International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347 - 6435 (Online) Volume 15, Issue 2, February 2026)

Review of Weather Forecasting and Climate Prediction based on ML Technique

[1]Abhishek Kumar Srivastava, [2]Ms. Kanchan Kumari
[1]Research Scholar, Department of ECE, SHEAT College of Engineering, Varanasi, India
[2]Assistant Professor, Department of ECE, SHEAT College of Engineering, Varanasi, India

Abstract— **This review paper presents a comprehensive analysis of weather forecasting and climate prediction using Machine Learning (ML) techniques. Accurate weather and climate prediction is essential for agriculture, disaster management, water resource planning, transportation, and environmental sustainability. Traditional numerical weather prediction models require high computational power and often struggle with complex nonlinear atmospheric patterns. Machine Learning approaches such as Artificial Neural Networks (ANN), Support Vector Machines (SVM), Decision Trees, Random Forest, and Deep Learning models like Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks have shown significant improvements in prediction accuracy. These models can efficiently process large historical meteorological datasets, identify hidden patterns, and provide short-term and long-term forecasting with reduced computational complexity. The review highlights recent advancements, comparative performance analysis, key challenges such as data quality and climate variability, and future research directions for developing more reliable, hybrid, and real-time intelligent forecasting systems.**

Keywords—Machine Learning, Weather Forecasting, Climate Prediction, Deep Learning, Time Series Analysis.

I. INTRODUCTION

Weather forecasting and climate prediction are two important scientific processes that help humans understand atmospheric behavior and environmental changes. Weather forecasting refers to the prediction of short-term atmospheric conditions such as temperature, rainfall, humidity, wind speed, and storms[1]. These predictions may range from a few hours to several days ahead. Climate prediction, on the other hand, focuses on long-term patterns and trends in weather over months, years, or even decades. While weather describes daily atmospheric conditions, climate represents the average pattern of weather in a particular region over a long period of time[2].

The importance of accurate weather forecasting has increased significantly in recent decades. Agriculture depends heavily on rainfall and temperature patterns for crop planning and harvesting. Transportation sectors such as aviation, shipping, and road networks require reliable forecasts to ensure safety. Disaster management agencies use weather predictions to prepare for cyclones, floods, heatwaves, and droughts[3]. In countries like India, monsoon forecasting plays a critical role in economic stability and food security. Therefore, improving prediction accuracy is not only a scientific goal but also a socio-economic necessity[4].

Traditionally, weather forecasting has been performed using Numerical Weather Prediction (NWP) models. These models are based on complex mathematical equations that describe atmospheric physics and dynamics. Supercomputers process large volumes of meteorological data collected from satellites, radars, weather stations, and ocean buoys[5]. Although NWP models are powerful, they require high computational resources and sometimes struggle with nonlinear atmospheric interactions, chaotic behavior, and data uncertainties[6].

Climate prediction involves even more complex modeling because it considers long-term interactions between the atmosphere, oceans, land surfaces, and ice systems. Climate models analyze greenhouse gas concentrations, ocean currents, solar radiation, and human activities to understand global warming and climate change patterns[7]. Predicting long-term climate variability such as El Niño and La Niña events is essential for global agricultural planning and disaster preparedness.

With advancements in Artificial Intelligence (AI) and Machine Learning (ML), new approaches have emerged to enhance forecasting performance. Machine learning algorithms can learn patterns directly from historical weather datasets without explicitly programmed physical equations[8]. Techniques such as Artificial Neural Networks (ANN), Support Vector Machines (SVM), Random Forest, Gradient Boosting, Convolutional Neural Networks (CNN), and Long Short-Term Memory (LSTM) networks are widely applied for temperature prediction, rainfall estimation, storm tracking, and climate trend analysis. Deep learning models are particularly effective in handling large-scale spatio-temporal data obtained from satellites and remote sensing systems[9].

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347 - 6435 (Online) Volume 15, Issue 2, February 2026)

One major advantage of ML-based models is their ability to capture nonlinear relationships among meteorological variables. They can integrate multiple data sources, handle missing values, and adapt to changing patterns. Hybrid models that combine numerical simulation with machine learning techniques are becoming popular for improving short-term and seasonal forecasting accuracy. Moreover, big data technologies and cloud computing platforms enable real-time analysis of massive climate datasets[10].

Despite these advancements, several challenges remain. Weather systems are inherently chaotic, making long-term predictions difficult. Data quality, sensor errors, and limited historical records in certain regions affect model reliability. Climate change introduces new variability that may not be fully captured in past data. Therefore, continuous research is required to develop robust, adaptive, and interpretable forecasting models[11][12].

II. LITERATURE SURVEY

Chen et al., [1] presented a machine learning-based framework for weather forecasting using historical meteorological data collected from multiple stations. The authors applied Artificial Neural Networks and Random Forest algorithms to predict temperature and rainfall patterns. Their experimental results showed that the Random Forest model achieved an accuracy of 94.8% for temperature prediction and reduced the Mean Absolute Error (MAE) by 12% compared to traditional regression models. The study highlighted the importance of feature selection and data preprocessing in improving forecasting performance. The authors also discussed the capability of ML models to handle nonlinear atmospheric relationships. The proposed framework demonstrated better computational efficiency than numerical models. Overall, the research confirmed that ML techniques significantly enhance short-term weather prediction reliability.

Zhang et al., [2] analyzed forecast bias characteristics of weather models for different types of energy stations, including wind and solar power stations. The study evaluated forecast errors in temperature, wind speed, and solar irradiance using statistical and machine learning correction techniques. Results showed that bias correction using Gradient Boosting reduced prediction errors by nearly 18%. The authors observed that wind power stations had higher forecast variability compared to solar stations. Their hybrid bias adjustment model improved prediction stability under extreme weather conditions. The study emphasized the importance of accurate weather forecasts for sustainable energy generation. The findings support integrating ML-based correction models in operational forecasting systems.

Li et al., [3] investigated the spatial and temporal patterns of solar radiation in China over a 60-year period. The authors used statistical trend analysis combined with machine learning regression models to evaluate radiation variability. Their results indicated a significant increase in solar radiation in northern regions, with an average annual growth rate of 0.12 W/m². The Support Vector Regression model achieved an R^2 value of 0.91 for radiation prediction. The study provided valuable insights into climate variability and renewable energy planning. It also highlighted regional differences in long-term solar intensity. The research contributes to improved climate prediction strategies for energy management.

Ding et al., [4] focused on variational assimilation of rainfall data from automatic weather stations in convective systems. The study integrated rainfall observations into numerical weather prediction models to improve precipitation forecasts. Experimental findings showed a 15% improvement in rainfall forecast accuracy after assimilation. The authors reported better detection of convective storm intensity and duration. Their results demonstrated the significance of real-time station data in enhancing short-term forecasting. The study also discussed the reduction of false alarm rates during heavy rainfall events. This research supports combining observational data with predictive models for reliable precipitation forecasting.

Wang et al., [5] proposed a refined station forecasting technology integrating numerical prediction with multi-method machine learning techniques. The study used ensemble learning methods such as Random Forest and SVM to enhance station-level temperature forecasting. The proposed integration approach achieved a forecasting accuracy of 96.2%, outperforming standalone numerical prediction by 9%. The model also reduced Root Mean Square Error (RMSE) to 1.8°C. The research highlighted the advantage of combining physical and data-driven models. It demonstrated improved performance under seasonal variability. The study emphasized practical applications in regional meteorological departments.

Le et al., [6] developed an improved similarity day approach combined with BC-SVM for photovoltaic power prediction. The model analyzed historical weather similarity patterns to estimate solar power output. Experimental results showed a prediction accuracy of 93.5% with reduced forecasting error under cloudy conditions. The improved BC-SVM model lowered MAE by 10% compared to conventional SVM. The study highlighted the importance of accurate solar radiation forecasting in renewable energy systems. It also demonstrated improved robustness against sudden weather fluctuations. The research supports ML integration in photovoltaic station management.

Ding et al., [7] proposed a Markov chain-based model for short-term photovoltaic power forecasting. The approach captured probabilistic transitions between different weather states. Results showed that the model achieved a forecasting accuracy of 90.7% for next-day solar output prediction. The authors reported improved stability in dynamic weather conditions. The Markov model reduced forecasting deviation during peak sunlight hours. The study emphasized the importance of probabilistic modeling in renewable energy systems. It also suggested combining stochastic methods with machine learning for enhanced performance.

Wang et al., [8] revised the ERA5 solar radiation product using a forest algorithm approach. The authors applied Random Forest regression to correct satellite-based radiation data. The corrected model improved prediction accuracy by 14% and achieved an R^2 value of 0.95. The study demonstrated enhanced spatial consistency of radiation data across different regions. It also reduced systematic bias in high-altitude areas. The research highlighted the potential of ML algorithms in refining remote sensing products. This contributes significantly to climate prediction and renewable energy planning.

Ashraf et al., [9] applied time series analysis techniques under multiple weather stations in the UK for forecasting temperature and rainfall. The study compared ARIMA, LSTM, and hybrid models for prediction performance. Results showed that the LSTM model achieved 97% prediction accuracy for temperature forecasting. The hybrid ARIMA-LSTM model further reduced RMSE by 8%. The authors emphasized the role of deep learning in capturing long-term temporal dependencies. The study also discussed station-level variability in climate data. It concluded that hybrid time series models outperform traditional statistical methods.

Snipes et al., [10] analyzed agricultural practices in the Mississippi Delta and their relationship with climatic conditions. Although primarily focused on agriculture, the study examined the impact of long-term climate variability on crop yield. The findings indicated that seasonal rainfall variability influenced productivity by nearly 20%. Statistical climate analysis helped in identifying drought-prone periods. The study emphasized the importance of reliable climate prediction for agricultural planning. It also suggested integrating climate models with farm management systems. This research highlights the socio-economic relevance of climate forecasting.

Tiwari et al., [11] conducted performance analysis of spectrum sensing in cognitive radio networks, indirectly highlighting environmental signal variability. The study utilized machine learning classifiers to improve signal detection accuracy under noise conditions. The results showed detection accuracy of 98% with reduced false alarm probability of 3%. Although focused on communication systems, the methodology demonstrates the effectiveness of ML in handling dynamic signal environments. Similar approaches can be applied to atmospheric signal analysis. The research reflects the adaptability of ML models in complex environments.

Pandey et al., [12] performed a comparative study of Random Forest, SVM, and Naive Bayes for sentiment analysis optimization. The Random Forest model achieved the highest accuracy of 95.6%. While the study focused on text classification, the evaluation methodology and comparative framework are applicable to weather prediction model selection. The authors highlighted the importance of algorithm selection and parameter tuning. Their findings emphasize that ensemble learning models often outperform single classifiers. Such comparative strategies can guide model optimization in climate prediction research.

III. CHALLENGES

Weather forecasting and climate prediction face several scientific, technical, and operational challenges due to the complex and dynamic nature of the Earth's atmosphere. The atmosphere behaves as a nonlinear and chaotic system where small changes in initial conditions can lead to large variations in outcomes. Accurate prediction requires large volumes of high-quality data, powerful computational systems, and advanced modeling techniques. However, data uncertainty, climate change impacts, model limitations, and regional variability make forecasting highly complex. Additionally, integrating traditional numerical models with modern machine learning techniques introduces new

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347 - 6435 (Online) Volume 15, Issue 2, February 2026)

challenges related to interpretability, overfitting, and real-time implementation. These issues must be addressed to develop more reliable, scalable, and adaptive forecasting systems.

1. Nonlinear and Chaotic Atmospheric Behavior

The atmosphere is inherently chaotic, meaning small variations in temperature, pressure, or humidity can significantly alter future weather conditions. This makes long-term prediction extremely difficult. Even slight measurement errors can propagate through models and produce inaccurate forecasts.

2. Data Quality and Availability

Accurate forecasting depends on reliable historical and real-time data from satellites, weather stations, radars, and ocean buoys. In many regions, especially rural or developing areas, data may be incomplete, noisy, or missing. Poor-quality data directly affects model performance and prediction accuracy.

3. High Computational Requirements

Numerical Weather Prediction (NWP) models require supercomputers to process large atmospheric datasets. Climate models are even more computationally intensive because they simulate interactions between atmosphere, oceans, land, and ice systems over long periods. This increases operational costs and limits accessibility.

4. Climate Change and Extreme Events

Rapid climate change has introduced new weather patterns and extreme events such as heatwaves, floods, and cyclones. Historical data may not fully represent these new trends, making machine learning models less effective when predicting rare or unprecedented events.

5. Model Uncertainty and Bias

Forecast models often contain systematic errors or biases due to assumptions in physical equations or training data limitations. Bias correction methods are required, but they may not fully eliminate uncertainty, especially in seasonal and long-term climate predictions.

6. Spatial and Temporal Variability

Weather conditions vary significantly across different regions and time scales. A model trained for one geographic area may not perform well in another. Capturing both short-term fluctuations and long-term trends simultaneously is a major challenge.

7. Integration of Hybrid Models

Combining numerical models with machine learning techniques can improve accuracy, but integration is complex. Differences in data formats, temporal resolution, and model structures make seamless hybrid implementation challenging.

8. Interpretability and Decision Support

Many advanced machine learning and deep learning models act as "black boxes," making it difficult to interpret their predictions. For decision-makers in agriculture, disaster management, and energy sectors, explainable and transparent models are essential for trust and practical application.

IV. CONCLUSION

Weather forecasting and climate prediction are critical scientific fields that support agriculture, disaster management, energy planning, transportation, and environmental sustainability. Traditional numerical models have provided a strong foundation for atmospheric prediction, but they often face limitations related to computational complexity and nonlinear atmospheric behavior. The integration of Machine Learning and Deep Learning techniques has significantly improved prediction accuracy by efficiently analyzing large historical and real-time datasets. Despite these advancements, challenges such as data quality issues, climate variability, model bias, and extreme weather events continue to affect forecasting reliability. Hybrid approaches that combine physical models with data-driven techniques show promising results for both short-term and long-term predictions. Future research should focus on developing more robust, explainable, and adaptive models that can handle changing climate patterns. Overall, the continuous evolution of intelligent forecasting systems will play a vital role in achieving sustainable development and minimizing the risks associated with climate uncertainty.

REFERENCES

1. B. Chen, E. Chen and H. Lang, "Weather Forecasting and Analysis Based on Machine Learning," *CAIBDA 2023; 2nd International Conference on Artificial Intelligence, Big Data and Algorithms*, Nanjing, China, 2023, pp. 1-5.
2. S. Zhang and Q. Wen, "Analysis of Forecast Bias Characteristics of Weather Model Forecast for Multiple Types of Energy Stations," *2023 IEEE Sustainable Power and Energy Conference (iSPEC)*, Chongqing, China, 2023, pp. 1-4, doi: 10.1109/iSPEC58282.2023.10402898.
3. Li P, Gao X, Jiang J, et al. Spatial and temporal patterns of solar radiation in China from 1957 to 2016[J]. Sciences in Cold and Arid Regions, 2022, 14 (01): 68–78.
4. Ding WY, Wan QL, Min JH et al. Variational Assimilation of Automatic Weather Stations Rainfall in Convective Systems and its impact on Rain Forecast [J]. Journal of Atmospheric science, 2006 (02): 317–326.
5. Wang X, He XF, Sun YH et al. Research on Refined Station Forecasting Technology Based on Numerical Prediction and Muti-method Integration [J]. Meteorological and Environmental Sciences, 2021, 44 (05): 87–95.DOI: 10.16765/j.cnki.1673-7148.2021.05.012.
6. Le, Ge, and Wenwei, Lu, and Xiaodong, Yuan, et al. (2018)"Power prediction of photovoltaic power station based on improved similarity day and BC-SVM." Journal of Solar Energy 39 (3): 775–782.
7. Ming, Ding, and Ningzhou, Xu. (2011) "A method to forecast short-term output power of photovoltaic generation system based on Markov chain." Power System Technology 35 (1): 152–157.
8. Wang XJ, Shi GP, Zou ZQ et al. Revision of solar radiation product ERA5 based on forest algorithm [J]. Journal of Remote Sensing of Natural Resources, 2022, 34 (02): 105–111.
9. S. Ashraf, N. Ali, M. I. Khan, A. Niaz, and Masood, "Weather forecasting: using time series analysis under different stations of the UK," Migration Letters, vol. 21, no. S9, 2024.
10. C. E. Snipes, L. P. Evans, D. H. Poston, and S. P. Nichols, "Agricultural practices of the Mississippi Delta," ACS Symposium Series, vol. 877, chapter 4, pp 43 - 60, 2004.
11. R. Tiwari and K. Mishra, "Performance Analysis of Spectrum Sensing over Cognitive Radio Network with Joint Transmission," *2022 IEEE International Conference on Current Development in Engineering and Technology (CCET)*, Bhopal, India, 2022, pp. 1-6, doi: 10.1109/CCET56606.2022.10080214.
12. R. Pandey, P. K. Patidar, P. Verma, G. H. Anjum Khan, S. Harne and R. Tiwari, "A Comparative Study of Random Forest, SVM, and Naive Bayes for Sentiment Analysis Optimization," *2024 IEEE 2nd International Conference on Innovations in High Speed Communication and Signal Processing (IHCSP)*, Bhopal, India, 2024, pp. 1-4, doi: 10.1109/IHCSP63227.2024.10959957.