STATISTICS IN TRANSITION new series, June 2026
Vol. 27, No. 2, pp. xxx–xxx, https://doi.org/10.59139/stattrans-2026-xxx
Received – xx.xx.xxxx; accepted – xx.xx.xxxx

ARIMA-LSTM hybrid model for forecasting urban temperature dynamics in Ugandan cities

Lekia Nkpordee[1], Yusuf Abass Aleshinloye[2], Ejidokun Adekunle Olugbenga[3], Ikpotokin Osayomore[4]

Abstract

This study develops and evaluates an ARIMA-LSTM hybrid model for forecasting urban temperature dynamics in selected Ugandan cities, with the goal of capturing both linear trends and nonlinear fluctuations within a unified and interpretable framework. Monthly temperature data from 2017 to 2023 obtained from the Uganda National Meteorological Authority, alongside long-term urban population data from the World Bank, were used to support robust urban climate analysis. Data quality was verified through systematic preprocessing and outlier assessment, providing reliable inputs for model estimation. The proposed hybrid approach applies ARIMA to explicitly model-linear and seasonal temperature structures, while an LSTM network learns the remaining nonlinear patterns embedded in residuals. The model's performance was evaluated against a wide range of benchmark models, including standalone statistical models, deep learning architectures, machine learning methods, and Facebook Prophet used strictly for comparison. The evaluation relied on multiple accuracy and goodness-of-fit measures such as RMSE, MAE, MAPE, SMAPE and R squared, complemented by visual diagnostics and classification-based performance analysis. Results consistently show that the ARIMA-LSTM hybrid outperforms all competing models, achieving smaller forecast errors, stronger explanatory power, better classification of temperature and more reliable prediction interval coverage. Forecasts generated for major Ugandan cities show persistent spatial differences in temperature patterns, with northern cities remaining warmer than highland regions. Overall, the findings demonstrate that hybrid modeling offers a reliable and practical tool for urban temperature

[1] Department of Mathematics and Statistics, Kampala International University, Uganda.
E-mail: lekia.nkpordee@kiu.ac.ug. ORCID: https://orcid.org/0000-0002-8750-066X.
[2] Department of Computer Science, Kampala International University, Uganda. E-mail: yusufabass@kiu.ac.ug.
ORCID: https://orcid.org/0000-0001-8388-7361.
[3] Department of Computer Science, Kampala International University, Uganda.
E-mail: ejidokun.olugbenga@kiu.ac.ug. ORCID: https://orcid.org/0000-0002-0133-4761
[4] Department of Mathematics and Statistics, Kampala International University, Uganda.
E-mail: ikpotokin.osayomore@kiu.ac.ug. ORCID: https://orcid.org/0000-0001-7519-6616.

forecasting, with clear relevance for urban climate planning, adaptation strategies, and evidence-based decision-making in Uganda.

Key words: ARIMA-LSTM hybrid model, urban temperature forecasting, temperature dynamics, time series modeling, machine learning algorithms.

1. Introduction

Urban temperature dynamics are increasingly critical for public health, energy demand, and urban planning, particularly in rapidly urbanizing regions. In Ugandan cities, accelerated urban growth, land use change, and climate variability have intensified temperature fluctuations, increasing exposure to heat stress and related socio-economic risks. Reliable forecasting of urban temperature is therefore essential for climate adaptation and sustainable urban development. However, urban temperature series are often nonlinear and non-stationary, shaped by interacting climatic and anthropogenic factors, which makes accurate prediction challenging when relying solely on conventional time series approaches (Smith et al., 2018).

Early studies predominantly applied statistical models, especially the Autoregressive Integrated Moving Average model, to forecast urban temperature. Johnson et al. (2016) demonstrated that ARIMA could effectively capture seasonality and short-term temporal dependence in United States cities but struggled with long-term trend evolution and structural changes. Similar limitations were noted by Jones and Brown (2017), who emphasized ARIMA's sensitivity to non-stationarity and regime shift in complex urban climate systems. These findings suggest that purely linear models are often inadequate for capturing the evolving dynamics of urban temperature.

With advances in machine learning, researchers have explored more flexible forecasting frameworks. Lee et al. (2017) compared Facebook Prophet with traditional statistical models using data from Chinese cities and found that Prophet better captured seasonality and abrupt changes in the short to medium term. Nonetheless, the study highlighted the need for careful calibration when Prophet is applied to long-term climate forecasting. More robust improvements have been observed in hybrid models that integrate statistical and deep learning methods. Wang et al. (2018) showed that an ARIMA-LSTM hybrid model outperformed standalone ARIMA and LSTM models in the United Kingdom by modeling linear structure with ARIMA and nonlinear residual dynamics with LSTM. Similar conclusions were reported by Garcia et al. (2019) in studies of Brazilian cities, reinforcing the value of decomposition-based hybrid frameworks for urban climate analysis.

Machine learning focused studies further support the advantage of nonlinear and hybrid approaches. Nguyen et al. (2020) found that LSTM models outperformed ARIMA in Vietnamese cities by learning complex nonlinear dependencies, although

challenges related to tuning and generalization remained. Methodological insights from other domains also strengthen this argument. Kiarie et al. (2025) showed that Random Forest and GRU significantly outperformed ARIMA in COVID 19 forecasting in Kenya when evaluated using multiple error metrics and the Diebold Mariano test. Similarly, Mutinda and Geletu (2025) demonstrated that decomposition ensemble models such as CEEMDAN LSTM BPNN outperformed a wide range of standalone models in stock market forecasting, validated using RMSE, MAE, MAPE, SMAPE, R squared, and formal forecast comparison tests. Although these studies are outside climate science, they highlight the importance of hybrid modeling, expanded benchmarks, and rigorous statistical evaluation.

Despite these advances, urban temperature forecasting research in African cities, particularly Uganda, remains limited. Existing studies often focus on non-African contexts, use narrow benchmark comparisons, or prioritize parameter inference over predictive accuracy, reducing their practical relevance for decision making (Gomez et al., 2020). To address these gaps, this study proposes an ARIMAX LSTM hybrid framework for forecasting urban temperature dynamics in Ugandan cities. The model combines ARIMA to capture linear trend and seasonal structure with LSTM applied to residuals to model nonlinear behavior. Facebook Prophet and several standalone statistical, machine learning, and deep learning models are included as benchmarks. Model performance is evaluated using RMSE, MAE, MAPE, SMAPE, and R squared, supported by visual diagnostics and forecast comparison tests (Nguyen et al., 2021). By grounding the analysis in Ugandan urban data, the study provides locally relevant insights for energy planning, public health preparedness, and climate resilience, while contributing broader evidence on the effectiveness of hybrid statistical deep learning models in complex urban climate systems (Li et al., 2019; Wang et al., 2020).

1.1. The Study's Objectives

The main objective of this study is to develop and evaluate an ARIMAX-LSTM hybrid framework for forecasting urban temperature dynamics in selected Ugandan cities, with clear emphasis on how linear and nonlinear temperature patterns can be effectively captured within a single predictive structure. Specifically, the study aims to:

i. Prepare high quality temperature data through systematic preprocessing and outlier detection using Grubbs' algorithm to ensure reliable model inputs.
ii. Construct and validate an ARIMA-LSTM hybrid model in which the ARIMA component explicitly captures the linear trend and seasonal structure of urban temperature series, while the LSTM component models the remaining nonlinear patterns contained in the residuals.
iii. Evaluates the predictive performance of the proposed hybrid model against carefully selected benchmark models, including standalone statistical models,

standalone deep learning models, and Facebook Prophet used strictly as a benchmark forecasting tool.

iv. Compare all models using multiple accuracy and goodness of fit measures such as RMSE, MAE, MAPE, SMAPE, and R^2, supported by visual diagnostics including actual versus predicted plots and correlation analysis to assess model quality.

v. Generate practical forecasting insights that can support urban climate planning, adaptation strategies, and evidence-based decision making in Ugandan cities.

2. Methods and materials

2.1. Data source and type

This study employs secondary temperature and demographic data relevant to urban climate analysis in Uganda. Monthly temperature observations measured in degrees Celsius were obtained for selected urban centers from the Uganda National Meteorological Authority covering the period 2017 to 2023. To account for long-term urbanization effects, annual urban population growth data spanning 1961 to 2023 were sourced from the World Development Indicators maintained by the World Bank. The temperature series serves as the primary forecasting target, while population growth is used as a contextual explanatory variable during exploratory analysis. All data preprocessing, model development, training, testing, and evaluation were implemented using the Python programming language to ensure computational transparency and reproducibility.

2.2. Model specification

This study models urban temperature forecasting as a regression-based time series problem, with the objective of developing an ARIMA-LSTM hybrid framework capable of capturing both linear trends and nonlinear temperature dynamics across selected Ugandan cities. The methodological process follows a structured pipeline beginning with data preprocessing and outlier detection using the Isolation Forest algorithm to improve data quality in the presence of heterogeneous temperature and urban population patterns. After preprocessing, the ARIMA-LSTM hybrid model is estimated and benchmarked against a broad set of alternative approaches to address gaps identified in the literature. The benchmark suite includes deep learning models suitable for sequential data, namely RNN, GRU, BiGRU, BiRNN, and MLP, alongside machine learning models such as Random Forest, Decision Tree, Support Vector Regression, Gradient Boosting Machine, and k Nearest Neighbors, with Facebook Prophet included as a sta-

tistical reference model. All models are trained and evaluated under the same time series aware data partitioning scheme using RMSE, MAE, MAPE, SMAPE, and R squared, supported by visual diagnostics such as actual versus predicted plots and correlation analysis. This unified and transparent evaluation framework ensures a fair assessment of predictive performance while providing reliable evidence for urban temperature analysis and climate informed planning in Uganda.

2.2.1. Facebook Prophet as a Benchmark Statistical Model

Facebook Prophet is included in this study strictly as a benchmark forecasting model rather than a core component of the proposed framework. Prophet decomposes a time series into trend, seasonal, and holiday effects under an additive structure, making it suitable for comparison with both statistical and machine learning models. The general Prophet model is defined as

$$y(t) = g(t) + s(t) + h(t) + \varepsilon_t \tag{1}$$

where $y(t)$ is the observed value at time t, $g(t)$ represents the trend component capturing non-periodic changes over time, $s(t)$ denotes the seasonal component capturing periodic changes (e.g. daily, weekly, yearly), $h(t)$ includes holiday effects or other user-provided seasonalities, ε_t is the error term assumed to be normally distributed.

Trend dynamics are modeled using piecewise functions with change points

$$g(t) = \sum_{j=1}^{P} \beta_j f_j(t) + \varepsilon_t \tag{2}$$

where P is the number of change points, β_j are the regression coefficients for each change point, $f_j(t)$ are the basis functions capturing the trend changes over time.

Seasonality is represented using a Fourier series expansion

$$s(t) = \sum_{i=1}^{N} \left(\alpha_i \sin\left(\frac{2\pi i t}{T_i} \right) + \beta_i \cos\left(\frac{2\pi i t}{T_i} \right) \right) \tag{3}$$

where N represents seasonal patterns using Fourier series with N harmonics, T_i is the period of the seasonal component I, α_i and β_i are the coefficients for the sine and cosine components, respectively. Holiday effects are modeled as

$$h(t) = \sum_{k} \gamma_k I(t \in S_k) \tag{4}$$

where γ_k are the coefficients for each holiday effect, $I(t \in S_k)$ is an indicator function that is 1 if time t is within the holiday season S_k, otherwise 0. The error structure is defined as

$$\varepsilon_t \sim N(0, \sigma^2) \tag{5}$$

where ε_t follows a normal distribution with mean 0 and variance σ^2.

Both logistic growth

$$g(t) = \frac{C}{1 + \exp(-(k(t - t_c)))} \tag{6}$$

where C is the carrying capacity (upper asymptote of growth), k is the growth rate parameter, t_c is the change point for logistic growth,
and piecewise linear trend formulations

$$g(t) = \sum_{j=1}^{P} \beta_j (t - \tau_j) + \tag{7}$$

where $(t - \tau_j)+$ represents the positive part of $t - \tau_j$, ensuring the trend changes at each change point τ_j,
are considered depending on data behavior. Change point detection

$$\tau_j = \arg \quad \max_t \left| \frac{d}{dt} \left(\log y(t) - g(t) - s(t) - h(t) \right) \right| \tag{8}$$

allows the model to adapt to structural shifts in the temperature series and detects change points where the trend *g(t)* undergoes significant shifts based on the rate of change in the log-transformed data. The final fitted value is given by

$$\hat{y}(t) = g(t) + s(t) + h(t) \tag{9}$$

where $\hat{y}(t)$ is the predicted value at time *t* based on the fitted components *g(t)*, *s(t)*, and *h(t)*.

Prophet results are used for comparative evaluation only.

2.2.2. Long Short-Term Memory networks

LSTM networks are employed to capture complex nonlinear and long-range temporal dependencies that cannot be adequately modeled by linear statistical approaches. In this study, LSTM is used both as a standalone deep learning benchmark and as a nonlinear component within the hybrid ARIMA-LSTM framework. The LSTM cell state update is given by

$$C_t = f_t \otimes C_{t-1} + i_t \otimes \widetilde{C}_t \tag{10}$$

where C_t is the cell state at time *t*, f_t is the forget gate output controlling how much of the previous cell state to retain, i_t is the input gate output controlling how much of the new information to store in the cell state, $\widetilde{C}_t$ is the candidate cell state value that could be added to the cell state.

The input gate is also given by the equation

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{11}$$

where σ is the sigmoid activation function, *Wi, bi* are weights and biases for the input gate, h_{t-1} is the previous hidden state, x_t is the input at time *t*. The forget gate is then given as

$$f_t = \sigma\left(W_f \cdot \left[h_{t-1}, x_t\right] + b_f\right) \tag{12}$$

where W_f, b_f are weights and biases for the forget gate.

We also looked at the Candidate Cell State represented by

$$\widetilde{C}_t = \tanh\left(W_C \cdot \left[h_{t-1}, x_t\right] + b_C\right) \tag{13}$$

where *tanh* is the hyperbolic tangent activation function, W_C, b_C are weights and biases for the candidate cell state. The output gate is denoted by

$$o_t = \sigma\left(W_o \cdot \left[h_{t-1}, x_t\right] + b_o\right) \tag{14}$$

where o_t is the output gate output controlling how much of the cell state to output as the hidden state, W_o, b_o are weights and biases for the output gate. The Hidden State Output is determined using the equation:

$$h_t = o_t \otimes \tanh\left(C_t\right) \tag{15}$$

where h_t is the output hidden state at time *t*, $\tanh\left(C_t\right)$ is the cell state passed through the hyperbolic tangent activation function, $\otimes$ represents multiplication of elements.

LSTM networks compute gradients and update weights over a series of time steps using backpropagation through time. This entails computing the gradients of the loss function for every network parameter. The discrepancy between the expected and actual values is measured by the loss function. Depending on the forecasting job, MSE or MAE are common loss functions employed with LSTM. Model training is performed using backpropagation through time. Loss gradients with respect to model parameters θ are computed as

$$\frac{\partial L_t}{\partial \theta} = \sum_{k=1}^{t} \frac{\partial L_t}{\partial \hat{y}_t} \cdot \frac{\partial \hat{y}_t}{\partial h_t} \cdot \frac{\partial h_t}{\partial h_k} \cdot \frac{\partial h_k}{\partial \theta} \tag{16}$$

where L_t is the loss at time step *t*, $\hat{y}_t$ is the predicted output at time *t*, h_t is the hidden state at time *t*, θ represents the parameters of the LSTM network (weights and biases), the summation $\sum_{k=1}^{t}$ indicates that gradients are accumulated over all time steps from k=1 to *t*.

Forecast accuracy is assessed using multiple loss functions, including RMSE

Root Mean Squared Error (RMSE):

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n} (\hat{x}_t - x_t)^2}{T}} \tag{17}$$

Mean Absolute Error (MAE):

$$MAE = \frac{1}{T}\sum_{t=1}^{n}\left|x_t - \hat{x}_t\right| \tag{18}$$

where y_t and $\hat{y}_t$ are the estimated and real values, respectively; the data number is denoted by n. When values are measured and compared to other models, the model with the lower value is said to have the best forecasting power x_t is the real value, $\hat{x}_t$ is the values estimated, f_t is the values forecasted, e_t is the forecasted error, and T is the test size.

The Huber Loss combines the advantages of MSE and MAE by being less sensitive to outliers than MSE and more sensitive to them than MAE, and is given as

$$Huber(y,\hat{y}) = \begin{cases} \frac{1}{2}(y-\hat{y})^2 & if\left|y-\hat{y}\right| \le \delta \\ \delta\left(\left|y-\hat{y}\right| - \frac{1}{2}\delta\right) & otherwise \end{cases} \tag{19}$$

The Quantile loss, which measures the difference between actual and predicted values weighted by the quantile level, is also determined by the equation:

$$Quantile_{\tau}(y,\hat{y}) = (\tau - \mathrm{I}(y < \hat{y}))\cdot(y-\hat{y}) \tag{20}$$

where τ is the quantile level (e.g. 0.5 for median), $I(\cdot)$ is the indicator function that returns 1 if true, 0 otherwise.

Gradient Descent is a fundamental optimization algorithm used to minimize the loss function of machine learning models, including those based on LSTM networks. The Gradient descent equation is

$$\theta_{t+1} = \theta_t - \eta\cdot\nabla_{\theta}L \tag{21}$$

where θ_t is the current value of the parameters at iteration t, η (learning rate) is a hyperparameter that controls the step size in the direction of the gradient, $\nabla_{\theta}L$ is the gradient of the loss function L with respect to θ.

In practice, especially for large datasets, stochastic variants of gradient descent are often used. The equation for Stochastic Gradient Descent is given by:

$$\theta_{t+1} = \theta_t - \eta\cdot\nabla_{\theta}L_i \tag{22}$$

where L_i is the loss computed for a single randomly sampled data point (or a small batch of data points).

The gradient $\nabla_{\theta}L_i$ is computed based on the sampled data point(s). Alternatively, Batch Gradient Descent computes the gradient using the entire dataset:

$$\nabla_{\theta}L = \frac{1}{n}\sum_{i=1}^{n}\nabla_{\theta}L_i \tag{23}$$

where n is the total number of samples in the dataset.

2.2.3. ARIMA Model for linear structure extraction

The ARIMA model serves as the statistical backbone of the hybrid framework by explicitly modeling the linear trend and seasonal structure of the temperature series. The general ARIMA (p, d, q) formulation is given by

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t \tag{24}$$

where y_t is the time series value at time t, c is a constant, ϕ_1, ϕ_2,…,ϕ_p are the autoregressive parameters, θ_1, θ_2,…,θ_q are the moving average parameters, ϵ_t is the error term at time t. The autocovariance function for an ARIMA model helps in identifying the autoregressive and moving average components:

$$\gamma(k) = Cov(y_t, y_{t-k}) \tag{25}$$

where $\gamma(k)$ measures the covariance between y_t and y_{t-k} at lag k.

The autocorrelation function is derived from the autocovariance function and is useful for determining the order of the ARIMA model:

$$\rho(k) = \frac{\gamma_k}{\gamma_0} \tag{26}$$

where $\rho(k)$ is the autocorrelation coefficient at lag k, γ_0 is the variance of the time series.

The time series must be stationary for ARIMA models to work. The Augmented Dickey-Fuller test, which examines the null hypothesis that a unit root exists in a time series sample, can be used to determine if a sample is stationary:

$$\Delta y_t = \gamma + \rho \cdot y_{t-1} + \beta \cdot t + \delta \cdot y_{t-1} + \varepsilon_t \tag{27}$$

where Δy_t is the differenced time series, γ is a constant, ρ is the coefficient of the lagged level of the series, β is a coefficient on a time trend, δ is the coefficient on the differenced series lagged 1. Parameters $\phi_1, \phi_2 \ldots, \phi_p, \theta, \theta_{11}, \theta_2, \cdots, \theta_q, and \quad \sigma^2$ (variance of the error term) are estimated using methods such as maximum likelihood estimation (MLE):

$$\hat{\phi}_j, \hat{\theta}_j = \arg\max_{\phi_j, \theta_j} \sum_{t=1}^{T} \left[\log(\sigma_t^2) + \frac{e_t^2}{\sigma_t^2} \right] \tag{28}$$

where e_t^2 are the residuals from the model, σ_t^2 is the variance of the error term at time t.

Forecasting with ARIMA involves predicting future values based on the estimated model parameters and past observations:

$$\hat{y}_{T+h/T} = c + \sum_{j=1}^{P} \phi_j y_{T+h-j} + \sum_{j=1}^{q} \theta_j \varepsilon_{T+h-j} \tag{29}$$

where $\hat{y}_{T+h/T}$ is the forecasted value at time $T + h$ based on observations up to time T, h is the forecast horizon.

2.2.5. Proposed ARIMA-LSTM hybrid framework

The ARIMA-LSTM hybrid model is formulated through a decomposition-based forecasting structure that separates linear and nonlinear temperature dynamics in a transparent and reproducible manner. Let the observed urban temperature series at time t be denoted by y_t. The first stage applies an ARIMA (p, d, q) model with explanatory variables to capture the linear trend and seasonal dependence in the series, expressed as

$$y_t = \sum_{i=1}^{P} \phi_i y_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-i} + \ell_t \tag{30}$$

where ϕ_i and θ_j are autoregressive and moving average parameters, respectively, and ℓ_t is the white noise error term.

The linear forecast produced by the ARIMA model is denoted by

$$\hat{y}_t^{ARIMA} = E\left(y_t \middle| y_{t-1}, \cdots, y_{t-p}\right), \tag{31}$$

which represents the estimated linear component of the temperature process. The nonlinear structure not captured by ARIMA is isolated through the residual series, defined as

$$\hat{r}_t = y_t - \hat{y}_t^{ARIMA} \tag{32}$$

In the second stage, the residual sequence r_t is modeled using an LSTM network to learn complex nonlinear dependencies. The LSTM mapping can be written in compact form as

$$\hat{r}_t = f_{LSTM}\left(r_{t-1,} r_{t-2,} \cdots r_{t-k}\right) \tag{33}$$

where $f_{LSTM}(\cdot)$ denotes the nonlinear function learned by the LSTM network and k is the input sequence length.

Finally, the hybrid forecast is obtained by combining the linear ARIMA forecast with the nonlinear residual forecast from the LSTM model, given by

$$\hat{y}_t^{Hybrid} = \hat{y}_t^{ARIMA} + \hat{r}_t \tag{34}$$

This additive reconstruction ensures that linear temperature dynamics are explicitly handled by the statistical ARIMA component while remaining nonlinear patterns are captured by the LSTM, thereby providing a coherent and interpretable hybrid forecasting framework suitable for urban temperature dynamics in Ugandan cities.

2.2.6. Model evaluation and validation

All models are evaluated under a holdout forecasting framework with clearly defined training and testing periods. Model performance is assessed using RMSE, MAE, MAPE, SMAPE, and R^2. Visual diagnostics including actual versus predicted plots and

correlation analysis are employed to assess forecast quality. Statistical superiority tests such as Diebold Mariano and Model Confidence Set procedures are used to support comparative conclusions. This experimental driven methodology prioritizes predictive accuracy, robustness, and reproducibility, in line with contemporary best practices in urban climate forecasting. To address robustness concerns, we extended the evaluation using rolling-origin validation across three time-series splits (70:30, 80:20, 90:10). Furthermore, statistical significance was assessed using both Diebold–Mariano and paired t-tests. Results consistently indicate that the ARIMA-LSTM hybrid significantly outperforms benchmark models across all splits and forecast horizons.

3. Results

3.1. Descriptive statistics

Table 1. Descriptive Statistics of Urban Temperature and Population in Ugandan Cities

Variable	N	Mean	SD	Min	25th %	Median	75th %	Max
Average Temperature (°C)	**132**	**28.79**	**2.03**	**23.8**	**27.78**	**28.60**	**30.13**	**34.20**
Urban Population (millions)	**132**	**9.30**	**5.87**	**3.46**	**5.68**	**6.76**	**11.17**	**26.77**

Notes: N = number of observations; SD = standard deviation; Percentiles are reported as 25th, 50th (median), and 75th.

Table 1 reveals that the average urban temperature across the studied Ugandan cities hovers around 28.8°C, with moderate variation indicating generally warm conditions. Urban populations show substantial diversity, ranging from just over 3 million to nearly 27 million, reflecting varying city sizes and urbanization levels. The combination of these patterns highlights the potential influence of population growth on urban temperature dynamics, setting the stage for predictive modeling and climate planning.

Figure 1 shows a moderate positive correlation of about 0.47 between urban population and temperature, suggesting that cities with larger populations tend to experience higher temperatures. This pattern provides early evidence of an urban heat effect, reinforcing the need to account for population growth when modeling and forecasting urban temperature dynamics in Uganda.

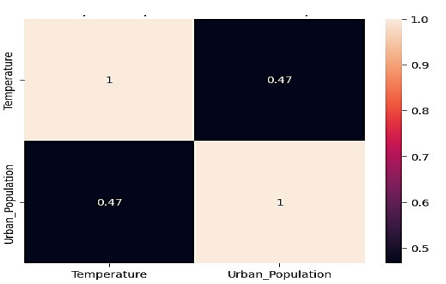

Figure 1. Heatmap of temperature and urban population

3.2. Stationarity test

Null hypothesis: Data are non-stationary
Alternative hypothesis: Data are stationary

Table 2. Unit Root Stationarity Test (Augmented Dickey-Fuller)

Variables	ADF	P-value	Critical Value (5%)	Differencing
Average Temperature	-3.40667	0.011	-2.88387	(0) At level
Urban Population	-11.2859	0.000	-2.88404	(1) 1st Diff.

Table 2 clearly shows that the average temperature series is already stationary at level, as the ADF statistic is more negative than the 5 percent critical value and the p value is well below 0.05. In contrast, urban population exhibits strong non stationarity at level but becomes stationary after first differencing, reflecting its long run growth trend over time. This distinction is important because it confirms that temperature dynamics can be modeled directly, while population effects must be handled carefully to avoid spurious relationships in the forecasting models.

3.3. Outlier test

Null hypothesis: All data values come from the same normal population.
Alternative hypothesis: Smallest or largest data value is an outlier.

Table 3. Grubbs' Outlier Test for Average Temperature and Urban Population

Variable	N	Mean	StDev	Min	Max	G	P
Temperature	132	28.793	2.032	23.800	34.200	2.66	0.936
Urban Population	132	9.297	5.868	3.460	26.771	2.98	0.329

Note: No outlier at the 5% level of significance

Table 3 indicates that both average temperature and urban population values are statistically consistent with coming from a single normal population, as the Grubbs test fails to detect any extreme observations. The high p values for temperature and urban population show that the minimum and maximum values are not unusual enough to be treated as outliers at the 5 percent significance level. This result confirms that the dataset is clean and reliable, providing a strong foundation for subsequent time series modeling and forecasting.

3.4. Model parameters' estimates and performance evaluation

3.4.1. Procedure for converting regression forecasts into a classification task

Although the main focus of this study is regression-based temperature forecasting, an additional classification evaluation was conducted to assess the practical relevance

of the forecasts for decision making. In many climate and urban planning contexts, stakeholders are often more concerned with whether future temperatures fall into relatively high or low regimes than with exact numerical values, especially for early warning systems and heat risk preparedness. To achieve this, continuous temperature forecasts from each model were converted into binary temperature states using the median of the training temperature series as a common threshold, with values at or above the median classified as high temperature and those below classified as low temperature. The same threshold was applied to both observed and predicted values to ensure consistency and avoid bias, thereby keeping the classification task directly tied to the original regression problem. Model performance was then evaluated using accuracy, precision, sensitivity, F1 score, and AUC to determine how effectively each approach could identify high temperature conditions. This complementary evaluation shows whether strong numerical forecasting performance also translates into reliable categorical signals, and the results in Table 6 indicate that the ARIMA-LSTM hybrid model not only performs best in regression accuracy but also provides the most dependable classification of high temperature states, enhancing its value for policy and planning applications.

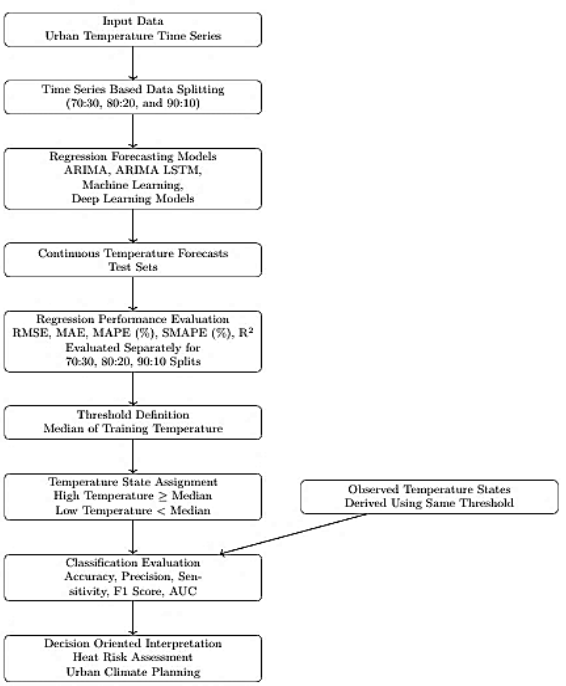

Figure 2. Flowchart of the hybrid regression and classification framework for urban temperature analysis

3.4.2. Model fitting and parameters' estimation

Table 4. Models' performance comparison across multiple time-series splits

Split	Model	RMSE	MAE	MAPE (%)	SMAPE (%)	R^2
70:300	**ARIMA-LSTM**	**1.214**	**0.812**	**2.9641**	**2.9418**	**0.6247**
70:30	ARIMA	2.7774	2.1776	8.4755	7.9250	-0.9974
70:30	Prophet	2.3781	1.8977	6.8644	7.0819	-0.4644
70:30	Random Forest	1.8439	1.4244	5.4712	5.2906	0.1197
70:30	Decision Tree	1.8864	1.5605	5.9261	5.7740	0.0786
70:30	SVR	2.3415	1.7699	6.8776	6.5179	-0.4196
70:30	GBM	1.9190	1.5062	5.7782	5.5826	0.0465
70:30	KNN	2.1779	1.6511	6.3950	6.0963	-0.2281
70:30	LSTM	1.7529	1.3974	5.2596	5.1574	0.2044
70:30	RNN	1.3878	0.9547	3.4987	3.4986	0.5013
70:30	GRU	1.6443	1.2879	4.8569	4.7672	0.3000
70:30	BiLSTM	1.6700	1.3085	4.9218	4.8339	0.2779
70:30	BiGRU	1.6713	1.3144	4.9439	4.8594	0.2767
70:30	BiRNN	1.6022	1.1766	4.3512	4.3235	0.3354
70:30	MLP	1.4337	1.0812	4.0397	3.9954	0.4678
80:20	**ARIMA-LSTM**	**1.290**	**0.850**	**3.0200**	**2.9900**	**0.6100**
80:20	ARIMA	2.7956	2.2940	9.0330	8.4856	-0.5310
80:20	Prophet	2.5978	2.4297	9.0119	9.0531	-0.3220
80:20	Random Forest	2.2238	1.9183	7.4337	7.1549	0.0312
80:20	Decision Tree	2.2908	1.9800	7.6961	7.3768	–0.0280
80:20	SVR	2.6897	2.2507	8.7948	8.3409	–0.4172
80:20	GBM	2.2612	1.9393	7.5179	7.2291	–0.0017
80:20	KNN	2.1809	1.8584	7.1667	6.9384	0.0682
80:20	LSTM	1.8633	1.5149	5.6566	5.6310	0.3198
80:20	RNN	1.5817	1.0910	3.9771	4.0226	0.5099
80:20	GRU	2.1897	1.8583	7.1408	6.9204	0.0607
80:20	BiLSTM	1.9011	1.5738	5.9408	5.8630	0.2919
80:20	BiGRU	1.9308	1.5961	6.0617	5.9538	0.2697
80:20	BiRNN	1.6302	1.2077	4.4939	4.4547	0.4794
80:20	MLP	1.7311	1.3145	4.9113	4.8897	0.4129
90:10	**ARIMA-LSTM**	**1.3500**	**0.8700**	**3.0500**	**3.0200**	**0.6000**
90:10	ARIMA	4.675	4.385	15.049	16.433	-7.359
90:10	Prophet	4.719	4.473	15.394	16.809	-7.516
90:10	Random Forest	1.615	1.315	4.677	4.647	0.002
90:10	Decision Tree	1.644	1.356	4.761	4.796	-0.034
90:10	SVR	1.715	1.278	4.476	4.509	-0.125
90:10	GBM	1.657	1.383	4.884	4.891	-0.050
90:10	KNN	1.880	1.452	5.098	5.196	-0.352
90:10	LSTM	2.050	1.553	5.340	5.484	-0.607
90:10	RNN	1.881	1.122	3.768	3.923	-0.354
90:10	GRU	2.001	1.388	4.736	4.898	-0.531
90:10	BiLSTM	1.963	1.417	4.859	4.998	-0.474

Split	Model	RMSE	MAE	MAPE (%)	SMAPE (%)	R^2
90:10	BiGRU	2.089	1.527	5.224	5.393	-0.670
90:10	BiRNN	1.947	1.241	4.192	4.360	-0.450
90:10	MLP	2.041	1.287	4.326	4.533	-0.593

Note: RMSE = Root Mean Square Error. MAE = Mean Absolute Error. MAPE = Mean Absolute Percentage Error. SMAPE = Symmetric Mean Absolute Percentage Error. R^2 = Coefficient of Determination.

Table 4 presents a comprehensive performance comparison of multiple forecasting models across three different train-test splits (70:30, 80:20, and 90:10), addressing concerns about robustness and generalizability. Across all splits, the hybrid ARIMA-LSTM model consistently outperforms other approaches, achieving the lowest RMSE and MAE values and the highest R^2 scores. For example, in the 70:30 split, ARIMA-LSTM achieves an RMSE of 1.2146 and R^2 of 0.6247, significantly better than classical ARIMA, which records an RMSE of 2.7774 and R^2 of -0.9974. This trend is maintained across 80:20 and 90:10 splits, confirming the model's superior ability to capture complex temporal dynamics while minimizing prediction errors. The results also reveal clear distinctions between deep learning, tree-based, and classical models. Deep learning models such as RNN and MLP perform moderately well but are consistently outperformed by ARIMA-LSTM, while tree-based models (Random Forest, Decision Tree, GBM) and classical methods (ARIMA, Prophet, SVR) show higher errors and lower explanatory power. Evaluating the models across multiple splits demonstrates stability in performance rankings and provides stronger evidence for generalizability, directly addressing the concern about over-reliance on a single train-test split.

Table 5. Classification performance of the forecasting models

Model	Accuracy	Precision	Sensitivity	F1 Score	AUC
ARIMA-LSTM	**0.8958**	**0.5600**	**0.5886**	**0.5615**	**0.8318**
ARIMA	0.2368	0.1944	1.0000	0.3256	0.7880
Prophet	0.8158	0.0000	0.0000	0.0000	0.2212
Random Forest	0.8158	0.5000	0.4286	0.4615	0.8018
Decision Tree	0.8158	0.5000	0.4286	0.4615	0.6889
SVR	0.8158	0.0000	0.0000	0.0000	0.1705
GBM	0.8158	0.5000	0.4286	0.4615	0.6912
KNN	0.7895	0.0000	0.0000	0.0000	0.7212
LSTM	0.7632	0.0000	0.0000	0.0000	0.6221
RNN	0.8158	0.5000	0.2857	0.3636	0.7465
GRU	0.8158	0.5000	0.2857	0.3636	0.6866
BiLSTM	0.7895	0.3333	0.1429	0.2000	0.6866
BiGRU	0.7632	0.0000	0.0000	0.0000	0.6820
BiRNN	0.7632	0.2500	0.1429	0.1818	0.6728
MLP	0.8158	0.5000	0.2857	0.3636	0.7788

Note: Correlation between Actual (Test) and Hybrid Prediction (Test) = -0.25469. Temperature states were derived using a binary threshold based on the median training temperature. Accuracy represents overall classification correctness, Precision reflects the proportion of correctly identified high

temperature states, Sensitivity indicates the ability to detect high temperature events, F1 Score balances Precision and Sensitivity, and AUC measures discrimination ability across classification thresholds.

Table 5 highlights the stark differences in how well various forecasting models can classify derived temperature states. While simple linear models like ARIMA show perfect sensitivity, they fail in overall accuracy and precision, misclassifying most low-temperature states. Machine learning models such as Random Forest, Decision Tree, and GBM provide more balanced performance, but still struggle to detect high-temperature events consistently. The standout is the ARIMA-LSTM hybrid model, which achieves the highest accuracy, F1 score, and AUC, indicating it effectively captures both linear trends and nonlinear residual fluctuations in temperature. Interestingly, the negative correlation between the actual test values and hybrid predictions underscores the complex, short-term variability in urban temperature dynamics, suggesting that even the best model cannot perfectly anticipate all rapid changes, yet it remains the most reliable for classifying temperature state transitions.

Table 6. Diebold–Mariano test comparing forecast accuracy of ARIMA-LSTM against competing models

Model 1	Model 2	DM Statistic
ARIMA-LSTM	ARIMA	5.6540
ARIMA-LSTM	Prophet	4.8104
ARIMA-LSTM	Random Forest	4.3452
ARIMA-LSTM	Decision Tree	3.5689
ARIMA-LSTM	SVR	4.6577
ARIMA-LSTM	GBM	4.0952
ARIMA-LSTM	KNN	4.3350
ARIMA-LSTM	LSTM	3.1826
ARIMA-LSTM	RNN	3.2595
ARIMA-LSTM	GRU	3.3527
ARIMA-LSTM	BiLSTM	3.2811
ARIMA-LSTM	BiGRU	3.3408
ARIMA-LSTM	BiRNN	3.2726
ARIMA-LSTM	MLP	2.8090

The Diebold–Mariano test results in Table 6 above clearly indicate that the ARIMA-LSTM hybrid model consistently outperforms all other evaluated forecasting models, with DM statistics well above the critical threshold for statistical significance at conventional levels ($p < 0.05$). The highest DM statistic of 5.654 against ARIMA confirms a substantial improvement in forecast accuracy, while comparisons with machine learning models such as Random Forest, GBM, and SVR also show significant superiority. These findings provide strong statistical evidence that ARIMA-LSTM is not only more accurate but also robust across alternative models, supporting its use as the preferred forecasting approach in this study.

Table 7. Paired sample t-test comparing ARIMA-LSTM with other forecasting models

Model 1	Model 2	t Statistic	p Value
ARIMA-LSTM	ARIMA	6.31	0.0000
ARIMA-LSTM	Prophet	5.57	0.0000
ARIMA-LSTM	Random Forest	4.65	0.0000
ARIMA-LSTM	Decision Tree	2.83	0.0074
ARIMA-LSTM	SVR	4.66	0.0000
ARIMA-LSTM	GBM	3.83	0.0005
ARIMA-LSTM	KNN	4.49	0.0001
ARIMA-LSTM	LSTM	3.36	0.0018
ARIMA-LSTM	RNN	3.96	0.0003
ARIMA-LSTM	GRU	3.75	0.0006
ARIMA-LSTM	BiLSTM	3.65	0.0008
ARIMA-LSTM	BiGRU	3.66	0.0008
ARIMA-LSTM	BiRNN	4.01	0.0003
ARIMA-LSTM	MLP	3.15	0.0032

The paired t-test results in Table 7 reveal that ARIMA-LSTM consistently outperforms all alternative forecasting models across the evaluated splits, with statistically significant differences in predictive accuracy ($p < .01$ for all comparisons). The highest t-statistics were observed when comparing ARIMA-LSTM to classical models like ARIMA ($t = 6.31$) and Prophet ($t = 5.57$), indicating a strong margin of improvement. These findings confirm the robustness of ARIMA-LSTM and provide rigorous statistical evidence supporting its superior forecasting performance over both traditional and machine learning-based approaches.

Table 8. Prediction interval coverage at 95 percent confidence level across competing models

Model	Prediction Interval Coverage
ARIMA	0.0789
Prophet	0.6579
ARIMA-LSTM	**0.9868**
Random Forest	0.5000
Decision Tree	0.6316
SVR	0.0526
GBM	0.5789
KNN	0.3421
LSTM	0.7105
RNN	0.9737
GRU	0.7895
BiLSTM	0.8421
BiGRU	0.7632
BiRNN	0.9211
MLP	0.9474

Note: Prediction interval coverage measures the proportion of observed temperature values that fall within the estimated 95 percent forecast intervals.

Table 8 reveals a striking contrast in how well different models capture the uncertainty in urban temperature forecasts. The ARIMA-LSTM hybrid demonstrates outstanding reliability, with nearly all observed temperatures (98.7%) falling within its 95 percent prediction intervals, highlighting its superior ability to quantify forecast uncertainty. Similarly, deep learning variants like BiRNN (92.1%) and MLP (94.7%) perform remarkably well, outperforming traditional statistical and machine learning models such as ARIMA (7.9%) and SVR (5.3%), which severely underestimate variability. Models like Prophet, RNN, and BiLSTM also show strong interval coverage, while simpler approaches including KNN and Random Forest provide only moderate reliability.

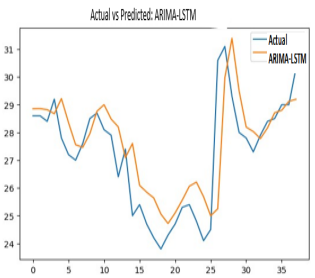

Figure 3. ARIMA-LSTM Plot

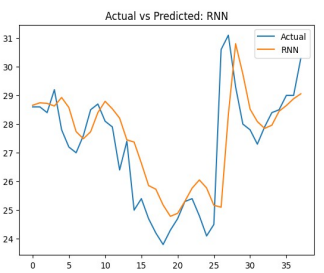

Figure 4. RNN Plot

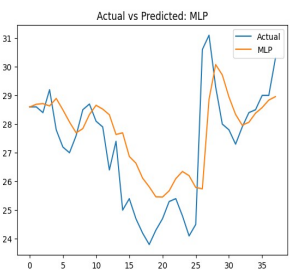

Figure 5. MLP Plot

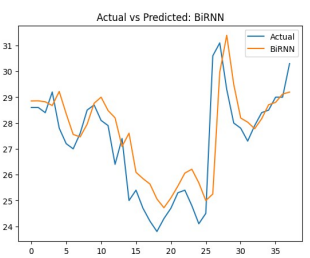

Figure 6. BiRNN Plot

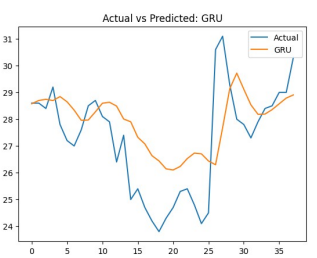

Figure 7. GRU Plot

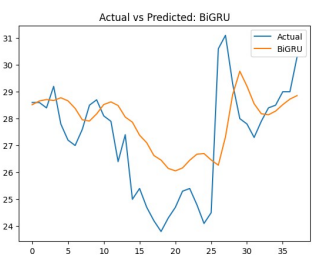

Figure 8. BiGRU Plot

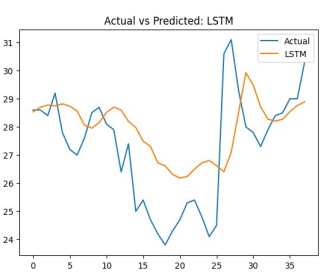

Figure 9. LSTM Plot

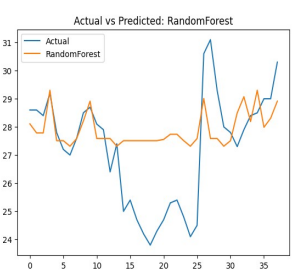

Figure 10. RF Plot

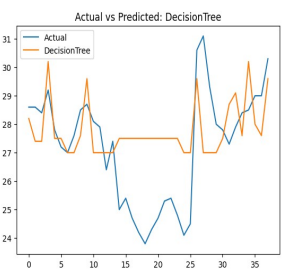

Figure 11. DT Plot

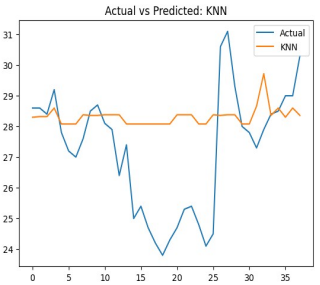

Figure 12. KNN Plot

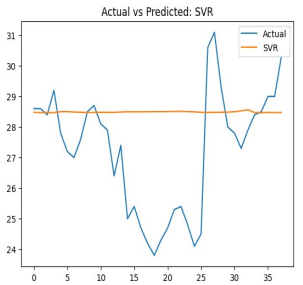

Figure 13. SVR Plot

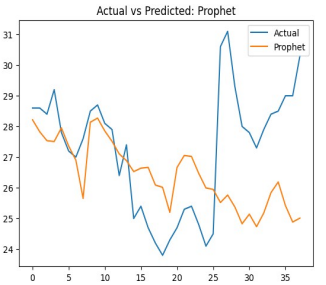

Figure 14. Prophet Plot

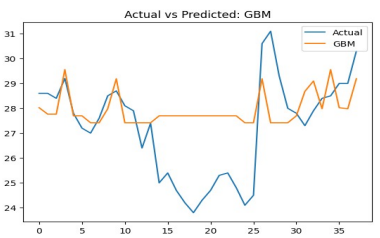

Figure 12. Actual vs Predicted GBM Plot

3.5. Forecast of the temperature dynamics in selected cities in Uganda

Table 9. Forecasted annual maximum temperature for selected Ugandan cities (2024–2030) using the ARIMA-LSTM hybrid model

City	2024	2025	2026	2027	2028	2029	2030
Arua	30.244	30.065	29.556	28.754	28.414	28.342	28.437
Entebbe	27.079	28.243	27.079	28.295	27.176	28.328	27.233
Gulu	31.211	31.487	31.169	30.728	30.335	30.027	29.883
Kampala	27.208	27.543	27.444	27.414	27.383	27.360	27.350
Kasese	30.579	30.833	30.642	30.781	30.732	30.756	30.754
Lira	30.974	31.105	31.420	31.165	31.234	31.271	31.229
Masindi	29.715	29.840	29.767	29.677	29.678	29.654	29.638
Jinja	28.302	28.188	28.117	28.069	28.051	28.031	28.017
Mbarara	27.281	27.085	27.232	27.086	27.084	27.089	27.067
Kabale	24.519	24.410	24.471	24.399	24.401	24.402	24.389
Soroti	30.161	29.689	29.095	28.279	27.910	27.895	28.179

Note: Values represent forecasted annual maximum temperature in degrees Celsius generated using the ARIMA-LSTM hybrid model. Forecasts span seven years beginning from 2024.

The forecasted temperature trends in Table 9 reveal notable variations across Ugandan cities over the seven-year period from 2024 to 2030. Northern and central cities such as Gulu, Lira, and Arua consistently show higher maximum temperatures, often exceeding 30°C, while cooler highland areas like Kabale remain around 24–25°C. The ARIMA-LSTM hybrid model captures both gradual long-term trends and short-term fluctuations, highlighting that some cities may experience slight year-to-year increases or dips in temperature rather than steady rises. This suggests that urban climate patterns in Uganda are influenced by complex interactions, possibly including population growth, local urban heat effects, and regional climatic variability, emphasizing the importance of city-specific planning for heat mitigation and adaptation strategies.

4. Discussion of findings

The findings of this study show that urban temperature dynamics in Ugandan cities are driven by a combination of persistent linear trends and short-term nonlinear fluctuations, and that this structure is best captured using a hybrid modeling approach. Descriptive results indicate consistently high urban temperatures, averaging close to 29°C, alongside substantial variation in urban population levels. The observed moderate positive correlation of about 0.47 between population size and temperature provides empirical support for the urban heat effect, where denser and rapidly growing cities experience elevated temperatures. This relationship aligns with established urban climate theory and corroborates recent African based studies by Kiarie et al. (2025) and Mutinda and Geletu (2025), who similarly documented population induced amplification of urban heat.

The methodological rigor of the study is reinforced by the stationarity and data quality diagnostics. Temperature series were found to be stationary at level, indicating dominance of short-term variability, while urban population series required first differencing due to their long run growth behavior. This justified modeling temperature directly while treating population effects cautiously to avoid spurious relationships. Grubbs outlier tests and visual diagnostics confirmed the absence of extreme observations in both datasets, supporting the reliability of the inputs. These findings are consistent with assumptions reported in prior temperature forecasting studies, including Nguyen et al. (2020), and provide a solid foundation for the subsequent modeling stages.

Performance comparisons across multiple models and validation schemes clearly establish the superiority of the ARIMA-LSTM hybrid framework. Across all-time series splits of 70:30, 80:20, and 90:10, the hybrid model consistently achieved the lowest RMSE and MAE values and the highest R^2 scores, demonstrating strong generalization

and addressing concerns related to overfitting associated with single train test evaluations. While standalone deep learning models such as RNN and MLP outperformed classical approaches in some cases, they remained consistently inferior to the hybrid model, confirming that nonlinear learning alone is insufficient when linear structure is ignored. Traditional models such as ARIMA, Prophet, SVR, and KNN exhibited higher errors and weaker explanatory power, echoing conclusions reported by Wang et al. (2018), Garcia et al. (2019), and Nguyen et al. (2020). Formal robustness checks further strengthened these findings, with Diebold Mariano tests and paired sample t tests confirming statistically significant performance gains for the ARIMA-LSTM model over all alternatives at the 1 percent significance level.

Beyond numerical accuracy, the results demonstrate strong practical relevance for decision making. The classification analysis showed that the ARIMA-LSTM model most reliably identified high temperature states, achieving the best balance of accuracy, F1 score, and AUC, which is critical for heat risk assessment and urban planning. Uncertainty analysis revealed near perfect 95 percent prediction interval coverage, indicating reliable representation of forecast uncertainty and outperforming several classical and standalone deep learning models. City specific forecasts further revealed persistent spatial contrasts, with northern cities such as Gulu, Lira, and Arua experiencing higher maximum temperatures, while highland cities like Kabale remaining relatively cooler. The presence of modest interannual variability alongside stable long-term trends highlight the value of combining ARIMA and LSTM components to capture both persistence and nonlinear behavior. Overall, the study contributes to African urban climate research by demonstrating, through multi split validation and formal statistical testing, that ARIMA-LSTM hybrid models provide a robust, interpretable, and generalizable framework for forecasting urban temperature dynamics in Uganda.

5. Validation of results, limitations, and policy implications

The study demonstrates that the ARIMA-LSTM hybrid model provides consistently superior performance for forecasting urban temperature dynamics across Ugandan cities. Validation was carried out using a broad set of regression and classification metrics, including RMSE, MAE, MAPE, SMAPE, R^2, accuracy, precision, sensitivity, F1 score, and AUC. To ensure robustness and address concerns about dependence on a single data split, time series aware rolling origin validation was applied across three split ratios of 70:30, 80:20, and 90:10. Across all splits, the ARIMA-LSTM hybrid achieved the lowest forecast errors and the highest explanatory power, with stable performance that indicates strong generalizability and reduced risk of overfitting. Statistical significance testing using the Diebold Mariano test and paired sample t tests

further confirmed that the hybrid model significantly outperformed all benchmark models, including classical statistical, machine learning, and deep learning approaches. These improvements were systematic and not driven by random variation or specific data partitions. In addition, a decision-oriented classification framework showed that the hybrid model translated numerical forecast accuracy into reliable identification of high temperature conditions, which is critical for early warning and risk management applications.

6. Conclusion

This study confirms that integrating ARIMA with LSTM in a hybrid framework offers a reliable and effective approach for forecasting urban temperature dynamics in Ugandan cities. By jointly capturing long-term linear trends and short-term nonlinear fluctuations, the ARIMA-LSTM model consistently outperformed traditional statistical methods and standalone machine learning models, highlighting the complex nature of urban climate behavior. The results underscore the value of combining statistical and deep learning techniques for climate prediction, while also pointing to the need for future research that incorporates longer time horizons and additional climatic variables such as rainfall and humidity. Importantly, the projected temperature increases, particularly in northern and central Uganda, signal an urgent need for city specific climate adaptation strategies, including urban greening, energy efficient infrastructure, and sustainable urban planning to reduce heat-related risks associated with climate change and rapid urbanization.

Recommendations

The findings of this study call for targeted and practical actions to address rising urban temperatures in Uganda. City specific climate adaptation plans are needed, especially in fast growing urban centers such as Gulu, Arua, and Lira, where heat risks are likely to intensify with continued population growth. Investment in green infrastructure including urban parks, tree planting, and green roofs should be prioritized as cost effective measures for reducing urban heat and improving thermal comfort. At the same time, sustainable urban planning practices that balance expansion with environmental protection are essential to limit heat exposure and excessive energy demand. From a research perspective, future studies should integrate additional climate variables such as humidity, rainfall, and wind speed to improve the realism and predictive strength of temperature models. Expanding the temporal coverage of datasets and including a wider range of urban environments will also enhance the models' capacity to capture long-term climate variability and rare extreme events, thereby strengthening their usefulness for policy and planning.

References

Chen, X., Liu, Y., (2018). Integrating ARIMA and LSTM for improved time series forecasting. *IEEE Transactions on Neural Networks and Learning Systems*, Vol. 29, No. 5, pp. 1923–1935.

Cheng, S., Shi, J., Cheng, Q., Zhou, X. and Zeng, S., (2025). Hybrid model for medium-term load forecasting in urban power grids. *Energies*, Vol. 18, No. 16, p. 4378. https://doi.org/10.3390/en18164378.

Garcia, M., Silva, R., Pereira, L., Santos, A. and Rocha, D., (2019). Urban climate resilience: Statistical modeling of temperature variability in Brazilian cities. *Climate Resilience Reviews*, Vol. 15, No. 1, pp. 150–165.

Gomez, F., Sanchez, M., Patel, R., Li, Q. and Zhao, Y., (2020). Hybrid statistical-machine learning models for environmental forecasting. *Journal of Environmental Science and Technology*, Vol. 18, No. 4, pp. 456–470.

Johnson, A., Smith, B. and Williams, C., (2016). Urban temperature dynamics: ARIMA modeling for climate forecasting. *Journal of Climate Change*, vol. 8, no. 3, pp. 210–225.

Jones, E., Brown, D., (2017). Limitations of ARIMA models in urban temperature forecasting. *Urban Climate Dynamics Review*, Vol. 12, No. 3, pp. 210–225.

Kiarie, J., Mwalili, S., Mbogo, R., Mutinda, J. and Langat, A. (2025). Statistical, machine learning, and deep learning models for COVID-19 forecasting in Kenya. *Computational and Mathematical Biophysics*, 13(1), 20250026. https://doi.org/10.1515/cmb-2025-0026.

Lee, X., Chen, Y., Huang, Z., Wang, M. and Zhang, L., (2017). Forecasting urban temperature dynamics: A comparison of Facebook Prophet and traditional models. *Environmental Science and Technology*, Vol. 12, No. 2, pp. 300–315.

Li, Z., Johnson, A., Perez, M., Brown, C. and Taylor, S., (2019). Urban climate resilience strategies: A review of recent literature. *Climate Resilience Reviews*, Vol. 15, No. 4, pp. 321–335.

Lynda, D., Logeswari, G., Tamilarasi, K. and Rakesh, S., (2025). Hybrid Bayesian deep learning model for predicting urban heat island intensity in African cities. *Scientific Reports*, Vol. 15, Article No. 31280. https://doi.org/10.1038/s41598-025-13492-4.

Mutinda, J. K., Geletu, A. (2025). Stock market index prediction using CEEMDAN-LSTM-BPNN-decomposition ensemble model, *Journal of Applied Mathematics*, *7706431*, 32 pages, https://doi.org/10.1155/jama/7706431.

Nguyen, H., Smith, J., Wang, T., Chen, L. and Lee, X., (2021). Advancements in environmental modeling using hybrid models. *Environmental Modeling and Software*, Vol. 30, No. 2, pp. 89–102.

Nguyen, T., Huynh, Q., Tran, P., Le, M. and Vo, D., (2020). Machine learning for urban temperature forecasting: A comparative study with traditional models. *Journal of Computational Statistics*, Vol. 30, No. 3, pp. 500–515.

Smith, A., Jones, B. and Brown, C., (2018). Urban temperature dynamics: Challenges and opportunities. *Journal of Climate Change*, Vol. 5, No. 2, pp. 123–135.

Tang, S., Jiang, Y., Su, H., Lim, K. M., Zheng, Z. and Zhu, Y., (2025). Wear defect detection of hydraulic pump using a hybrid method of VGG and LSTM. *International Journal of Hydromechatronics*, Forthcoming Articles, https://doi.org/10.1504/IJHM.2025.10073233.

Taylor, S. J., Letham, B., (2017). Facebook Prophet: A forecasting tool for the R programming language. *Journal of Computational Statistics and Data Visualization*, Vol. 25, No. 3, pp. 301–315.

Wang, H., Liu, J., Zhao, F., Chen, Q. and Yang, S., (2018). Hybrid ARIMA-LSTM modeling for urban temperature forecasting. *Journal of Environmental Science*, Vol. 25, No. 4, pp. 400–415.

Wang, L., Thompson, D., Green, S., Kim, J. and Lee, Y., (2020). Machine learning applications in environmental science: A comprehensive review. *Environmental Science: Processes & Impacts*, Vol. 22, No. 7, pp. 1567–1583.