Optimized Deep Learning Models for Monthly Weather Forecast in Kogi, Nigeria

Shaibu Hamza | Abisoye Opeyemi Aderiike | Joshua Babatunde Agbogun

Malik Adeiza Rufai | Bello Ojochide Joy

ABSTRACT
Accurate weather forecasting is essential for agriculture, disaster preparedness, and economic planning in Nigeria, yet existing approaches such as Numerical Weather Prediction (NWP) face challenges of computational intensity and limited accuracy for localized monthly predictions. This study develops optimized deep learning models for monthly weather forecasting using Nigerian meteorological data from 2014 to 2023, with a focus on the Kogi region. Three deep learning architectures: Artificial Neural Networks (ANN), Recurrent Neural Networks (RNN), and Long Short-Term Memory (LSTM), were implemented and optimized using Bayesian hyperparameter tuning. To further enhance predictive performance, a Boosting Ensemble approach integrating the three models was proposed. Model evaluation employed Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) as benchmarks. Results showed that while RNN outperformed ANN and LSTM individually, the Boosting Ensemble achieved the best accuracy, with the lowest RMSE (52.351) and MAE (29.475), consistently capturing both stable and transitional weather patterns. The findings demonstrate that ensemble deep learning methods significantly improve monthly weather forecasting accuracy compared to standalone models. This study contributes a scalable, data-driven framework tailored to Nigeria's climatic conditions, offering practical value for farmers, policymakers, and disaster management agencies, while also providing a foundation for future research incorporating additional climatic variables and advanced attention-based models.

Keywords
Forecasting, Deep Learning (DL), Artificial Neural Network (ANN), Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), Ensemble Learning, Bayesian Optimization

1. INTRODUCTION
Weather forecasting involves predicting variations in atmospheric conditions such as temperature, precipitation, wind, and humidity using historical data, scientific models, and real-time observations [13]. Forecasts are categorized into short-term (hours to days), medium-term (weeks), and long-term (months to years) predictions [9;11]. Accurate forecasts are critical to agriculture, transportation, disaster management, and other sectors that rely on weather-sensitive planning [20].

Traditional forecasting has relied on Numerical Weather Prediction (NWP) models, which simulate atmospheric processes mathematically [4]. While widely adopted in operational systems such as the Nigerian Meteorological Agency (NiMET), NWP methods are computationally intensive and struggle with the uncertainty inherent in modeling complex atmospheric dynamics [17]. Recent advances in machine learning (ML) have introduced data-driven alternatives capable of handling nonlinear dependencies in weather data. In particular, deep learning (DL) methods such as Artificial Neural Networks (ANN), Recurrent Neural Networks (RNN), and Long Short-Term Memory (LSTM) networks have shown strong performance in time-series weather forecasting [16;18;19;21]

Despite progress, existing Nigerian-focused forecasting models remain limited. The NiMET system is constrained by its reliance on NWP methods, while recent ML-based approaches have been restricted to annual forecasting [7;15], offering little insight into monthly trends crucial for agriculture and climate-sensitive decision-making. Furthermore, many studies rely on outdated datasets, reducing accuracy in the context of recent climate variability [2].

This study addresses these gaps by developing optimized deep learning models for monthly weather forecasting in Nigeria. Specifically, it applies ANN, RNN, and LSTM architectures to monthly weather data from Kogi State, employing Bayesian optimization for hyperparameter tuning and constructing a hybrid ensemble model for improved predictive accuracy. Performance is evaluated using standard metrics including Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE).

The contribution of this research is threefold: (i) the development of Nigeria-specific DL models tailored for monthly weather forecasting, (ii) the optimization of model performance through Bayesian tuning and ensemble methods, and (iii) the provision of a scalable, cost-effective framework for accurate weather prediction in data-limited regions. By bridging methodological and regional gaps, this work aims to support agriculture, disaster preparedness, and sustainable resource planning in Nigeria and other developing nations vulnerable to climate variability.

2. LITERATURE REVIEW
2.1 Theoretical Background
Weather forecasting relies on traditional physics-based numerical weather prediction (NWP) and emerging machine learning (ML) and deep learning (DL) methods. While NWP achieves reliable short- to medium-term forecasts, it is computationally intensive and sensitive to initial conditions

[3;12] In contrast, DL models learn complex spatiotemporal patterns from large datasets, making them suitable for nonlinear, high-dimensional weather data [1]. CNNs capture spatial dependencies in satellite imagery, while RNNs, LSTMs, and Transformers handle temporal sequences [10]. Hybrid and physics-informed approaches are increasingly explored for improved accuracy [17].

Despite progress, challenges remain. DL models require vast, high-quality datasets and computational resources [5]. They often function as "black boxes," limiting interpretability, and they struggle with extreme or rare events. Key performance metrics include MAE, RMSE, and CSI, with growing interest in probabilistic evaluation and explainable AI (XAI).

2.2 Forecasting Approaches

- NWP Models: Offer detailed climate simulations but demand significant computational resources and may fail to capture localized phenomena [5].
- Statistical and ML Models: Include regression, ARIMA, random forests, and SVMs, which improve efficiency but often lack the ability to model nonlinear dynamics fully [14].
- Deep Learning Models: CNNs, LSTMs, GRUs, and hybrid CNN-LSTM/GRU frameworks have shown superior performance for rainfall, air temperature, soil moisture, and wind speed forecasting [9;10;19;21]. Attention-based and ensemble models further improve accuracy but add computational overhead.

2.3 Related Studies

Recent works emphasize hybrid DL architectures. Guo et al. (2024) found CNN-LSTM models superior for climate forecasting in Jinan, China. [19] confirmed CNN-LSTM-GRU hybrids improve one-day-ahead air temperature prediction in Türkiye. [21] demonstrated attention-based LSTMs outperform conventional models for soil moisture. [3] LSTM and Stacked-LSTM excel in rainfall forecasting for UK cities. Similar advances include wind speed prediction with CNN-BLSTM and WRF integration [10], uncertainty reduction via mixed models [8], and extreme event prediction with Capsule Networks [6].

Across these studies, DL consistently outperforms statistical baselines. However, most are geographically narrow (India, China, Türkiye, Brazil, UK), computationally expensive, and lack explicit model optimization (e.g., grid search, Bayesian optimization).

2.4 Summary and Research Gaps

The literature reveals four main gaps:

1. Regional limitation – Most models are validated outside Africa, limiting generalizability to Nigerian contexts [9;14].
2. Data challenges – Use of outdated or single-source datasets reduces robustness against climate change trends [15]
3. Model optimization – Few studies apply hyperparameter tuning (e.g., Bayesian optimization, AutoML) to maximize performance.
4. Efficiency and interpretability – Many models achieve accuracy at high computational cost, with limited emphasis on interpretability or real-time application [10;5].

This study addresses these gaps by applying optimized DL techniques to Nigerian monthly weather forecasting, integrating diverse datasets, and emphasizing both efficiency and accuracy.

3. RESEARCH METHODOLOGY

3.1 Research Design

This study employs a quantitative experimental design aimed at developing accurate monthly weather forecasting models for Nigeria. A quantitative approach is suitable because the objectives revolve around numerical data analysis, model training, and performance evaluation.

The design follows a sequential process:

1. Collection of monthly meteorological records from NiMET for a 10-year period (2014–2023).
2. Preprocessing of the dataset, including missing data handling, normalization, and feature engineering.
3. Model development using advanced machine learning and deep learning architectures.
4. Optimization of models through hyperparameter tuning and ensemble learning.
5. Evaluation of results with statistical performance metrics.

This structure ensures a replicable, data-driven approach that aligns with best practices in computational forecasting research.

3.2 Methodological Framework

The methodological framework is iterative, with each phase designed to strengthen the next (Figure 1). It comprises:

1. Data Collection – Reliable monthly weather data from NiMET and climatological archives.
2. Preprocessing – Cleaning, transformation, and preparation of the dataset.
3. Modeling – Selection and training of ANN, RNN, and LSTM models.
4. Optimization – Bayesian optimization, genetic algorithms, and boosting ensembles.
5. Evaluation – Performance measurement with RMSE, MAE, and accuracy metrics.

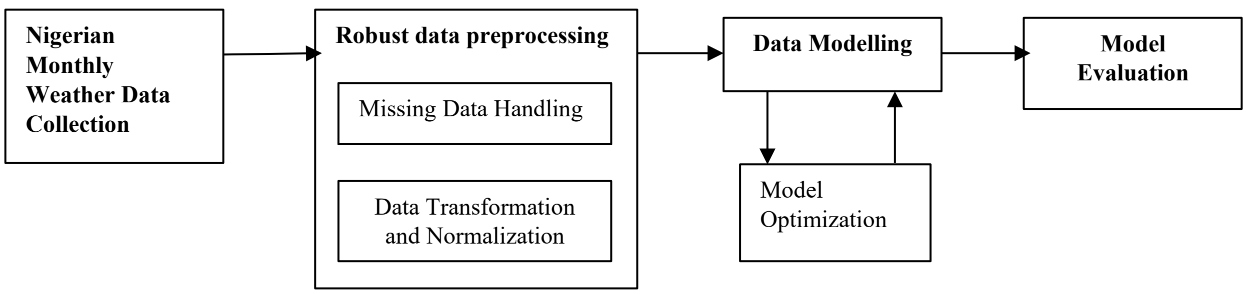

Fig. 1: Research Methodological Framework

This pipeline ensures that data quality, modeling rigor, and evaluation validity are preserved throughout the study.

3.3 Data Collection

The study utilizes secondary meteorological data obtained from the Nigerian Meteorological Agency (NiMET). A formal request was submitted, and data covering the Kogi region from 2014–2023 were provided.

The dataset includes:

- Temperature (°C)
- Rainfall (mm)
- Humidity (%)
- Wind speed (m/s)
- Atmospheric pressure (hPa)

The Kogi region was selected due to its diverse climatic characteristics and central geographic location, making it suitable for evaluating regional prediction performance. The long time span (10 years) enhances temporal robustness and enables training of deep learning models with adequate seasonal variation.

3.4 Data Analysis Techniques

The data analysis phase employed both individual deep learning models and a hybrid ensemble approach.

3.4.1 Artificial Neural Networks (ANN)

ANNs (Multilayer Perceptrons) learn nonlinear input-output relationships through interconnected layers of neurons. They are suitable for extracting general patterns but lack sequential memory, which limits their use in time-series forecasting.

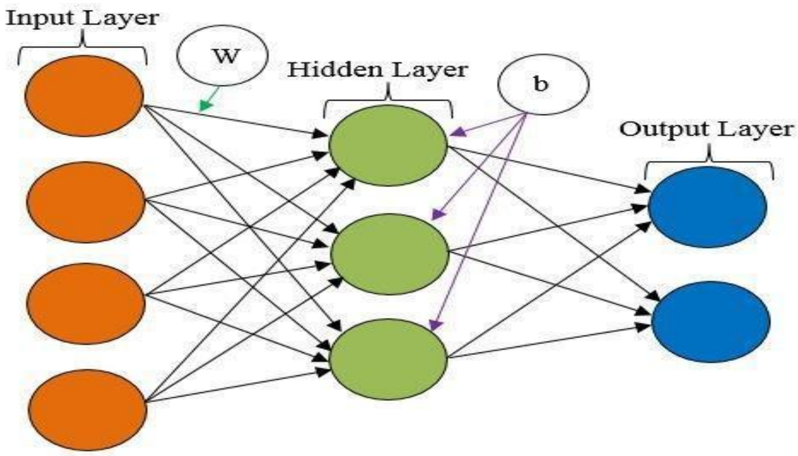

Fig. 2: Multilayer Perceptrons (MLP) Architecture

3.4.2 Recurrent Neural Networks (RNN)

RNNs introduce hidden states that allow the network to learn temporal dependencies. However, standard RNNs often struggle with vanishing gradients, which reduce their effectiveness for long-term dependencies.

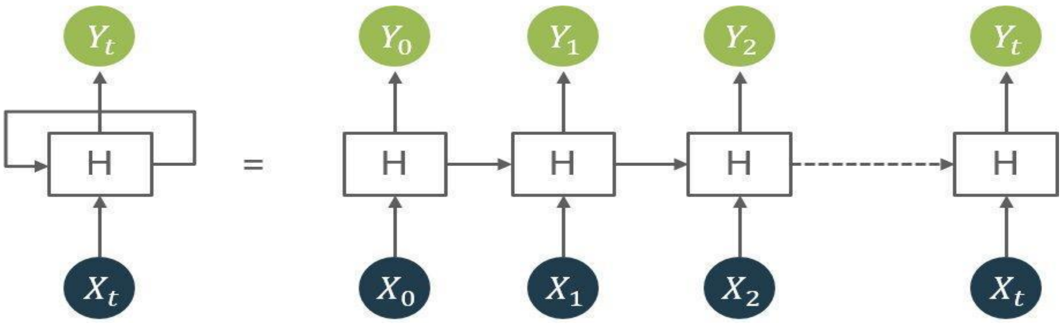

Fig. 3: Recurrent Neural Network (RNN) Architecture

3.4.3 Long Short-Term Memory (LSTM)

LSTMs overcome RNN limitations through gating mechanisms (input, forget, and output gates). These allow the network to selectively retain long-term information, making LSTM particularly effective for monthly weather prediction.

3.4.4 Hybrid Ensemble Model

To maximize predictive power, a Boosting Ensemble was designed that integrates ANN, RNN, and LSTM models. The ensemble works by training models sequentially: ANN first, followed by RNN on residual errors, and finally LSTM to capture remaining complexity. Predictions from all models are combined using weighted integration.

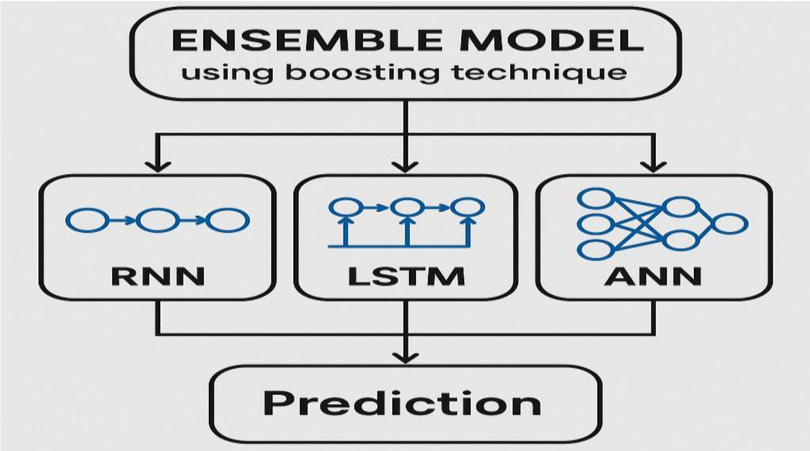

Fig. 4: Boosting Ensemble Architecture

3.5 Summary of Models

Table 1: Summary of Models and Their Strengths/Limitations

Model	Strengths	Limitations
ANN	Captures nonlinear relationships	No memory of sequential data
RNN	Learns short-term temporal dependencies	Struggles with long-term dependencies
LSTM	Handles long-term dependencies	Computationally expensive
Ensemble	Combines strengths of all models	Requires more training resources

3.6 Model Optimization

Optimization was carried out through:

- Hyperparameter tuning (learning rate, hidden layers, batch size).
- Bayesian optimization for efficient parameter search.
- Metaheuristic methods (e.g., Genetic Algorithm) for exploring broader parameter spaces.
- Ensemble boosting for error reduction and robustness.

These techniques ensured improved generalization and minimized overfitting.

3.7 Model Evaluation

The models were evaluated using standard forecasting metrics:

- Root Mean Squared Error (RMSE): Penalizes large errors more heavily.
- Mean Absolute Error (MAE): Provides an interpretable average error magnitude.
- Forecast Accuracy (%): Allows comparison across models.

Baseline models were compared with the optimized ensemble to quantify improvements.

3.8 Justification of Methodology

The methodology is justified on the following grounds:

1. Suitability for time-series forecasting: Deep learning models capture nonlinear, dynamic weather patterns better than traditional methods.
2. Robustness through ensembles: Boosting integrates multiple models, improving accuracy and reducing overfitting.
3. Reliability of dataset: Use of NiMET official data (2014–2023) ensures validity.
4. Scientific rigor: The inclusion of systematic optimization and evaluation strengthens reproducibility.

4. RESULTS AND DISCUSSION

4.1 Introduction

This chapter presents a comprehensive analysis of the experimental results for monthly weather forecasting in the Kogi region of Nigeria. We evaluated three individual deep learning models—Artificial Neural Network (ANN), Recurrent Neural Network (RNN), and Long Short-Term Memory (LSTM), as well as a Boosting Ensemble approach that sequentially integrates these models. Performance metrics (RMSE and MAE) and graphical analyses are presented to demonstrate model accuracy and robustness.

4.2 Description of the Collected Dataset

The dataset was obtained from NiMET and consists of monthly weather records for 10 years (2014–2023), comprising 120 entries. It contains three columns:

Table 2: Structure of the Collected Monthly Weather Dataset (2014–2023)

Column	Type	Non-null	Description
YEARS	int64	120	Observation year
MONTH	object	120	Month name

Column	Type	Non-null	Description
TEMPERATURE	float64	102	Average monthly temperature (°C)

Eighteen missing temperature values (out of 120) were imputed using mean values, ensuring temporal continuity for time-series modeling. This dataset structure is suitable for learning seasonal and long-term patterns relevant to Nigerian climate conditions.

4.3 Dataset Preparation for Modeling

Before model development, the dataset underwent preprocessing and transformation:

1. Missing Value Handling: Mean imputation replaced missing temperature values, preserving the overall data distribution.
2. Normalization: The "temperature" column was scaled to [0,1] using MinMaxScaler for stable neural network training.
3. Time-Series Structuring: A sliding window of 12 months was used to create supervised input-output pairs, where the previous 12 months predict the next month.
4. Train-Test Split: Data was divided chronologically, with 80% for training and 20% for testing to maintain temporal dependencies.

These steps ensured a clean, structured, and temporally consistent dataset suitable for ANN, RNN, LSTM, and ensemble modeling.

4.4 Model Development and Optimization

Each model was optimized using Bayesian Optimization for three key hyperparameters: number of neurons (units), learning rate, and batch size.

Table 3: Optimized Hyperparameter Configurations for ANN, RNN, and LSTM Model

Model	Best Configuration	Observations
ANN	85 neurons, LR=0.00167, Batch=16	Stable learning, moderate trend capture, slight overreactions
RNN	128 units, LR=0.00063, Batch=57	Better sequential learning, smoother updates, slower adaptation to sharp shifts
LSTM	21 units, LR=0.00886, Batch=27	Quick convergence, potential overfitting, limited long-term dependency capture

The Boosting Ensemble sequentially combines ANN, RNN, and LSTM outputs, with each model correcting residual errors from its predecessor.

4.5 Model Performance Evaluation

Table 4: Model Performance on Test Data

Model	RMSE	MAE
ANN	54.531	31.668
RNN	53.461	28.653
LSTM	59.791	35.476
Boosting Ensemble (ANN+RNN+LSTM)	52.351	29.475

Analysis:

- ANN: Captures general trends but overreacts to minor fluctuations.
- RNN: Performs better than ANN and LSTM by learning sequential dependencies effectively, though it lags during abrupt trend shifts.
- LSTM: Underperforms due to small network size and high learning rate, resulting in overfitting.
- Boosting Ensemble: Achieves the lowest RMSE and MAE by integrating complementary strengths and correcting residuals, producing stable and accurate forecasts.

4.6 Graphical Analysis

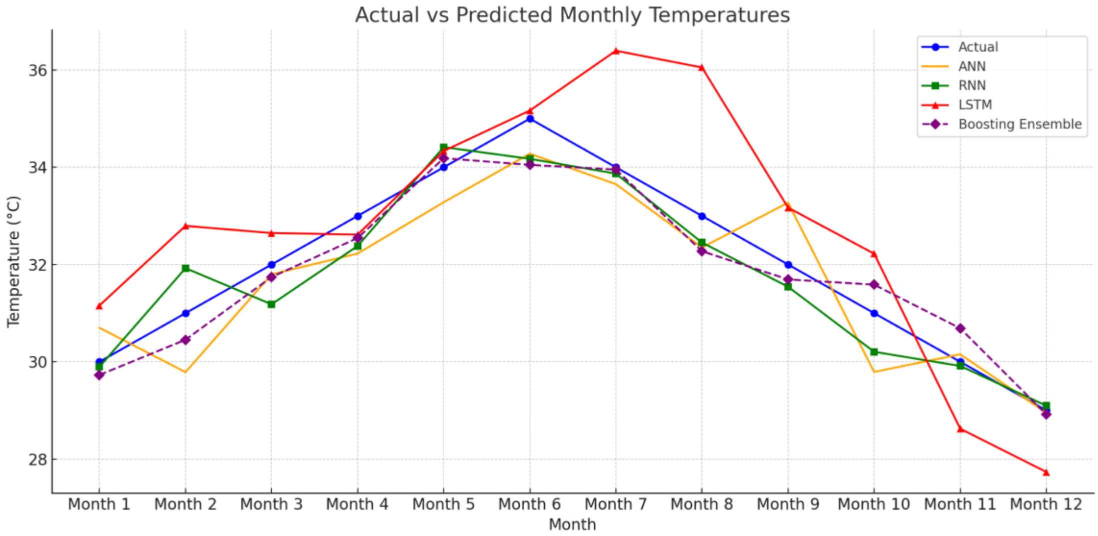

Fig. 5: Comparison of Actual vs Predicted Temperatures (1st Feature)

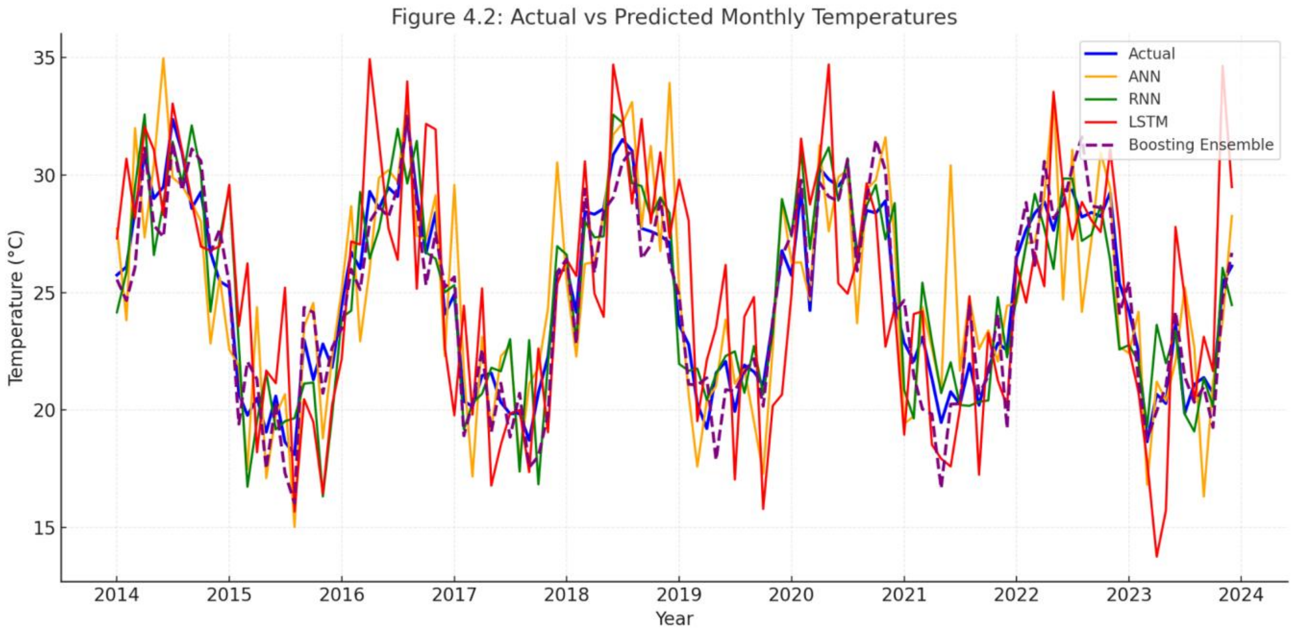

Fig. 5: Comparison of Actual vs Predicted Temperatures (2014 -2024)

- Blue Line: Actual temperature values
- Orange Line (ANN): Captures trend direction but fluctuates in stable periods
- Green Line (RNN): Smooth early predictions but fails to adjust to trend shifts
- Red Line (LSTM): Tracks transitions better than ANN/RNN but deviates slightly post-shift
- Purple Dashed Line (Boosting Ensemble): Closely follows actual data, with minimal fluctuations

Observations:

- The Boosting Ensemble consistently tracks both stable and transitional patterns.
- ANN overfits short-term variations, RNN lags on shifts, and LSTM partially overfits due to hyperparameter choice.
- Integration in the ensemble mitigates individual model weaknesses.

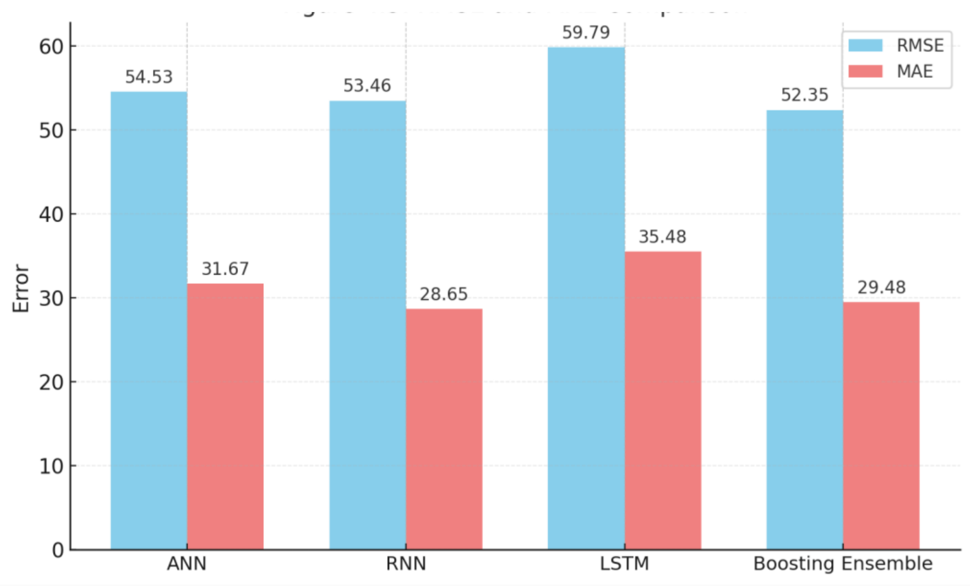

Fig 6: Error Metrics Comparison (RMSE & MAE)

- Ensemble achieves lowest RMSE and competitive MAE.
- Visual comparison reinforces numerical findings and confirms ensemble superiority.

4.7 Discussion

The results highlight that no single deep learning model was universally superior. While ANN was simple and responsive, it overfitted local variations. RNN excelled in short-term dependencies but lacked long-memory capacity. Surprisingly, the LSTM, despite its theoretical advantage in long-term learning, did not surpass RNN, likely due to architecture size and learning rate trade-offs.

The Boosting Ensemble provided the best compromise, integrating the strengths of ANN, RNN, and LSTM while mitigating their individual weaknesses. This approach aligns with prior findings e.g [6;8], which emphasize ensemble learning as a pathway to more accurate climate and weather forecasting.

Overall, the ensemble's improved performance underscores the value of model diversity and adaptive weighting strategies for weather prediction in Nigeria. These results contribute evidence that data-driven forecasting, when enhanced by ensemble methods, can support regional climate resilience and policy planning

5. CONCLUSION AND FUTURE WORK

This study addressed the challenge of localized weather forecasting in Nigeria by developing optimized deep learning models tailored to the Kogi region. Using a 10-year dataset (2014–2023) from NiMET, three models, ANN, RNN and LSTM, were implemented and optimized with Bayesian Optimization. While the RNN showed slightly stronger standalone performance, none of the individual models consistently captured regional weather patterns. The proposed Boosting Ensemble model outperformed all single models, achieving the lowest RMSE and demonstrating robust predictive accuracy.

The findings confirm that ensemble learning, when applied to region-specific meteorological data, significantly enhances forecasting performance compared to individual deep learning architectures. This contribution is particularly relevant for Nigerian agriculture, energy management, and policy planning, where accurate monthly forecasts are critical for decision-making.

Future research can expand on this work by incorporating additional climatic variables (e.g., rainfall, humidity, wind speed), extending the approach to multiple ecological zones, and exploring higher-frequency forecasts (weekly or daily). Furthermore, integrating attention-based architectures such as Transformers may improve temporal pattern recognition and boost predictive accuracy.

6. REFERENCES

[1] Abdalla, A. M., Ghaith, I. H., & Tamimi, A. A. (2021, July). Deep learning weather forecasting techniques: literature survey. In 2021 International Conference on Information Technology (ICIT) (pp. 622-626). IEEE.

[2] Almazroui, M., Saeed, F., Saeed, S., Nazrul Islam, M., Ismail, M., Klutse, N. A. B., & Siddiqui, M. H. (2020). Projected change in temperature and precipitation over Africa from CMIP6. Earth Systems and Environment, 4, 455-475.

[3] Barrera-Animas, A. Y., Oyedele, L. O., Bilal, M., Akinosho, T. D., Delgado, J. M. D., & Akanbi, L. A. (2022). Rainfall prediction: A comparative analysis of modern machine learning algorithms for time-series forecasting. Machine Learning with Applications, 7, 100204.

[4] Bauer, P., Thorpe, A., & Brunet, G. (2015). The quiet revolution of numerical weather prediction. Nature, 525(7567), 47-55.

[5] Brotzge, J. A., Berchoff, D., Carlis, D. L., Carr, F. H., Carr, R. H., Gerth, J. J., ... & Wang, X. (2023). Challenges and opportunities in numerical weather prediction. Bulletin of the American Meteorological Society, 104(3), E698-E705.

[6] Chattopadhyay, A., Nabizadeh, E., & Hassanzadeh, P. (2020). Analog forecasting of extreme-causing weather patterns using deep learning. Journal of Advances in Modeling Earth Systems, 12(2), e2019MS001958.

[7] Egbunu, C. O., Ogedengbe, M. T., Yange, T. S., Rufai, M.

A., & Muhammed, H. I. (2021). Towards food security: the prediction of climatic factors in Nigeria using random forest approach. *Journal of Computer Science and Information Technology*, 70-81.

[8] Godwin, J. A., Singh, S., & Kumar, R. (2024). Prediction of Rainfall Using Data Mining Techniques: Evidence from Nigeria. SSRN, 4829865.

[9] Grönquist, P., Yao, C., Ben-Nun, T., Dryden, N., Dueben, P., Li, S., & Hoefler, T. (2021). Deep learning for post-processing ensemble weather forecasts. Philosophical Transactions of the Royal Society A, 379(2194), 20200092.

[10] Guo, Q., He, Z., & Wang, Z. (2024). Monthly climate prediction using deep convolutional neural networks and long short-term memory. Scientific Reports, 14(1), 17748. https://doi.org/10.1038/s41598-024-68906-6

[11] Han, Y., Mi, L., Shen, L., Cai, C. S., Liu, Y., Li, K., & Xu, G. (2022). A short-term wind speed prediction method utilizing novel hybrid deep learning algorithms to correct numerical weather forecasting. Applied Energy, 312, 118777.

[12] Hewage, P., Behera, A., Trovati, M., & Pereira, E. (2019, May). Long-short term memory for an effective short-term weather forecasting model using surface weather data. In IFIP International Conference on Artificial Intelligence Applications and Innovations (pp. 382-390). Cham: Springer International Publishing.

[13] Holmstrom, M., Liu, D., & Vo, C. (2016). Machine learning applied to weather forecasting. Meteorological Applications, 10(1), 1-5.

[14] Karevan, Z., & Suykens, J. A. (2020). Transductive LSTM for time-series prediction: An application to weather forecasting. Neural Networks, 125, 1-9.

[15] Meenal, R. M. P. A., Michael, P. A., Pamela, D., & Rajasekaran, E. (2021). Weather prediction using random forest machine learning model. Indonesian Journal of Electrical Engineering and Computer Science, 22(2), 1208-1215.

[16] Ojo, O. S., & Ogunjo, S. T. (2022). Machine learning models for prediction of rainfall over Nigeria. Scientific African, 16, e01246.

[17] Salman, A. G., Kanigoro, B., & Heryadi, Y. (2015, October). Weather forecasting using deep learning techniques. In 2015 International Conference on Advanced Computer Science and Information Systems (ICACSIS) (pp. 281-285). IEEE.

[18] Schultz, M. G., Betancourt, C., Gong, B., Kleinert, F., Langguth, M., Leufen, L. H., ... & Stadtler, S. (2021). Can deep learning beat numerical weather prediction? Philosophical Transactions of the Royal Society A, 379(2194), 20200097.

[19] Singh, S., Kaushik, M., Gupta, A., & Malviya, A. K. (2019, March). Weather forecasting using machine learning techniques. In Proceedings of 2nd International Conference on Advanced Computing and Software Engineering (ICACSE).

[20] Uluocak, I., & Bilgili, M. (2023). Daily air temperature forecasting using LSTM-CNN and GRU-CNN models. Acta Geophysica, 1-20.

[21] Vourlidas, A. (2021). Improving the Medium-Term Forecasting of Space Weather: A Big Picture Review From a Solar Observer's Perspective. Frontiers in Astronomy and Space Sciences, 8, 651527.

[22] Wang, Y., Shi, L., Hu, Y., Hu, X., Song, W., & Wang, L. (2023). A comprehensive study of deep learning for soil moisture prediction. Hydrology and Earth System Sciences Discussions, 2023, 1-38.

[23] Xu, S., Zhang, Y., Chen, J., & Zhang, Y. (2025). Short-to medium-term weather forecast skill of the AI-based Pangu-Weather model using automatic weather stations in China. Remote Sensing, 17(2), 1.