American Journal of Artificial Intelligence
2025; 9(1): 46-54
http://www.sciencepublishinggroup.com/j/ajai
doi: 10.11648/j.ajai.20250901.15
ISSN: 2639-9717 (Print); ISSN: 2639-9733 (Online)

Comparative Analysis of ARIMA and LSTM Models for Temperature Forecasting in Semi-Arid Regions: A Case Study of Machakos County, Kenya

Noah Mutavi Mutiso*, Martin Mutwiri Kithinji, Victor Muthama Musau

Department of Mathematics, Division of Statistics and Data Science, Kirinyaga University, Kerugoya, Kenya

Email address:

mutisonoah@gmail.com (Noah Mutavi Mutiso)
*Corresponding author

To cite this article:

Noah Mutavi Mutiso, Martin Mutwiri Kithinji, Victor Muthama Musau. (2025). Comparative Analysis of ARIMA and LSTM Models for Temperature Forecasting in Semi-Arid Regions: A Case Study of Machakos County, Kenya. *American Journal of Artificial Intelligence, 9*(1), 46-54. https://doi.org/10.11648/j.ajai.20250901.15

Received: 28 April 2025; **Accepted:** 16 May 2025; **Published:** 3 June 2025

Abstract: Accurate temperature forecasting is vital for agriculture, disaster management, and climate resilience, particularly in semi-arid regions like Machakos County, Kenya. Traditional forecasting models, such as the Auto Regressive Integrated Moving Average (ARIMA), have been widely used for both short- and long-term temperature predictions. However, with advancements in machine learning, there is a growing need to evaluate how modern methods compare to traditional approaches. This research focused on comparing the predictive accuracy of ARIMA and the Long Short-Term Memory (LSTM) model, a deep learning algorithm that uses a gating mechanism to capture non-linear patterns in data. The study used Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), and the Diebold-Mariano statistical test to evaluate and compare the performance of models. Temperature data was obtained from the National Aeronautics and Space Administration's Prediction of Worldwide Energy Resource (NASA POWER) database, using GPS coordinates to retrieve location-specific data for Machakos County. To ensure robust analysis, Python libraries such as Statsmodels, Pandas, TensorFlow, NumPy, Keras, and Matplotlib were used for data processing, fitting of models, and visualization of results. Findings showed that LSTM performed better in long-term predictions, achieving higher accuracy for 30-day forecasts, while both models performed significantly equivalently in the short term of 7 days. These findings highlight the complementary strengths of traditional and deep learning models in addressing different forecasting needs.

Keywords: Temperature Forecasting, Long Short-Term Memory (LSTM), Auto-regressive Integrated Moving Average (ARIMA), Machine Learning, Deep Learning

1. Introduction

Accurate temperature forecasting is essential for effective planning in sectors such as agriculture, water resource management, energy distribution, and disaster preparedness. Reliable predictions enable stakeholders to mitigate climate-related risks, optimize resource allocation, and enhance resilience to extreme weather conditions. Traditional statistical models, particularly the Autoregressive Integrated Moving Average (ARIMA), have been widely used for time series forecasting due to their ability to model linear dependencies in weather data. However, ARIMA struggles to capture the nonlinear patterns, seasonality, and long-term dependencies present in meteorological variables, limiting its effectiveness for extended forecasts.

Recent advances in deep learning have introduced more sophisticated approaches to time series forecasting, particularly Long Short-Term Memory (LSTM) networks, a variant of Recurrent Neural Networks (RNNs). LSTM models are designed to learn complex temporal relationships and retain long-term dependencies, making them particularly suited for weather prediction tasks. Studies have shown that LSTMs outperform traditional statistical models in capturing intricate temperature variations, yet their effectiveness compared to ARIMA remains underexplored, particularly in modeling temperature in semi-arid regions such as Machakos County, Kenya.

Machakos County experiences high temperature variability and seasonal fluctuations that significantly impact agricultural productivity and water availability. The region's reliance on rainfed agriculture makes it vulnerable to extreme temperature deviations, necessitating accurate and reliable forecasting models. However, challenges persist due to limited meteorological station coverage by the Kenya Meteorological Department (KMD), resulting in spatial gaps in observed data. In such cases, satellite-based climate datasets, such as the NASA POWER project, offer a viable alternative for obtaining high-resolution historical weather records.

This study conducted a comparative analysis of ARIMA and LSTM models for daily temperature forecasting in Machakos County. Using historical temperature data from the National Aeronautics and Space Administration - Prediction Of Worldwide Energy Resources (NASA POWER) project, the study aims to evaluate the predictive performance of both models based on the statistical metrics Mean Absolute Error(MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error(MAPE) and statistical comparison using the Diebold-Mariano (DM) test. By determining which model provides higher accuracy and robustness in forecasting temperature variations, this research contributes to the growing field of machine learning applications in climate science and provides valuable insights for policymakers, meteorologists, and agricultural planners seeking to enhance climate adaptation strategies.

2. Literature Review

2.1. Traditional Statistical Models in Temperature Forecasting

The Autoregressive Integrated Moving Average. (ARIMA) model has been widely used in temperature prediction due to its ability to model linear dependencies in time series data. Box and Jenkins [1] introduced ARIMA as a statistical approach for sequential data analysis, demonstrating its effectiveness in short-term temperature forecasting. Several studies have explored the application of ARIMA in meteorology, highlighting its strengths and limitations. Zhao and Magoules [2] utilized ARIMA to forecast ambient temperature for air conditioning systems, showing improved energy efficiency through accurate short-term predictions.

Aye and Karaman [3] conducted a comparative study between ARIMA and ARIMAX (ARIMA with exogenous variables) for daily temperature forecasting in Istanbul, Turkey. Their results indicated that ARIMAX performed better by incorporating additional weather variables such as humidity and wind speed. Similarly, [4] found that Seasonal ARIMA (SARIMA) provided better forecasts than standard ARIMA when applied to temperature data from different regions of Spain.

However, multiple studies have pointed out ARIMA's limitations, particularly its inability to model nonlinear dependencies in temperature patterns. Bianco et al. (2009) [5] and Maia et al. (2009) [6] demonstrated that while ARIMA models can capture seasonality, they struggle with long-term temperature trends and rapid fluctuations, which are crucial for climate change analysis.

2.2. Emergence of Deep Learning Models for Time Series Forecasting

To address ARIMA's limitations, researchers have increasingly turned to deep learning techniques, particularly LSTM networks. LSTM models, introduced by Hochreiter & Schmidhuber (1997) [7], are designed to handle long-range dependencies in time series data, making them suitable for weather forecasting.

Recent studies have demonstrated LSTM's superior performance in temperature forecasting compared to ARIMA. Sagheer & Kotb (2019) [8] found that LSTM models outperformed ARIMA in predicting extreme heat conditions in Upper Egypt, particularly for longer forecast windows. Similarly, Sezer et al. (2022) [9] showed that LSTMs achieved lower prediction errors than traditional models in Turkey's urban and rural areas, especially for five-day forecasts.

Sajo et al. (2021) [10] compared VAR and LSTM models for predicting canopy conductance in cocoa trees. The VAR model showed poor performance with low predictive accuracy. In contrast, the LSTM achieved a high accuracy of 97.4% and a low RMSE of 0.026. This highlights LSTM's superior ability to model complex temporal patterns in plant physiology.

Tugal & Sevgin (2023) [11] conducted a study in Mus Province, Turkey, comparing ARIMA, PROPHET, and LSTM using 12 years of historical temperature data. Their findings indicated that LSTM models consistently outperformed ARIMA and PROPHET in all evaluation metrics, particularly for long-term forecasts. Likewise, Siami-Namini et al. (2018) [12] demonstrated an 85% improvement in prediction accuracy using LSTM over ARIMA when applied to financial time series, highlighting LSTM's broader applicability in sequential data forecasting.

Further research by Sumanta & Shymapada (2022) [13] found out that ARIMA's performance remained relatively stable regardless of dataset size, whereas LSTM required larger datasets (>2000 samples) to achieve optimal accuracy. This aligns with findings from Qiu et al. (2021) [14], who applied LSTM for river water temperature prediction and found that deep learning methods provided more accurate forecasts than traditional statistical techniques.

2.3. Comparative Studies on ARIMA vs. LSTM for Temperature Forecasting

Multiple comparative studies have evaluated ARIMA vs. LSTM for temperature prediction. Salman et al. (2018) [15] benchmarked LSTMs against ARIMA and PROPHET for daily temperature predictions in Iraq, showing that LSTM models were more robust during rapid temperature fluctuations. Similarly, Zhang et al. (2019) [16] found that LSTM's error rate was 20% lower than ARIMA's for one-day forecasts and 35% lower for seven-day forecasts, reinforcing its advantage in long-term predictions.

In a large-scale study, Elsner & Schmertmann (2020) [17] compared LSTM with 16 statistical and machine learning models for daily temperature forecasting across 210 U.S. cities. LSTM consistently outperformed all models, particularly for longer forecast windows.

A recent study by Karevan & Suykens (2020) [18] applied LSTM, Support Vector Machines (SVMs), and Gaussian Processes to Belgian temperature data, concluding that while all models degraded in accuracy over time, LSTM degraded the least, making it the most reliable for extended forecasts.

2.4. Research Gap

While ARIMA and LSTM models have been widely studied, their application to temperature forecasting in Machakos County remains largely unexplored, particularly using NASA POWER data. Semi-arid regions like Machakos present distinct climatic challenges that necessitate a deeper examination of model performance. This research introduces a novel comparative analysis of ARIMA and LSTM models, specifically tailored to the region's unique temperature dynamics. By leveraging high-resolution large historical climate data from NASA POWER (a requirement in fitting deep learning models), this study provides new insights into the effectiveness of LSTM over ARIMA in capturing local temperature variations, advancing the understanding of machine learning applications in semi-arid climate forecasting.

3. Materials and Methods

3.1. Model Fitting and Evaluation Workflow

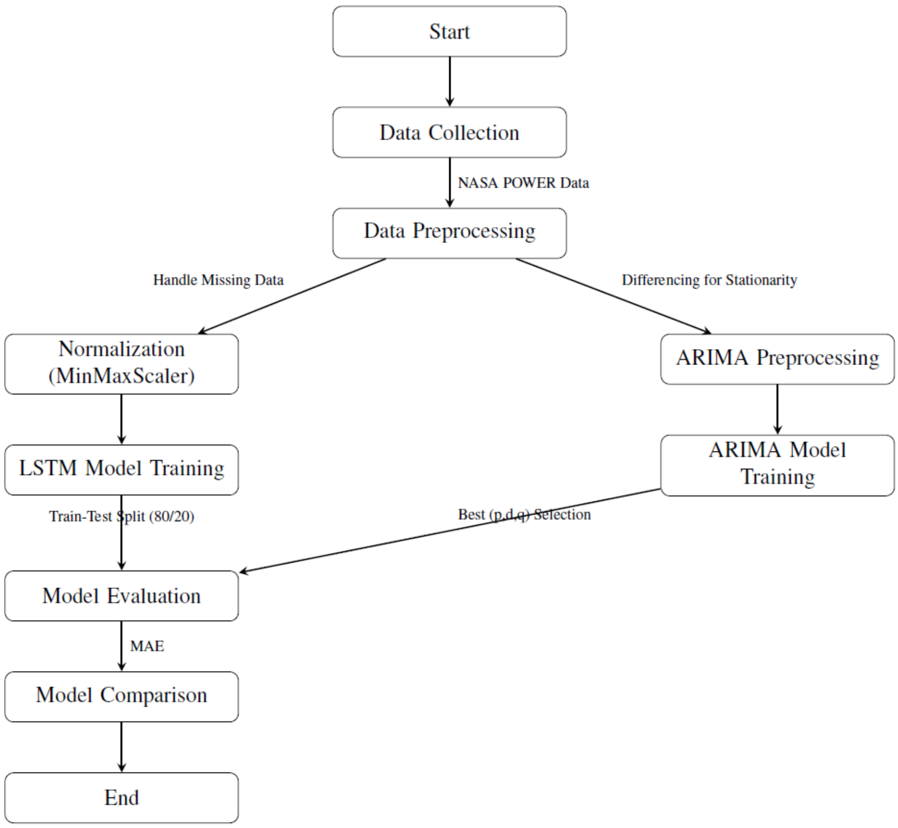

Figure 1. *Temperature Forecasting Methodology Workflow.*

3.2. Data Collection

The research utilized secondary data obtained through an Application Programming Interface (API), specifically from the NASA POWER project. The dataset includes daily recorded weather data, comprising of Temperature (primary focus variable),Precipitation, humidity and frost Point The dataset spans 12 years, from June 1, 2012, to May 31, 2024, consisting of 4,383 rows and 4 columns. The data was retrieved using latitude and longitude coordinates approximately at -1.518122 and 37.36823, corresponding to

Machakos County.

NASA POWER has been instrumental in providing global satellite-based atmospheric observation data. The Surface Meteorological and Solar Energy (SSE) climatological resource database, originally developed for renewable energy applications, has been expanded under the NASA POWER project to support various sectors, including agriculture and climate research. Studies such as Monteiro et al. (2018) [19] have demonstrated minimal root mean square error (RMSE) values when comparing NASA POWER data with ground-based weather station measurements, validating its reliability.

3.3. Model Development

3.3.1. LSTM Model

Long Short-Term Memory (LSTM) is a deep learning architecture designed to effectively model sequential data by capturing long-term dependencies, patterns and correlations that persist over extended time intervals, where earlier data points significantly influence current and future outcomes. The model has the following components.

Forget gate

Equation (1) was used to carry out forget operations in LSTM.

Input gate

Equation (2) was used to calculate input gate in long short-term memory cell. Input gate oversees the flow of temperature data into the cell state C_t

Candidate cell state

Candidate cell state computed using (3) and was used to update cell state in each timestamp.

Cell state

The cell C_t computed using (4) represents the memory of the LSTM cell.

Output gate

Output gate shown as (5) determines terminal temperature output of LSTM cell at every timestep. It controls how much information from the cell state should be utilized in computation of hidden state h_t

Hidden state

Hidden state h_t shown by (6) is the output of a LSTM cell at each timestep. This output is used as an input in next cell sequentially or model's final output.

LSTM Mathematical equations

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad \text{Forget Gate} \quad (1)$$

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad \text{Input Gate} \quad (2)$$

$$\tilde{C}_t = \tanh(W_C[h_{t-1}, x_t] + b_C) \quad \text{Cell State Proposal} \quad (3)$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \quad \text{Cell State Update} \quad (4)$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad \text{Output Gate} \quad (5)$$

$$h_t = o_t \cdot \tanh(C_t) \quad \text{Hidden State Update} \quad (6)$$

3.4. ARIMA Model

The ARIMA model was optimized using grid search to select the best (p, d, q) parameters based on the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). The model was tested with differencing to handle non-stationarity.

ARIMA Equation

$$y'_t = c+\phi_1 y'_{t-1}+\cdots+\phi_p y'_{t-p}+\theta_1\varepsilon_{t-1}+\cdots+\theta_q\varepsilon_{t-q}+\varepsilon_t \quad (7)$$

Where:

y'_t differenced value at time t

c constant term

$\phi_1, \phi_2, \ldots, \phi_p$ - autoregressive parameters

$y'_{t-1}, y'_{t-2}, \ldots, y'_{t-p}$ - lag p differenced values

$\theta_1, \theta_2, \ldots, \theta_q$ - moving average parameters

$\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots, \varepsilon_{t-q}$ are the previous q error terms, ε_t -error term at time t, (white noise).

3.5. Evaluation Metrics

The model evaluation and comparison was done using Mean Absolute Error (MAE) (8), Root Mean Absolute Error(RMSE) (9), Mean Absolute Percentage Error(MAPE) (10) and statistical Diebold Mariano(DM) test(11).

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad (8)$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (9)$$

$$\text{MAPE} = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \quad (10)$$

$$DM = \frac{\bar{d}}{\sqrt{\frac{1}{n}\gamma_0 + \frac{2}{n}\sum_{k=1}^{h-1}\gamma_k}} \quad (11)$$

where

$\bar{d}$ is the mean of the loss differentials

γ_k is the k-th order autocovariance of d_t,

n is the sample size,

h is the forecast horizon,

$g(\cdot)$ is a loss function,

y_i is the actual (observed) value,

$\hat{y}_i$ is the predicted (forecasted) value,

n is the number of observations.

4. Results

4.1. Model Training

To train the model, the LSTM model expects training data to be in a three-dimensional shape consisting of [number of samples, number of days to look back, number of features], in this case, the shape of the data is transformed from (4383,1) to (4323, 60, 1) since the model forecasts the temperature values in degrees Celsius for the next 7 and 30 days by looking back at the past 60 days.

The LSTM model architecture consisted of four stacked LSTM layers with 50, 50, 50, and 48 units respectively. To mitigate overfitting, dropout regularization was applied after each layer, with dropout rates of 0.3 for the first three LSTM layers and 0.2 after the final layer. The model concluded with a fully connected (Dense) output layer for predicting the next temperature value. The model was compiled using the Adam optimizer with a learning rate of 0.01 and trained over 45 epochs with a batch size of 64. The loss function used was Mean Squared Error (MSE), and model performance was evaluated on a separate validation set to monitor generalization capability. Dropout regularization method was used to prevent overfitting of the LSTM model during training [20] and maintain model robustness Validation loss was used to track overfitting.

4.2. Model Evaluation

After training the model, both training and validation losses were plotted to evaluate performance. Loss values reflect how well the model fits the data, with values closer to zero indicating a better fit. The validation set, which consists of data not used during training, helps assess the model's generalization ability. In this case, the training loss was 0.0048 and the validation loss was 0.0045, both calculated using the Mean Squared Error (MSE) metric. The low training loss suggests that the model fits the training data well, while the slightly lower validation loss indicates strong forecasting performance and minimal signs of overfitting. The figure 2 shows training loss vs. validation loss.

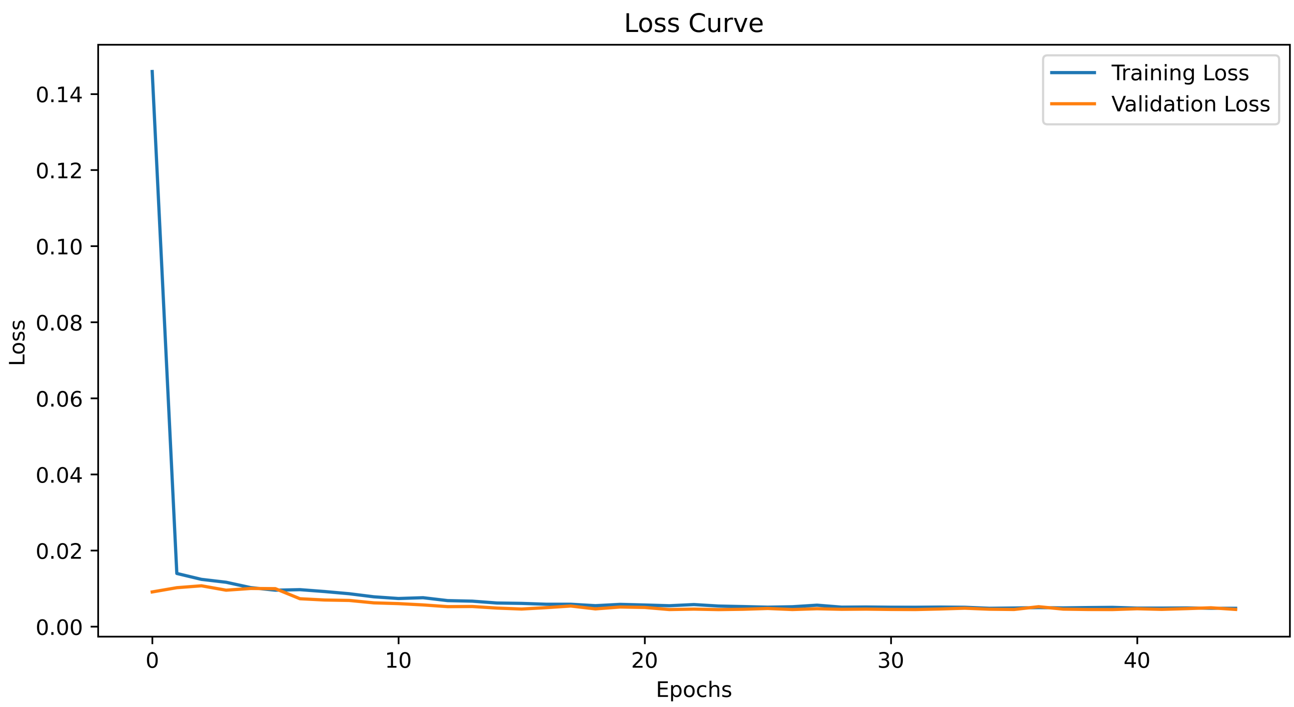

Figure 2. *Validation Loss vs. Training Loss.*

4.3. ARIMA Model

To assess the stationarity of the temperature time series data, a combination of Auto-correlation Function (ACF) fig: 4 and Partial Auto-correlation Function (PACF) fig: 3 plots were analyzed, along with the Augmented Dickey-Fuller (ADF) statistical test. These diagnostics, supported by results from the ndiffs() function, confirmed that the series was trend-stationary, meaning no differencing was required (d = 0).

Once stationarity was achieved, the ACF and PACF plots were further examined to identify potential orders for the autoregressive (AR) component (p) and the Moving Average (MA) component (q) of the model. Based on these visual diagnostics, several candidate ARIMA (p,0,q) models were identified for further evaluation.

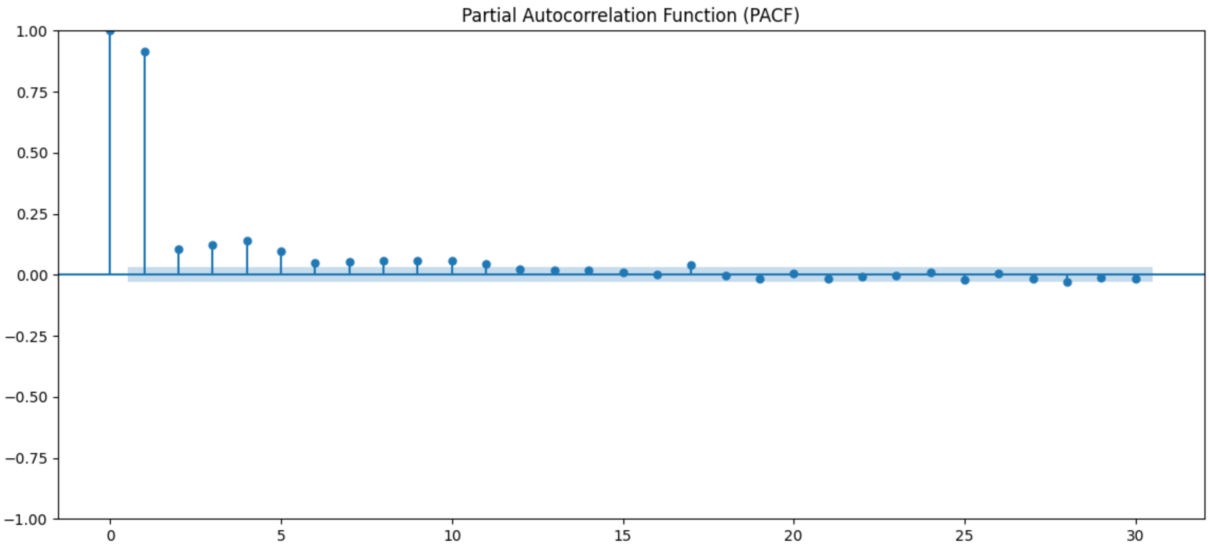

Figure 3. *Partial Auto-Correlation Function (PACF) Plot.*

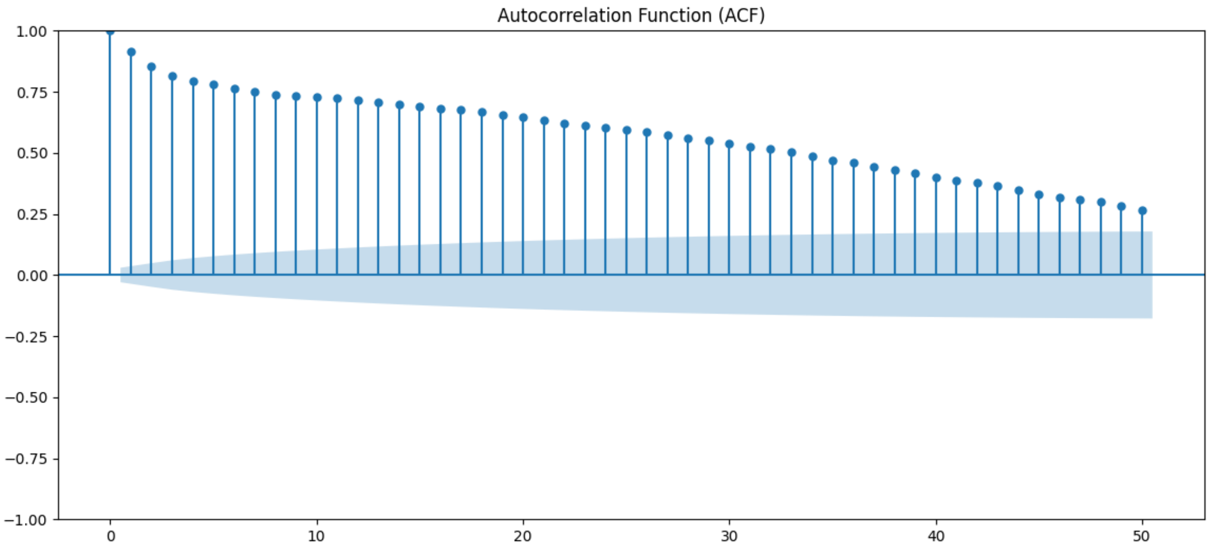

Figure 4. *Auto-Correlation Function (ACF) Plot.*

To determine the optimal model, each candidate was fitted and compared using two key criteria: the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). The AIC measures the overall goodness of fit, while the BIC applies an additional penalty for model complexity to encourage selection of the most parsimonious model.

Table 1. *Candidate ARIMA Models and Their Evaluation Criteria.*

p	q	AIC	BIC
1	3	7824.24	7862.55
3	3	7825.23	7876.31
4	1	7825.73	7870.42
3	1	7826.02	7864.33
2	3	7826.13	7870.83

Among these models, ARIMA(1,0,3) was selected as the optimal model lowes AIC and BIC.

The selected ARIMA(1,0,3) model was then used to generate seven-day and thirty-day temperature forecasts. The accuracy of these forecasts was evaluated using performance metric Mean Squared Error(MAE) .

4.4. Comparison of LSTM and ARIMA Models

Table 2 and 3 provides a comparative analysis of the forecasting performance of ARIMA and LSTM models over two different time periods: 7 days and 30 days. The evaluation is based on MAE,RMSE, MAPE and Diebold-Mariano which measure the accuracy of the models in predicting temperature. Figures 5 6, illustrate the comparison of the performance between the LSTM and ARIMA model.

Table 2. *Forecast Performance Metrics for ARIMA and LSTM Models.*

Model	Forecast Horizon	MAE °C	RMSE °C	MAPE (%)
ARIMA	7 Days	0.5756	0.6602	2.7
LSTM	7 Days	0.5577	0.6288	2.60
ARIMA	30 Days	1.1460	1.2646	6.7
LSTM	30 Days	0.9169	1.0170	4.70

Table 3. *Diebold-Mariano Test Results for Forecast Accuracy Comparison.*

Forecast Horizon	DM Statistic	p-value
7 Days	1.0489	0.3346
30 Days	6.7129	0.0000

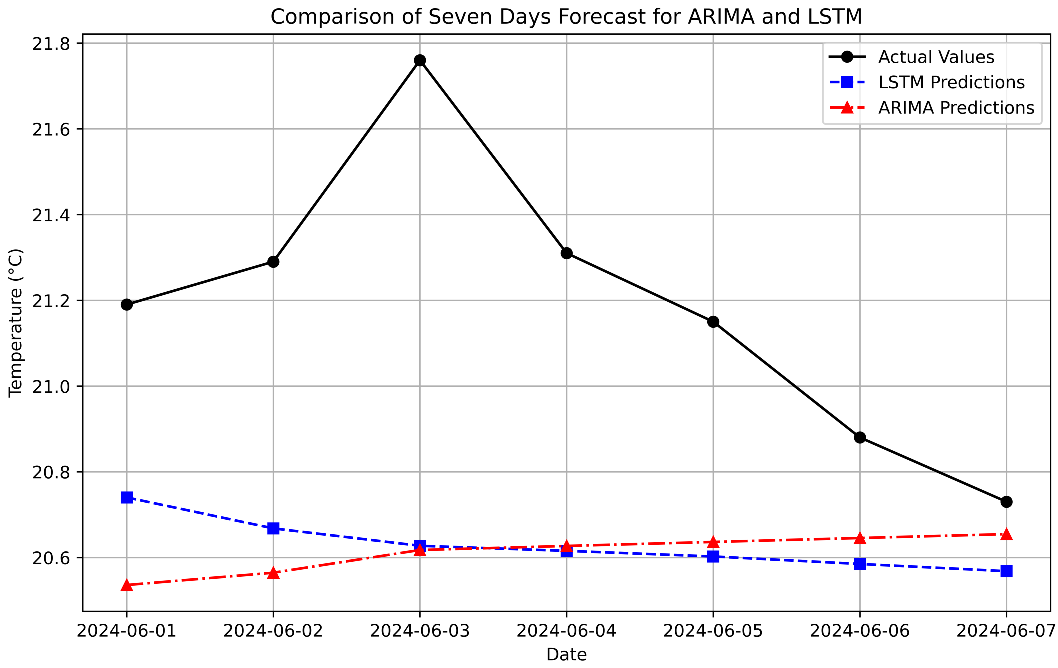

Figure 5. *Seven-Day ARIMA and LSTM Predictions.*

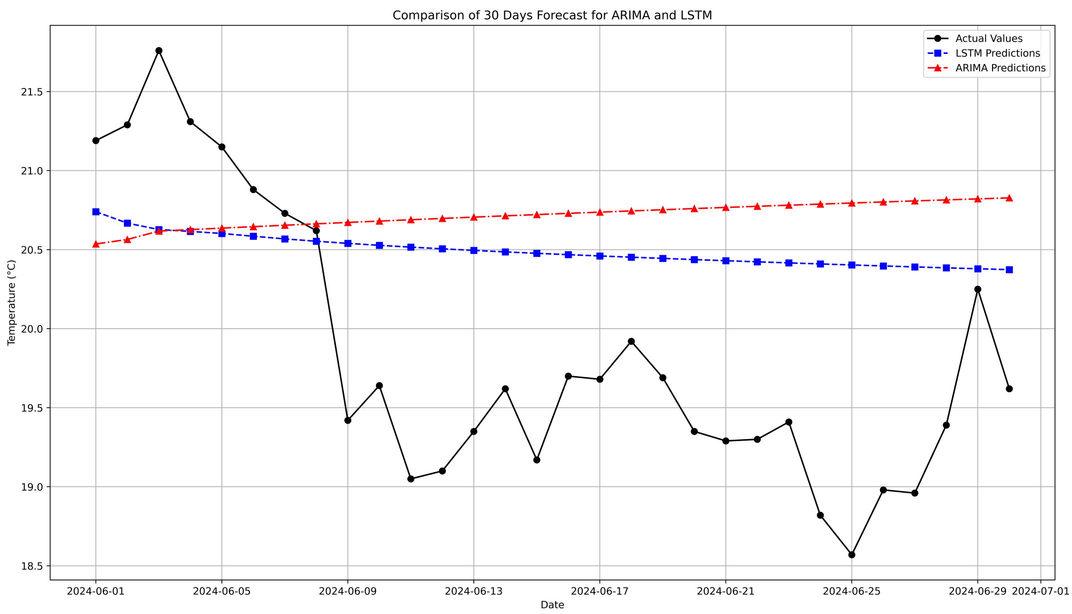

Figure 6. *Thirty-Day ARIMA and LSTM Predictions.*

5. Discussion

The evaluation of ARIMA and LSTM models was carried out using forecast performance metrics which are Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE) for 7-day and 30-day forecast horizons. As highlighted in table 2, the LSTM model performed better than ARIMA model in all evaluated metrics across both time horizons.

In the 7 day forecast, LSTM produced a MAE of 0.5577°C, an RMSE of 0.6288°C, and a MAPE of 2.60%, all slightly lower than ARIMA's corresponding values (MAE: 0.5756°C, RMSE: 0.6602°C, MAPE: 2.70%). These slight differences reflect the LSTM model's capacity to capture short-term patterns more accurately than ARIMA. The performance gap becomes more pronounced in 30 day forecast time-frame. LSTM recorded a MAE of 0.9169°C and RMSE of 1.0170°C, while ARIMA showed increased error values (MAE: 1.1460°C, RMSE: 1.2646°C). In addition, LSTM attained a lower MAPE of 4.70%, compared to 6.70% for ARIMA, indicating improved relative accuracy over extended forecasts.

To assess whether these differences are statistically significant, Diebold-Mariano (DM) test was used, as shown in Table 3. For the 7-day forecast, the DM statistic was 1.0489 with a p-value of 0.3346, indicating that the observed difference in forecast accuracy is not statistically significant at $\alpha = 0.05$ confidence level. In contrast, the 30-day forecast produced a DM statistic of 6.7129 with a p-value of 0.0000, providing strong evidence to reject the null hypothesis of equal predictive accuracy. This result confirms that the LSTM model significantly outperforms ARIMA in long-term temperature forecasting but performs equivalently in short term forecast.

In summary, while both models deliver acceptable short-term forecasts, the LSTM model exhibits superior performance in both absolute error metrics and statistical robustness over longer horizons. These findings validate the suitability of deep learning models such as LSTM for time series applications that require both precision and scalability across different forecast lengths.

6. Conclusion

This study compared the performance of ARIMA and LSTM models for temperature forecasting over short- and long-term horizons. The evaluation was based on standard forecast error metrics and statistical significance testing using the Diebold-Mariano framework. While both models demonstrated reasonable accuracy in short-term forecasts, LSTM consistently performed better compared to ARIMA, especially as the forecast time frame extended.

The statistical analysis further confirmed that the difference in predictive performance became significant over longer time-frames, reinforcing the LSTM model's strength in capturing nonlinear patterns and long-range dependencies inherent in temperature data. These findings highlight the practical advantage of deep learning approaches like LSTM for applications where forecast reliability over extended periods is critical.

Future work may consider hybrid models or ensemble techniques to further enhance predictive performance, as well as the integration of exogenous variables such as humidity, wind speed, or seasonal indicators to improve model robustnes

Abbreviations

ARIMA	Autoregressive Integrated Moving Average
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
RMSE	Root Mean Squared Error
MAPE	Mean Absolute Percentage Error
DM	Diebold-Mariano
NASA POWER	National Aeronautics and Space Administration‘s Prediction of Worldwide Energy Resources
RNN	Recurrent Neural Network
ADF	Augmented Dickey-Fuller (Test)
API	Application Programming Interface
AIC	Akaike Information Criterion
BIC	Bayesian Information Criterion
PACF	Partial Auto-Correlation Function
ACF	Auto-Correlation Function
KMD	Kenya Meteorological Department
SSE	Surface Meteorological and Solar Energy
SARIMA	Seasonal Autoregressive Integrated Moving Average
VAR	Vector Autoregression
SVM	Support Vector Machine
MSE	Mean Squared Error
DNN	Deep Neural Network
AR	AutoRegressive
MA	Moving Average

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] G. E. P. Box and G. M. Jenkins, *Time Series Analysis: Forecasting and Control*. San Francisco, CA: Holden-Day, 1970.

[2] H.-X. Zhao and F. Magoules, "A review on the prediction of building energy consumption," *Renew. Sustain. Energy Rev.*, vol. 16, no. 6, pp. 3586–3592, 2012, https://doi.org/10.1016/j.rser.2012.02.049

[3] S. A. Aye and O. Karaman, "Forecasting daily temperature using ARIMA and ARIMAX models," in *Proc. 2018 Int. Conf. Artif. Intell. Data Process. (IDAP)*, Malatya, Turkey, 2018, pp. 1–5.

[4] O. Claveria and S. Torra, "Forecasting tourism demand to Catalonia: Neural networks vs. time series models," *Econ. Model.*, vol. 36, pp. 220–228, 2014, https://doi.org/10.1016/j.econmod.2013.09.024

[5] V. Bianco, O. Manca, and S. Nardini, "Electricity consumption forecasting in Italy using linear regression models," *Energy*, vol. 34, no. 9, pp. 1413–1421, 2009, https://doi.org/10.1016/j.energy.2009.06.005

[6] A. L. Maia, F. L. de Oliveira, and F. G. da Silva, "Forecasting models for ambient temperature of unshaded environments," *Energy Buildings*, vol. 41, no. 8, pp. 874–878, 2009, https://doi.org/10.1016/j.enbuild.2009.03.007

[7] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997, https://doi.org/10.1162/neco.1997.9.8.1735

[8] A. Sagheer and M. Kotb, "Time series forecasting of petroleum production using deep LSTM recurrent networks," *Neurocomputing*, vol. 323, pp. 203–213, 2019, https://doi.org/10.1016/j.neucom.2018.09.075

[9] O. B. Sezer, V. M. Sezer, V. S. Gürısık, A. A. Önder, Ö. Meydanoğlu, and M. R. Şenyüz, "Bagging LSTM ensemble model for multi-day temperature forecasting," *Environmental Science and Pollution Research*, vol. 29, no. 24, pp. 36264–36278, 2022, https://doi.org/10.1007/s11356-022-18637-z

[10] O. S. Sajo *et al.*, "Modelling the canopy conductance of cocoa tree using a recurrent neural network," *Am. J. Neural Netw. Appl.*, vol. 7, no. 2, pp. 22–27, 2021.

[11] İ. Tuğal and F. Sevgin, "Analysis and forecasting of temperature using time series forecasting methods: A case study of Muş," *Thermal Science*, vol. 27, pp. 3081–3088, 2023, https://doi.org/10.2298/TSCI2304081T

[12] S. Siami-Namini, N. Tavakoli, and A. S. Namin, "A comparison of ARIMA and LSTM in forecasting time series," in *Proc. 17th IEEE Int. Conf. Mach. Learn. Appl. (ICMLA)*, Orlando, FL, USA, 2018, pp. 1394–1401, https://doi.org/10.1109/ICMLA.2018.00227

[13] D. Sumanta and M. Shymapada, "A comparative study of seasonal-ARIMA and RNN (LSTM) on time series temperature data forecasting," in *Pervasive Comput. Social Netw.*, Singapore: Springer, 2022, pp. 263–273, https://doi.org/10.1007/978-981-16-5640-8_25

[14] R. Qiu *et al.*, "River water temperature forecasting using a deep learning method," *J. Hydrol.*, vol. 595, p. 126016, Apr. 2021, https://doi.org/10.1016/j.jhydrol.2021.126016

[15] A. G. Salman, Y. Heryadi, E. Abdurahman, and W. Suparta, "Single-layer and multi-layer LSTM model for weather forecasting," in *Proc. 2018 Int. Semin. Res. Inf. Technol. Intell. Syst. (ISRITI)*, Yogyakarta, Indonesia, 2018, pp. 112–117, https://doi.org/10.1109/ISRITI.2018.8864437

[16] J. Zhang, Y. Zheng, D. Qi, R. Li, and X. Yi, "DNN-based prediction model for spatio-temporal data: A traffic data case," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 3, no. 2, pp. 1–24, 2019, https://doi.org/10.1145/3351261

[17] J. B. Elsner and C. P. Schmertmann, "Assessing forecasts on a global scale from the lens of local behavior: A conceptual discussion with insights from daily temperature forecasts over the United States," *Weather Forecast.*, vol. 35, no. 4, pp. 1539–1555, 2020, https://doi.org/10.1175/WAF-D-19-0245.1

[18] Z. Karevan and J. A. Suykens, "Transductive multiview learning for multi-output regression," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 12, pp. 5423–5437, 2021, https://doi.org/10.1109/TNNLS.2020.3042273

[19] L. A. Monteiro, P. C. Sentelhas, and G. U. Pedra, "Assessment of NASA/POWER satellite-based weather system for Brazilian conditions and its impact on sugarcane yield simulation," *Int. J. Climatol.*, vol. 38, no. 3, pp. 1571–1581, 2018, https://doi.org/10.1002/joc.5249

[20] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1929–1958, 2014.

[21] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv: 1412.6980*, 2014.