms. Their work discussed recent innovations, including deep learning architectures, and how these methods enhance predictive accuracy and efficiency. The study emphasized the transformative potential of machine learning in modern meteorology, particularly in addressing challenges posed by complex and rapidly changing weather systems.

A. Research Gap

Most ML models are trained on region-specific data and struggle to generalize across diverse climatic zones (e.g., tropical vs. temperate). ML models often ignore physical laws of atmospheric science, leading to unrealistic predictions. Weather datasets often contain missing values, noise, or irregular sampling intervals. Robust ML techniques for gap-filling and uncertainty quantification are still evolving. ML models excel in short-term forecasting but struggle with long-term predictions due to chaotic atmospheric dynamics. Temporal modeling techniques (e.g., LSTM, transformers) need refinement for extended forecasts.

Table 1: Summary of Literature Survey

Author(s)	Year	Objectives	Findings	Research Gap
Ogochukwu Ejike [4]	2025	To compare the performance of machine learning algorithms in predicting rainfall across tropical and temperate climatic zones using historical meteorological data.	Random Forest outperformed SVM and ANN in both climates. Tropical regions showed higher prediction variability due to complex atmospheric dynamics, while temperate zones yielded more stable forecasts.	Limited integration of physical atmospheric principles and absence of standardized benchmark datasets hinder model robustness and cross-study comparison.
Jankauskas et al. [5]	2024	To assess the impact of weather forecast models on AI-based power generation predictions using BiLSTM.	BiLSTM improved prediction accuracy when integrated with weather forecast data.	Limited to power generation; lacks generalization to broader meteorological applications.
Guo et al. [6]	2024	To evaluate rainfall forecasting using GNSS-assisted NWP models.	GNSS data enhanced rainfall prediction accuracy in NWP models.	Needs validation across diverse climatic regions and longer time scales.
Teixeira et al. [7]	2024	To integrate LSTM with Genetic Algorithms for enhanced weather forecasting.	LSTM-GA hybrid improved forecast precision over standalone models.	Computational complexity and scalability to real-time systems remain unaddressed.
Zhang et al. [8]	2024	To improve atmospheric temperature and humidity profiling using microwave radiometers	Enhanced retrieval accuracy of atmospheric profiles.	Focused on vertical profiling; lacks integration with surface-level

Author(s)	Year	Objectives	Findings	Research Gap
		and cloud radar.		forecasting.
Sim et al. [9]	2024	To improve short-term ocean fog prediction using AutoML and satellite data.	AutoML models outperformed traditional methods in fog detection.	Limited to oceanic regions; generalizability to inland fog conditions is unclear.
Shiferaw et al. [10]	2024	To analyze WRF model sensitivity to parameterization over Ethiopia.	Identified optimal parameterization schemes for regional accuracy.	Results are region-specific; broader applicability needs exploration.
Boloukasli Ahmadgourabi et al. [11]	2024	To enhance water demand forecasting using LSTM networks.	LSTM provided high accuracy in short-term water demand prediction.	Focused on hydrological forecasting; indirect implications for weather modeling.
Ayyash et al. [12]	2024	To apply neural networks for short-term water demand forecasting.	Neural networks showed reliable performance in demand prediction.	Lacks integration with meteorological variables for weather-linked forecasting.
Xian et al. [13]	2024	To assess the impact of assimilating infrared sounder data on weather forecasts.	Improved forecast accuracy with geostationary infrared data assimilation.	High-resolution models require further optimization for global scalability.
Díaz-Ramírez & Badilla-Torrico [14]	2024	To compare ML techniques for weather forecasting in a case study.	Identified LSTM and RF as top performers in accuracy.	Case-specific results; lacks cross-validation across datasets and regions.
Bajad et al. [15]	2024	To develop a machine learning model for general weather prediction.	Demonstrated feasibility of ML in short-term weather forecasting.	Requires deeper evaluation of long-term forecasting capabilities.
Gupta & Sharma [16]	2024	To review advancements in ML algorithms for weather forecasting.	Highlighted recent progress in deep learning and hybrid models.	Lacks empirical benchmarking and real-world deployment analysis.

3. OBJECTIVES OF WORK

A. Problem Statement

Despite growing interest in applying machine learning techniques to rainfall prediction, significant challenges persist in achieving reliable and climate-sensitive forecasting across diverse geographical zones. Existing models often lack adaptability to the complex atmospheric dynamics of tropical regions and exhibit limited generalization when transferred to temperate climates. Furthermore, inconsistencies in data quality, absence of standardized benchmarks, and inadequate integration of domain-specific meteorological knowledge hinder the robustness and scalability of predictive systems. This study addresses the need for a comparative evaluation of machine learning algorithms tailored to distinct climatic conditions, aiming to identify performance disparities and inform the development of more accurate, climate-aware forecasting frameworks.

B. Research Objectives

To systematically review and categorize machine learning models applied to weather forecasting, including traditional algorithms (e.g., SVM, Random Forest) and deep learning architectures (e.g., LSTM, CNN, hybrid models), based on their structure, input data types, and forecasting capabilities.

4. MACHINE LEARNING MODELS FOR WEATHER FORECASTING

Machine learning (ML) models have become increasingly popular in weather forecasting due to their ability to learn complex, nonlinear relationships from large datasets without relying on explicit physical equations. These models can be broadly categorized into traditional ML algorithms and deep learning architectures, each offering unique advantages depending on the forecasting task and data characteristics.

A. Traditional Machine and Deep Learning Models

These models are often used for short-term forecasting and structured data analysis (Ejike, 2025):

- Linear Regression Used for predicting continuous weather variables like temperature or humidity. It assumes a linear relationship between input features and the target variable, making it simple but limited in handling complex patterns.
- Decision Trees and Random Forests These models split data into branches based on feature thresholds. Random Forests, an ensemble of decision trees, are particularly effective in handling noisy weather data and capturing nonlinear relationships.
- Support Vector Machines (SVM) SVMs are used for both classification (e.g., predicting rain/no rain) and regression tasks. They are effective in high-dimensional spaces and can model complex boundaries between weather conditions.
- K-Nearest Neighbors (KNN) A non-parametric method that predicts weather outcomes based on the similarity to historical data points. It is simple but computationally expensive for large datasets.

B. Deep Learning Models

Deep learning models are particularly well-suited for time-series forecasting and spatial-temporal data, such as satellite imagery and sensor streams (Jankauskas, 2024).

- Artificial Neural Networks (ANNs) Basic feedforward networks that can model nonlinear relationships. They are often used for temperature and wind speed prediction but lack memory of past inputs.
- Recurrent Neural Networks (RNNs) Designed for sequential data, RNNs maintain a memory of previous inputs, making them suitable for time-series weather forecasting. However, they suffer from vanishing gradient issues over long sequences.
- Long Short-Term Memory (LSTM) Networks A specialized type of RNN that overcomes the limitations of standard RNNs by using memory cells and gates. LSTMs are widely used for forecasting temperature, rainfall, and wind patterns due to their ability to capture long-term dependencies in weather data.
- Convolutional Neural Networks (CNNs) Primarily used for spatial data like satellite images or radar

maps. CNNs can detect patterns such as cloud formations or storm systems, making them valuable for visual-based weather prediction.

- Hybrid Models (e.g., CNN-LSTM) Combine the spatial feature extraction capabilities of CNNs with the temporal modeling power of LSTMs. These models are effective for forecasting based on sequences of weather maps or satellite imagery.

C. Emerging and Hybrid Approaches

- Ensemble Learning Combines predictions from multiple models to improve accuracy and robustness. For example, blending Random Forests with LSTM outputs can yield more stable forecasts.
- Transfer Learning Uses pre-trained models on large datasets and fine-tunes them for specific weather tasks, especially useful when labeled data is limited.
- Reinforcement Learning Though still emerging in meteorology, it holds potential for adaptive forecasting systems that teach optimal prediction strategies over time.

Machine learning models offer flexible, scalable, and increasingly accurate alternatives to traditional numerical weather prediction. While traditional ML models are useful for structured, tabular data, deep learning models—especially LSTM and CNN architectures—excel in capturing temporal and spatial dependencies in complex weather systems. The choice of model depends on the forecasting objective, data type, and required prediction horizon. As data availability and computational power continue to grow, ML models are poised to play an even more central role in the future of weather forecasting.

D. Performance Evaluation

Model performance will be evaluated using metrics such as accuracy, precision, recall, F1-score, and confusion matrix analysis. ROC curves and Area Under the Curve (AUC) are used to assess classification robustness. Accuracy, precision, recall, F1-score, and confusion matrix analysis are key performance metrics used to evaluate machine learning models, especially in classification tasks like EEG-based emotion detection.

E. Tools and Environment

The implementation will be carried out using Python-based libraries such as TensorFlow, Keras, and MNE for weather forecasting. Data visualization and analysis will be supported by tools like Matplotlib, Seaborn, and Pandas.

5. TRENDS IN MACHINE LEARNING FOR WEATHER FORECASTING

1. Shift from Physics-Based to Data-Driven Models Traditional weather forecasting relies on numerical weather prediction (NWP) models, which solve complex physical equations. However, these models are computationally intensive and sensitive to initial conditions. Machine learning (ML), particularly deep learning, is now being used to complement or replace parts of these models by learning patterns directly from historical and real-time data.
2. Rise of Deep Learning Architectures Deep learning models such as LSTM (Long Short-Term Memory), CNN (Convolutional Neural Networks), and hybrid models (e.g., CNN-LSTM) are increasingly used for spatiotemporal forecasting. These models excel at capturing nonlinear dependencies and long-range temporal patterns, making them suitable for predicting variables like temperature, rainfall, and wind speed.
3. Integration of Remote Sensing and Satellite Data ML models are now being trained on high-resolution satellite imagery and geospatial data, enabling more accurate and localized forecasts. This trend is particularly useful for predicting extreme weather events such as cyclones, fog, and flash

floods.

4. AutoML and Foundation Models Automated machine learning (AutoML) frameworks are being adopted to optimize model selection and hyperparameter tuning. Additionally, foundation models—large pre-trained models adapted for weather prediction—are emerging as powerful tools for global-scale forecasting.
5. Hybrid and Ensemble Approaches Researchers are combining ML models with traditional NWP outputs to create hybrid systems that leverage the strengths of both. Ensemble learning techniques are also being used to improve robustness and reduce prediction uncertainty (Jorge, 2024).

A. Challenges in Machine Learning-Based Weather Forecasting

- Data Quality and Availability: ML models require large volumes of high-quality, labeled data. In many regions, especially developing countries, weather data is sparse, inconsistent, or unavailable. This limits the generalizability of models trained on localized datasets.
- Model Interpretability: Deep learning models often function as "black boxes," making it difficult to understand how predictions are made. This lack of transparency poses challenges for trust, especially in high-stakes applications like disaster management.
- Computational Demands: While ML models can be faster than full-scale NWP simulations once trained, the training process, especially for deep learning and foundation models requires significant computational resources and infrastructure.
- Generalization Across Regions and Scales: A model trained on data from one geographic region may not perform well in another due to differences in climate patterns, terrain, and data characteristics. Ensuring cross-regional robustness remains a major hurdle.
- Real-Time Forecasting and Deployment: Integrating ML models into operational forecasting systems requires real-time data ingestion, low-latency processing, and seamless deployment pipelines challenges that are still being actively addressed.

Machine learning is revolutionizing weather forecasting by offering faster, more flexible, and potentially more accurate alternatives to traditional models. However, to fully realize its potential, researchers must overcome challenges related to data, interpretability, scalability, and integration. Continued innovation in hybrid modeling, explainable AI, and global data sharing will be key to advancing the field (Gupta, 2024).

6. CONCLUSION

Machine learning has revolutionized weather forecasting by providing data-driven alternatives to numerical models. From classical algorithms like Support Vector Machines and Random Forests to advanced deep learning architectures like LSTM and CNN, ML techniques have shown promise in predicting meteorological variables with improved accuracy and efficiency. These models capture nonlinear patterns and temporal relationships well, making them ideal for complicated atmospheric systems. Recent trends suggest a move toward hybrid and ensemble models, satellite and sensor data integration, AutoML and foundation models for scalable forecasting. Recent advances enable localized, real-time, and adaptive weather forecast systems. The discipline still confronts major problems, including data scarcity in some places, deep learning model interpretability, and large-scale system training and deployment processing demands. Meteorologists, data scientists, and legislators must work together to solve these problems. Future research should construct explainable AI models, improve data accessibility, and build strong frameworks that can generalize across varied geographic and climatic situations. Operational forecasting systems using machine learning will provide more accurate, timely, and actionable weather insights, improving preparedness and resilience to climate variability and extreme weather occurrences.

7. FUTURE SCOPE

Future research on machine learning-based rainfall prediction can benefit from several strategic advancements. First, the integration of hybrid models that combine machine learning with physics-based atmospheric simulations could enhance predictive realism and reliability. Second, the development of climate-specific architectures tailored to tropical and temperate dynamics may improve generalization across diverse geographies.

ACKNOWLEDGMENT

The authors gratefully acknowledge the contributions of researchers and institutions whose pioneering work in machine learning and weather forecasting analysis has laid the foundation for this review. It extends our sincere thanks to the developers of open-access datasets and platforms, which have enabled reproducible and data-driven research in this domain.

REFERENCES

1. Breiman, L. Random forests. Mach. Learn. 2001, 45, 5–32.
2. Wilks, D.S.; Hamill, T.M. Comparison of ensemble-MOS methods using GFS reforecasts. Mon. Weather. Rev. 2007, 135, 2379–2390
3. Akaike, H. A New Look at the Statistical Model Identification. IEEE Trans. Automat Contr. 1974.
4. Ejike, O.; Ndzi, D.; Shakir, M.Z. Comparative Study of Machine Learning-Based Rainfall Prediction in Tropical and Temperate Climates. Climate 2025, 13, 167. https://doi.org/ 10.3390/cli13080167
5. Jankauskas, M.; Serackis, A.; Paulauskas, N.; Pomarnacki, R.; Hyunh, V.K. (2024) The Impact of the Weather Forecast Model on Improving AI-Based Power Generation Predictions through BiLSTM Networks. Electronics, 13, 3472. https://doi.org/10.3390/ electronics13173472
6. Guo, H.; Ma, Y.; Li, Z.; Zhao, Q.; Zhai, Y. (2024) The Evaluation of Rainfall Forecasting in a Global Navigation Satellite System-Assisted Numerical Weather Prediction Model. Atmosphere 15, 992. https://doi.org/ 10.3390/atmos15080992
7. Teixeira, R.; Cerveira, A.; Pires, E.J.S.; Baptista, J. (2024) Enhancing Weather Forecasting Integrating LSTM and GA. Appl. Sci., 14, 5769. https://doi.org/10.3390/ app14135769
8. Zhang, L.; Ma, Y.; Lei, L.; Wang, Y.; Jin, S.; Gong, W. (2024) Improving Atmospheric Temperature and Relative Humidity Profiles Retrieval Based on Ground-Based Multichannel Microwave Radiometer and Millimeter-Wave Cloud Radar. Atmosphere 15, 1064. https:// doi.org/10.3390/atmos15091064
9. Sim, S.; Im, J.; Jung, S.; Han, D. (2024) Improving Short-Term Prediction of Ocean Fog Using Numerical Weather Forecasts and Geostationary Satellite-Derived Ocean Fog Data Based on AutoML. Remote Sens., 16, 2348. https://doi.org/10.3390/ rs16132348
10. Shiferaw, A.; Tadesse, T.; Rowe, C.; Oglesby, R. (2024) Weather Research and Forecasting Model (WRF) Sensitivity to Choice of Parameterization Options over Ethiopia. Atmosphere, 15, 974. https://doi.org/10.3390/ atmos15080974
11. Boloukasli ahmadgourabi, F.; Khashei Varnamkhasti, M.; Nosrati Habibi, M.; Hedaiaty Marzouny,

N.; Dziedzic, R. (2024) Enhancing Water Demand Forecasting: Leveraging LSTM Networks for Accurate Predictions. Eng. Proc., 69, 120. https://doi.org/10.3390/ engproc2024069120

12. Ayyash, F.; Hayslep, M.; Ko, T.; Kalumba, M.; Simukonda, K.; Farmani, R. (2024) Application of a Neural Network Model to Short-Term Water Demand Forecasting. Eng. Proc., 69, 123. https://doi.org/10.3390/ engproc2024069123
13. Xian, Z.; Zhu, J.; Lin, S.-J.; Liang, Z.; Chen, X.; Chen, K. (2024) Impact of Assimilating Geostationary Interferometric Infrared Sounder Observations from Long- and Middle-Wave Bands on Weather Forecasts with a Locally Cloud-Resolving Global Model. Remote Sens. 16, 3458. https:// doi.org/10.3390/rs16183458
14. Jorge Díaz-Ramírez, Ximena Badilla-Torrico (2024), Comparative Analysis of Machine Learning Techniques For Forecasting Weather: A Case Study, International Journal of Recent Advances In Multidisciplinary Topics, 12(2), 305-313.
15. Bajad T., Rawat, A., Kumar, A., & Singh, M. (2024). Weather prediction using machine learning. SSRN Electronic Journal 160-163. https://doi.org/10.2139/ssrn.4489040
16. Gupta Y., Sharma S. (2024), Advancements in Weather Forecasting Through Machine Learning Algorithms, International Journal Of Recent Advances In Multidisciplinary Topics, 145-148.