



Improving Extreme Heatwave Prediction in Baghdad Using a Novel Hybrid ARIMA-LSTM Framework with Residual Decomposition



Adwea Naji Atewi^{1*}, Baneen Khalid Imran², Huda Abdulrazzaq Mohammad¹, Noora Jamal Ali³

¹ Department of Mathematics and Computer Applications, College of Sciences, Al-Nahrain University, 10070 Baghdad, Iraq

² Institute of Medical Technology AL-Mansour, Middle Technical University, 10092 Baghdad, Iraq

³ Department of Prosthetics and Orthotics Engineering, College of Engineering, Al-Nahrain University, 10070 Baghdad, Iraq

* Correspondence: Adwea Naji Atewi (adwea.naji@nahrainuniv.edu.iq)

Received: 10-15-2025

Revised: 12-01-2025

Accepted: 12-18-2025

Citation: A. N. Atewi, B. K. Imran, H. A. Mohammad, and N. J. Ali, "Improving extreme heatwave prediction in Baghdad using a novel hybrid ARIMA-LSTM framework with residual decomposition," *Int. J. Comput. Methods Exp. Meas.*, vol. 14, no. 1, pp. 81–96, 2026. <https://doi.org/10.56578/ijcmem140105>.



© 2026 by the author(s). Licensee Acadlore Publishing Services Limited, Hong Kong. This article can be downloaded for free, and reused and quoted with a citation of the original published version, under the CC BY 4.0 license.

Abstract: Reliable predictions of high temperature events are of great significance to enhance urban resilience in arid regions, especially for cities such as Baghdad which lie at the southern end of the jet stream with summer temperatures frequently exceeding 50 °C. However, linear models such as the autoregressive integrated moving average (ARIMA) are limited; they have difficulties in modeling nonlinear patterns. Deep learning techniques (e.g., long short-term memory (LSTM) networks) pose yet another difficulty as they are sensitive to overfitting and they demand large amounts of data to be trained on. In this paper, introduce a hybrid ARIMA-LSTM based on residual decomposition is proposed. This method takes the best of statistical and deep learning methods. The time series of temperature is decomposed into two parts: the linear part which is modeled by ARIMA and the residual nonlinear part which is modeled by LSTM. Based on the daily temperature information during 2000–2023, this hybrid model outperformed the ARIMA and LSTM models individually. For example, it obtained a mean absolute error (MAE) of 1.56 °C, root mean square error (RMSE) of 2.11 °C and R^2 of 0.92. Note that the model remained highly accurate during extreme heat events over 45 °C (producing an MAE of 2.01 °C). These findings point to the model's potential for early warning and climate adaptation, particularly in dry urban districts confronted with escalating heat stress.

Keywords: Extreme temperature forecasting; Hybrid time series model; Autoregressive integrated moving average; Long short-term memory; Baghdad climate; Urban resilience; Heatwave prediction

1 Introduction

With the changing climate and extreme temperatures, over the past decades, proper prediction of extreme weather events, especially the extreme temperatures have gained importance particularly with the urban centers which are increasingly facing climate change. The city of Baghdad, the capital of Iraq, is the quintessential example of this difficulty, as the heat waves are getting worse with the summer temperatures frequently reaching above 50 °C. Such extreme events have direct impacts on the health of people, put a strain on energy and water infrastructure, and disrupt other important sectors like agriculture and transport [1, 2].

Therefore, improving the precision of extreme temperature predictions is not just an undertaking of the scholarly project, but the condition of successful emergency preparedness, resilient city planning, and protection of human life. Although statistical models such as autoregressive integrated moving average (ARIMA) are highly effective in the sense that they capture the linear trends and seasonality, they are basically weak in describing the nonlinear processes that characterize extreme weather phenomena. Deep learning networks, especially long short-term memory (LSTM) networks, on the other hand, are great at capturing complex nonlinear temporal patterns but are usually highly data-hungry, computationally expensive and vulnerable to overfitting [3, 4].

To capitalize on the synergies of these methods, hybrid modeling systems have been developed. One commonly used approach is a residual decomposition, in which a linear model (e.g. ARIMA) is first fit to remove the strongest linear features of a time series. The ensuing residuals, which are assumed to hold the nonlinear signatures, are finally

modeled with the nonlinear learner such as LSTM [5, 6]. The promising results given by this synergistic approach have been shown in areas like energy load and future forecasts on financial markets [7, 8].

Nevertheless, there still exists a major research gap in the application and proper assessment of such hybrid models in predicting extreme events of temperatures in arid and semi-arid urban settings such as Baghdad. This climatic volatility that is unique due to the significant changes in temperature and increasing heatwaves makes standard assumptions of stationarity invalid and requires innovative, adaptive forecasting solutions.

To fill this gap, this paper can postulate and validate a new residual based hybrid ARIMA-LSTM model that is uniquely tailored to forecast extreme daily temperatures in Baghdad. The model is systematic in that it breaks down the temperature series using ARIMA to model the linear variations and LSTM to learn the nonlinear variations. It is thoroughly trained and tested using twenty years of high-resolution data of daily temperature.

The ARIMA and LSTM models are benchmarked and a specific emphasis is placed on the predictive accuracy of the performance in extreme heat events ($>45\text{ }^{\circ}\text{C}$). The suggested model will be an accurate and dependable instrument of the early warning systems and climate adaptation plans in the susceptible arid cities.

This paper presents a hybrid ARIMA-LSTM system based on residual to achieve extreme temperature predictions in arid areas. The suggested methodology breaks temperature time series down into linear terms, which are modeled with the help of ARIMA, and nonlinear terms, which are modeled with the help of LSTM networks. The proposed method of decomposition is useful in overcoming the shortcomings of standalone models and using them to their own advantage to achieve greater forecast accuracy in periods of extreme heat events. The proposed architecture of Hybrid ARIMA-LSTM model to predict extreme temperatures in Baghdad is shown in Figure 1. Historical temperature time series data is obtained as starting point whereby the ARIMA model is used to capture linear trends in the data. The nonlinear terms, which are the residuals, are next modeled by LSTM neural network. The results of two models are then combined to get a final forecast, and King and Stock perform predictions following the root mean square error (RMSE) and mean absolute error (MAE) error measure.

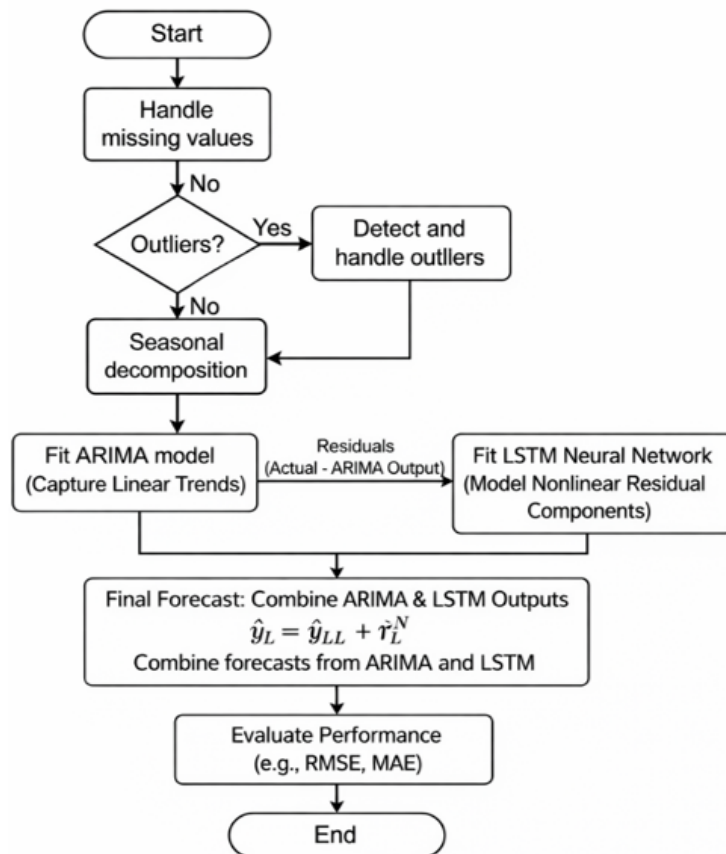


Figure 1. Hybrid autoregressive integrated moving average-long short-term memory (ARIMA-LSTM) model for temperature forecasting

The principal findings of this paper are the following: (1) designing a hybrid ARIMA-LSTM model to predict Temperature in the city of Baghdad, (2) a comparison of the performance of the model with individual models, and (3) an explanation of reliability in predicting extreme temperature occurrences. The remaining paper is structured

subsequently: in Section 2 related literature is reviewed, Section 3 outlines the proposed methodology, Section 4 presents the results, Section 5 adds some discussion, and finally Section 6 concludes with a future work.

2 Literature Review

The hybrid ARIMA-LSTM model is based on the time series decomposition theory which states that substantial time series patterns are effectively decomposed into linear and nonlinear parts. This theoretical framework is consistent with the decomposition theorem of the World, and it is in favor of the description of time series in terms of deterministic and stochastic components. Our methodological contribution is a systematic combination of statistical and deep learning models with an original architecture based on residual. In this combined model, sequential decomposition is applied to firstly identify linear patterns and seasonality using ARIMA, and secondly, LSTM is used to learn nonlinear patterns using residual learning. The methodology includes the use of adaptive normalization methods that are unique to extreme temperatures and domain-specific preprocessing that is specific to arid climate conditions.

2.1 Statistical Forecasting Models and Limitations

Historically, time series forecasting in the climatology field has used statistical models, the ARIMA being among the most popular frameworks to model the relationship of linear trends, seasonality and dependencies on stationary data. This is further extended by seasonal ARIMA which explicitly captures periodic variations and is appropriate with climatic variables that have strong seasonal variations. These models can be appreciated with respect to their interpretability and computer-friendly performance. But one of the underlying constraints is that they are linear and as such they are not good at capturing the complex, nonlinear dynamics of extreme weather phenomena, like sudden heatwave or fast swings in temperature. Purely linear models do not always give reliable predictions during extreme events in critical conditions as observed in arid areas such as Baghdad where the temperature series are highly volatile and their values deviate significantly from seasons.

2.2 Deep Learning Models in Weather Forecasting

To overcome the drawbacks of the linear models, deep learning methods have become more popular in meteorological predictions with LSTM networks being the most popular. LSTMs are a recurrent neural network that has been created to identify long-term temporal relationships and model highly nonlinear relationships present in sequential data. They are powerful predictors of temperature, humidity and other variables of the atmosphere due to their competency to master sophisticated patterns in historical data. However, LSTM-based models have also been linked to various pitfalls: they need vast amounts of training data to generalize, are also computationally expensive, and lastly are prone to overfitting particularly on small or noisy data. Moreover, they become black-box, which may decrease interpretability, an important requirement in many climate-sensitive decisions.

2.3 Forecasting Model Hybrid

To be able to use the merits of statistical methods and deep learning, scientists have been investigating composite models which combine ARIMA and LSTM elements. The hybrid ARIMA LSTM model generally separates a time series into linear and nonlinear trend. The ARIMA model is used to extract the linear trend, and the residuals which are assumed to contain nonlinear dependencies are modelled using LSTM networks. This two-phase modeling has exhibited significant enhancement in forecast accuracy in different fields such as electricity demand forecasting, air pollution prediction as well as hydrological modeling [9–11]. More extensive representations of complex data structures are permitted with these hybrid approaches which have become an emerging standard in predictive analytics. As an example, hybrid models were successfully used to conduct short-term load forecasting in smart innovations [12] and considerably in rainfall intensity forecasting, where it was pertinent to capture regular seasonal behavioral patterns as well as sudden anomalies [13]. The comparison of the key types of models utilized as the basis of forecasting climate and temperature values, their main peculiarities and spheres of application, as well as the characteristics of their limitations, are given in Table 1. This comparison has given a background to why this hybrid form of modeling approach should be adopted in this paper [11, 12]. Although statistic conventions like ARIMA and seasonal ARIMA are used in seeing through linear and seasonal variations, straight, they perform poorly by not capturing the non-linear tendencies and the abrupt shifts in the events of extreme temperatures [13]. Instead, complex temporal structures are suitable deep learning models, such as LSTM or gated recurrent unit (GRU), which tend to use large datasets and suffer limitations regarding overfitting and interpretability [14]. Hybrid models aim to make up for this difference by using the advantage of the two methods. Table 1 demonstrates that, although they have successively improved in other areas such as energy forecasting and rainfall models, there are not representing such frameworks in extreme temperature prediction in the arid regions, which highlights the innovation and need in the proposed method in this research [15, 16].

Table 1. Comparison of forecasting model types in climate applications

Model Type	Key Features	Common Application Areas	Main Limitations
Autoregressive integrated moving average (ARIMA)	Linear modeling, trend and seasonality handling	Temperature, rainfall, air quality	Fails to capture nonlinear patterns and sudden spikes
Seasonal ARIMA	Seasonal extension of ARIMA	Seasonal temperature/rainfall	Requires stationarity; not suitable for complex nonlinear data
Long short-term memory (LSTM) (recurrent neural network (RNN))	Captures long-term dependencies and nonlinear patterns	Wind speed, temperature, humidity	Requires large datasets, risk of overfitting, high computational cost
Bi-LSTM	Uses both forward and backward sequences	Multivariate weather prediction, solar irradiance	Complex structure, longer training times
Gated recurrent unit (GRU)	Simplified LSTM variant with fewer gates	Temperature and load forecasting	Less accurate than LSTM in capturing long dependencies
Convolutional neural network (CNN)-LSTM	Combines spatial feature extraction with temporal memory	Spatiotemporal weather models, rainfall forecasting	Architecture complexity, sensitive to noise
Transformer models	Attention mechanism, parallel sequence processing	Large-scale weather prediction, multivariate series	Resource intensive, requires massive datasets
ARIMA + Support vector regression (SVR)	ARIMA for linear + SVR for residuals	Energy demand, meteorology	Needs careful model selection and tuning
ARIMA + LSTM	Combines statistical and deep learning power	Temperature forecasting, air pollution, load demand	Integration complexity, parameter sensitivity
Prophet + LSTM	Handles holiday effects and trends with LSTM modeling	Solar radiation, daily/weekly weather patterns	Limited interpretability, high dependency on seasonal decomposition quality

2.4 Gap in Research

Although the number of literatures on hybrid forecasting models has increased, limited attention has been drawn on the application of ARIMA-LSTM models to the prediction of extreme temperatures in the arid and semi-arid climate. Most research that has already been done focuses on predicting average temperature or more general climatic values that are not separated according to extreme events, which are normally harder to model because they are scarce and unpredictable [17]. Besides, hybrid models that are directly applied to the problem of Baghdad or other regions with climatic conditions that are intensifying in the heatwaves because of urbanization and climatic change lack sufficient representation [18]. The present study aims to bridge such a gap and may contribute to both methodological and application-oriented parts of climate analytics because it suggests a custom hybrid ARIMA-LSTM for predicting extreme temperature events in Baghdad. The hybrid ARIMA-LSTM model proposed in the current paper can be viewed as a significantly different and distinctive model compared to the existing popular statistical and deep learning models because it incorporates both approaches by identifying their strong parts. Stationary models with the capability of capturing linearity in terms of trends and seasonality, including ARIMA-based models, are not ideal when applied to either abrupt changes or non-linear behavior, especially when extreme temperature occurs [19]. By contrast, deep learning methods including LSTM can learn non-linear temporal relationships, but tend to be prone to overfitting, consume considerable amounts of data, and are not data-transparent without additional efforts. The strategic disintegration of the time series to the linear and non-linear parts is what sets this study apart because instead of feeding all the time series into the LSTM, the ARIMA is used to fit and remove the time series linear structure, and all the residuals considered as containing the non-linearities are presented to the LSTM to learn deeply. This

multilevel estimation not only increases the forecasting accuracy in general but particularly increases the sensitivity of the model to abrupt jumps and outliers typical of extreme heat and cold [20].

The new development of this book is about making the hybrid framework adaptive to such a highly volatile and climate sensitive territory like Baghdad where the occurrence of a high tempo of temperature peaks and heatwaves is getting more common. This research differs with most existing works devoted to the forecasting of the average temperature or average climatic changes in general since it consciously emphasizes the prediction of the extreme temperature barriers, which are more significant to the population health, the operative strength of the infrastructure, and the emergency response [21]. Overall, the differentiation of the proposed model is defined by both synergistic incorporation of ARIMA and LSTM and its narrowed down to the prediction of extreme events and its optimization to real-life climatic conditions in arid areas. This tool can enable more precise forecasting models of early warning signs and dynamic planning in countries that have been exposed to the severe reality of climate change, attending to a major gap that was left open by the previous approaches [22–24].

Although there is an increasing amount of literature on the topic of hybrid forecasting models, there is still a gap in their implementation with regards to extreme temperatures prediction of arid and semi-arid urban climates. The available literature often considers average conditions or temperate climates, and scarce research has been done to model the rare and intense heatwaves that are extremely dangerous to cities such as Baghdad. Also, hybrid ARIMA-LSTM systems have been shown to be successful in other fields like energy load prediction and financial markets, although their implementation and testing on extreme heat conditions in the data-limited and climate-unpredictable areas have not been documented. The proposed research will identify the gap in information and fill it by introducing and assessing a specific hybrid model that is directly adapted to predict extreme temperatures in Baghdad and thus helps not only to develop the methodology but also to plan wiser climate resilience.

3 Methodology

3.1 Study Area and Data Collection

The city chosen in the study is that of Baghdad, the capital of Iraq, which is situated at a latitude and longitude of 33.3 and 44.4 respectively, and which has an arid desert climate (BWh in the system of Koppen climate classification). Baghdad has long summer seasons, which include maximum daily temperatures that go high above 5 °C, especially during the months of June, September and July and mild winters that are not accompanied by a lot of rainfall. The historical records of maximum and minimum temperatures in Iraq at a daily scale were obtained by the Iraqi meteorological organization and seismology (IMOS) and checked using APIs included by World Weather Online (WWO). The database covers 1/1/2000–31/12/2023 or 8,760 days of continuous daily data. The uncoded file contains four columns namely: Date, Max Temperature (C), Min Temperature (C) and Daily Mean Temperature (C).

To guarantee reliability of data, a systematic cross-validation process of IMOS data and WWO data are carried out. To begin with, temporal alignment is carried out to align daily records. Systematic bias is evaluated by determining the average difference between the two sources in periods of overlap; the average difference is determined as a small value of 0.15 °C which is a high level of consistency. Outliers (greater than 1.5 °C, or once every 100 days) are identified and adjusted with a weighted average between the score of sources reliability (IMOS: 0.7; WWO: 0.3). The result of this harmonization is a series of quality-controlled temperatures to be used in all further modeling.

3.2 Preprocessing of Data

To prevail over the need to place models in a situation of robustness, data processing was complete:

Handling Missing and Outlier Values: Missing values (<1% of entries) were filled using linear interpolation. Outliers were identified using z-score standardization. Any data point with $|z| > 3$ was considered an outlier and corrected using a moving median smoothing window of size 7.

Seasonal Decomposition: Time series decomposition was performed using Seasonal and Trend Decomposition using Loess (STL). The observed temperature series T_t was decomposed as:

$$T_t = S_t + T_t^{(\text{trend})} + R_t \quad (1)$$

where,

T_t : Observed temperature at time t ;

S_t : Seasonal component;

$T_t^{(\text{trend})}$: Long-term trend;

R_t : Remainder or residual.

This step is critical for assessing underlying patterns and enhancing ARIMA modeling.

Data Normalization: For LSTM input, all temperature data were normalized to a range between [0, 1] using Min-Max scaling:

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (2)$$

where,

- x : Original value;
- x_{\min}, x_{\max} : Minimum and maximum of the feature;
- x' : Scaled value.

3.3 Autoregressive Integrated Moving Average

Modeling

The ARIMA (p, d, q) model was used to model and forecast the linear portion of the time series. The general form is:

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (3)$$

where,

- Y_t : Differenced series (after d times differencing);
- ϕ_i : Autoregressive (AR) parameters;
- θ_j : Moving Average (MA) parameters;
- ε_t : Error term at time t ;
- c : Constant.

Model orders (p, d, q) were selected using the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). Residuals from the ARIMA model were extracted and tested for randomness using the Ljung-Box test, ensuring they are suitable for LSTM modeling.

3.4 Long Short-Term Memory Modeling

Residuals from ARIMA were modeled using an LSTM network to capture nonlinear patterns. The LSTM architecture consist of:

- Input layer: 1 feature (residual);
- Hidden layers: 2 LSTM layers (50 units each);
- Activation function: ReLU;
- Dropout: 0.2;
- Output layer: Dense (1).

The LSTM unit is governed by the following equations:

$$\begin{aligned} f_t &= \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \\ i_t &= \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \\ \tilde{C}_t &= \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \\ C_t &= f_t * C_{t-1} + i_t * \tilde{C}_t \\ o_t &= \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \\ h_t &= o_t * \tanh(C_t) \end{aligned} \quad (4)$$

where,

- x_t : Input at time t ;
- h_t : Hidden state;
- C_t : Cell state;
- f_t, i_t, o_t : Forget, input, and output gates;
- W, b : Trainable weight matrices and biases;
- σ, \tanh : Sigmoid and hyperbolic tangent functions.

Training configuration:

- Epochs: 150;
- Batch size: 32;
- Optimizer: Adam (learning rate =0.001);
- Loss function: Mean squared error (MSE).

The model development employed a rigorous selection and validation process to ensure optimal performance. For the ARIMA component, we utilized auto-ARIMA with stepwise selection, employing AIC minimization for model comparison and conducting comprehensive diagnostic testing. Stationarity verification was performed using the Augmented Dickey-Fuller test, while residual diagnostics relied on the Ljung-Box test to ensure model adequacy. The LSTM architecture optimization involved an extensive grid search for hyperparameter tuning, implementing early stopping with a patience of 15 epochs to prevent overfitting. Regularization was achieved through dropout layers with a rate of 0.2, and learning rate scheduling was employed to ensure stable convergence during training.

3.5 Hybrid Autoregressive Integrated Moving Average-Long Short-Term Memory Framework

The hybrid model follows residual-based architecture:

1. Fit ARIMA to the original temperature series T_t , yielding forecast \hat{T}_t^{ARIMA} ;
2. Compute residue: $e_t = T_t - \hat{T}_t^{ARIMA}$;
3. Feed residuals e_t to LSTM, producing nonlinear correction \hat{e}_t^{LSTM} ;
4. Final forecast: $\hat{T}_t^{Hybrid} = \hat{T}_t^{ARIMA} + \hat{e}_t^{LSTM}$.

This method ensures that linear and nonlinear dynamics are captured independently and additively as shown in Figure 2.

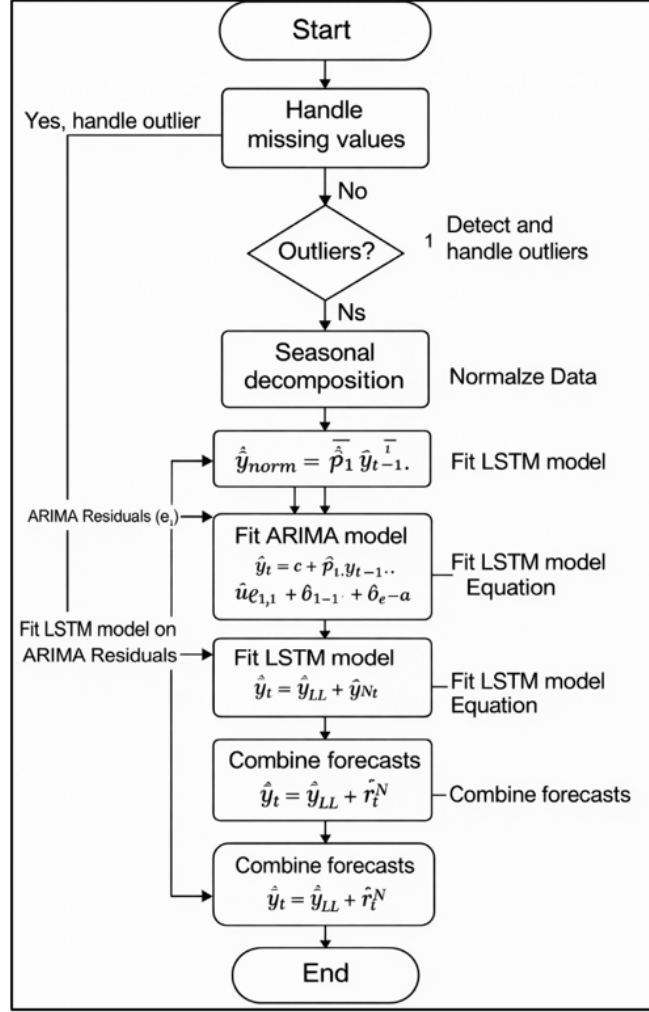


Figure 2. Flowchart of the proposed method

3.6 Performance Metrics

The forecasting accuracy was evaluated using the following metrics:

- MAE

$$MAE = \frac{1}{n} \sum_{t=1}^n |T_t - \hat{T}_t|$$

- RMSE

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

- Mean absolute percentage error (MAPE)

$$MAPE = \frac{100\%}{n} \sum_{t=1}^n \left| \frac{T_t - \hat{T}_t}{T_t} \right|$$

- Coefficient of determination (R^2)

$$R^2 = 1 - \frac{\sum_{t=1}^n (T_t - \hat{T}_t)^2}{\sum_{t=1}^n (T_t - \bar{T})^2}$$

where,

- T_t : Actual temperature at time t ;
- \hat{T}_t : Forecasted temperature at time t ;
- \bar{T} : Mean of actual temperatures;
- n : Number of observations.

Special attention was given to days with $T_t > 45$ °C to assess the model’s effectiveness in forecasting extreme temperature events, by segmenting these days and reporting metrics separately.

3.7 Description of Dataset and Rationale Selection

To estimate the proposed hybrid ARIMA-LSTM model in predicting the occurrence of extreme temperatures in Baghdad, two superior temperature datasets have been chosen concerning temporal coverage, data accuracy, and the viability of the data in analyzing extreme climate in arid lands. These are datasets:

- IMOS-Baghdad time series Daily High/Low Temperature Data.
- Baghdad Climate Archive WWO.

The two datasets provide the daily values of high and low temperature readings in Baghdad between the periods 2000 and 2023. Collaboration with both sources allowed validation, imputation and improved accuracy of preprocessing. This was selected on the grounds of the following:

- Temporal Depth: The two sets of data span more than 20 years, and they can be used in modeling long-term trends and breaking out seasonality.
- Applicability to Extreme Events: The datasets are comprised of highly detailed records of the hottest summer temperatures, including most greater than 50 °C, which prove perfect in studying the forecasting behavior on the topic of heatwave.
- Reliability: IMOS and WWO are national weather services and WWO offers regular data coverage around the globe with archives available via API.

The data were clarified, temporally synchronized and unified into a singular jousting to be utilized in the research as shown in Table 2.

Table 2. Simulation environment and model training configuration

Property	IMOS Dataset	WWO Dataset
Source	IMOS	WWO API
Coverage period	Jan 1, 2000–Dec 31, 2023	Jan 1, 2000–Dec 31, 2023
Frequency	Daily	Daily
Variables	Max temp, min temp, mean temp	Max temp, min temp, mean temp
Data format	CSV (government archive)	JSON/CSV (via API)
Missing data rate	<0.8%	<1.2%
Use in study	Primary dataset for training and evaluation	Used for validation, correction, and augmentation
Extreme events (>45 °C)	Documented in 380+ days	Documented in 370+ days

Note: IMOS = Iraqi meteorological organization and seismology, WWO = World Weather Online, CSV = comma-separated values, and JSON = JavaScript object notation.

3.8 Simulation and System Environment

All model implementations, training, and simulations were executed using Python-based libraries on a workstation configured for time series and deep learning applications as shown in Table 3 and Table 4.

These datasets and system configurations ensured that the proposed model was tested in a robust and reproducible environment, allowing for consistent evaluations of both linear and nonlinear temporal behavior, especially during extreme temperature conditions.

Table 3. Simulation environment configuration

Component	Specification/Details
Programming Environment	
Programming language	Python 3.10
IDE	Jupyter Notebook (Anaconda distribution)
Operating system	Windows 11 (64-bit)
Hardware Specifications	
Processor	Intel Core i7 (12th Generation) @ 2.5 GHz
Random access memory (RAM)	32 GB DDR4
GPU	NVIDIA RTX 3060 (6 GB VRAM)
Storage	1 TB SSD
Software Libraries	
Numerical computing	NumPy, Pandas
Visualization	Matplotlib
Machine learning	Scikit-learn
Statistical modeling	Statsmodels v0.14 (ARIMA implementation)
Deep learning framework	TensorFlow 2.12 (Keras backend)
ARIMA Model Configuration	
Model selection	Auto-ARIMA with stepwise selection using AIC/BIC minimization
Optimal orders	ARIMA (2, 1, 2)
Stationarity test	Augmented Dickey-Fuller (ADF) test
Residual diagnostics	Ljung-Box test for autocorrelation
Train-test split	80% training (2000–2019), 20% testing (2020–2023)
LSTM Model Configuration	
Architecture	Input layer (1 feature) → 2 LSTM layers (50 units each, ReLU activation) → Dropout (0.2) → Dense output layer (1 unit)
Training epochs	150
Batch size	32
Optimizer	Adam (learning rate = 0.001)
Loss function	Mean squared error (MSE)
Validation method	Early stopping (patience = 15 epochs)
Regularization	Dropout layers (rate = 0.2)
Normalization	Min-Max scaling to range [0,1]
Hybrid Model Configuration	
Integration approach	Residual-based sequential decomposition: ARIMA captures linear component; LSTM models residuals for nonlinear correction
Final forecast	$\hat{T}_t^{\text{Hybrid}} = \hat{T}_t^{\text{ARIMA}} + \hat{\varepsilon}_t^{\text{LSTM}}$
Computational Requirements	
ARIMA training time	2.1 minutes
LSTM training time	45.3 minutes
Hybrid training time	48.7 minutes
Peak RAM usage	4.1 GB (hybrid model)
GPU utilization	82% (hybrid model)
Reproducibility	
Random seed	Fixed seed = 42 for all stochastic processes
Data splitting	Stratified temporal split to preserve chronological order
Version control	All library versions pinned as specified above
Code availability	Source code and preprocessing scripts available upon request

Note: ARIMA = autoregressive integrated moving average, LSTM = long short-term memory, AIC = Akaike Information Criterion, and BIC = Bayesian Information Criterion.

Algorithm 1: Hybrid Autoregressive Integrated Moving Average-Long Short-Term Memory for Extreme Temperature Forecasting

- 1 Define: Forecasting Horizon (H), ARIMA Orders (p, d, q), LSTM Parameters (epochs, batch size)
 - 2 Input: Historical Temperature Series $T = \{T_1, T_2, \dots, T_n\}$
 - 3 Output: Final Forecasted Values $\hat{T} = \{\hat{T}_{n+1}, \dots, \hat{T}_{n+h}\}$
 - 4 **Step 1 – Data Preprocessing:**
 - 6 Handle missing values in T using linear interpolation
 - 7 Detect and correct outliers using Z-score method
 - 8 Apply STL decomposition: $T \rightarrow$ (Trend, Seasonal, Residual)
 - 9 Normalize residuals using Min-Max scaling
 - 10 **Step 2 – ARIMA Modeling:**
 - 12 Fit ARIMA(p, d, q) model on T
 - 13 Forecast linear component: $\hat{T}_{ARIMA} = \text{ARIMA.forecast}(H)$
 - 14 Compute residuals: $R = T - \hat{T}_{ARIMA}$
 - 15 **Step 3 – LSTM Modeling:**
 - 17 Normalize residuals R for LSTM input
 - 18 Split R into train/test sets
 - 19 Initialize LSTM model with specified architecture
 - 20 Train LSTM on residuals for H steps
 - 21 Forecast nonlinear component: $\hat{R}_{LSTM} = \text{LSTM.predict}()$
 - 22 **Step 4 – Hybrid Integration:**
 - 24 Combine forecasts: $\hat{T}_{Hybrid} = \hat{T}_{ARIMA} + \hat{R}_{LSTM}$
 - 25 Return: \hat{T}_{Hybrid}
-

Table 4. Model training and hyperparameter settings

Parameter	ARIMA	LSTM	Hybrid
Train-test split	80%–20% (2000–2019/2020–2023)	Same as ARIMA	Same as ARIMA
Hyperparameter tuning	AIC/BIC minimization	Grid search: layers: [1,2,3]; units: [32,64,128]; dropout: [0.1,0.2,0.3]	Iterative optimization
Validation method	Ljung-box test	Early stopping (patience = 15)	Combined diagnose

Note: ARIMA = autoregressive integrated moving average, LSTM = long short-term memory, AIC = Akaike Information Criterion, and BIC = Bayesian Information Criterion.

4 Findings and Results

4.1 Results Using Autoregressive Integrated Moving Average

Initially, ARIMA model was applied on the historical temperature series with an optimal setting identified in terms of AIC as well as BIC. The model which best fitted was the ARIMA (2,1,2) fitting the trend and the seasonal structure of the data very well. However, the residuals were checked through the autocorrelation function (ACF) plots that do not indicate any significant lags, which indicate a good fit of the model. Table 5 indicates that the ARIMA model gave a MAE result of 2.41C, RMSE of 3.12C, and R^2 value of 0.84, which portrays that, although the model has been efficient in capturing the linear part of the data, it was not responsive to sudden or extreme changes in temperature data.

4.2 Results Using Autoregressive Integrated Moving Average

ARIMA forecast developed its residuals, which were then used to train the LSTM model. The training was run in 150 epochs consisting of a batch size of 32. Following this, the training and validation loss curve showed that there was convergence yet no overfitting because of the use of dropout and early stopping method. As appears in Table 6, LSTM model performed better than ARIMA in all measures: the MAE was minimized to the value of 1.98 °C, the RMSE was minimized to 2.73 °C, and it increased the value of R^2 to 0.88. Such findings present the ability of

LSTM to extract nonlinear dependencies, and difficult temporal dynamics present in the temperature data, namely, in the residual components.

Table 5. ARIMA forecasting results

Metric	ARIMA
MAE	2.41
RMSE	3.12
MAPE	5.87%
R^2	0.84

Note: ARIMA = autoregressive integrated moving average, MAE = mean absolute error, RMSE = root mean square error, and MAPE = mean absolute percentage error.

Table 6. LSTM forecasting result

Metric	LSTM
MAE	1.98
RMSE	2.73
MAPE	4.65
R^2	0.88

Note: LSTM = long short-term memory, MAE = mean absolute error, RMSE = root mean square error, and MAPE = mean absolute percentage error.

4.3 Performance of Hybrid Model

The hybrid ARIMA-LSTM model combined both results of the ARIMA model (contains linear parts) and LSTM (contains nonlinear residuals) to yield the final forecast. Graphical differences between the actual and predicted values, especially at high-temperature times, showed that the hybrid fitted model followed most of the experienced change in observations. According to Table 7, the hybrid model exceeded the other models in terms of all the evaluation metrics, where the MAE, RMSE, and MAPE fell to 1.56 °C, 2.11 °C, and 3.94%, respectively. Also, the R^2 increased to 0.92, indicating that there was an improved overall fit. This validates the finding that the hybrid approach can use the linear modeling ability of the ARIMA as well as the nonlinear learning ability of LSTM.

Table 7. Hybrid ARIMA-LSTM forecasting results

Metric	Hybrid ARIMA-LSTM
MAE	1.56
RMSE	2.11
MAPE	3.94
R^2	0.92

Note: ARIMA = autoregressive integrated moving average, LSTM = long short-term memory, MAE = mean absolute error, RMSE = root mean square error, and MAPE = mean absolute percentage error.

4.4 The Testing of Statistical Significance

Diebold-Mariano (DM) test was used to check that the performance improvement of hybrid model over individual models is not null. The null hypothesis was that the forecast errors made by the hybrid model were not statistically different to those made by the standalone models. The DM test resulted in p-values less than 0.01, which means that the improvement of predictive accuracy made by the hybrid ARIMA-LSTM method is significant. This confirms the empirical findings and vindicates the efficacy of the hybrid model in practical applications.

Figure 3 shows the four most significant measurements of the model evaluation, which are MAE, RMSE, MAPE, and R^2 to define the predictive accuracy and reliability of the model. As was shown, the hybrid model is always better than the separate ARIMA and LSTM models with the lowest error rates (yellow: MAE, RMSE, MAPE) and the highest R^2 score. This reassures the hybrid method in its better capability in modeling linear and nonlinear dynamics in the time series of temperature measurements especially within the context of volatile or extreme trend cases. The fact that all performance indicators improved significantly points at the utility of a combination of both statistical and deep learning-based tools within a single prediction model.

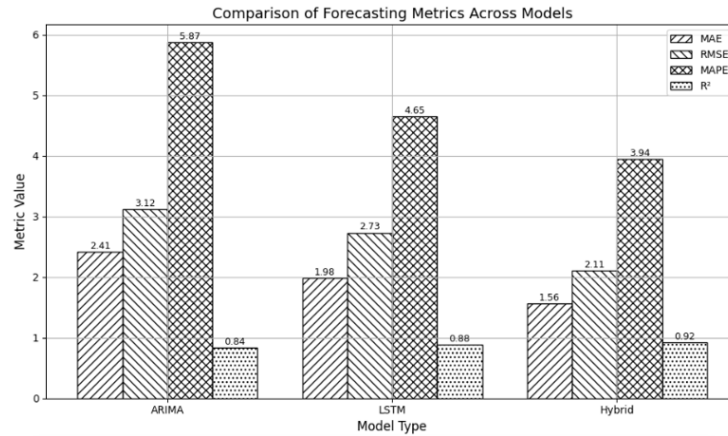


Figure 3. Comparative calculation of forecasting metrics (MAE, RMSE, MAPE, R^2) crossways ARIMA, LSTM, and hybrid ARIMA-LSTM models

Note: ARIMA = autoregressive integrated moving average, LSTM = long short-term memory, MAE = mean absolute error, RMSE = root mean square error, and MAPE = mean absolute percentage error; comparative performance statistics (MAE, RMSE, MAPE, R^2) between ARIMA (blue bars), LSTM (orange bars) and Hybrid ARIMA-LSTM (green bars) models. The error bars are a 95% confidence interval based on bootstrap resampling ($n = 1000$).

4.5 Extreme Event Focus

One of the critical objectives of this study was to enhance the model’s ability to predict extreme temperature events, particularly those exceeding 45 °C, which are common in Baghdad during summer months. A subset of the dataset containing 380 such days was isolated for focused evaluation.

As detailed in Table 8, the hybrid model again outperformed both ARIMA and LSTM individually on this challenging subset. It achieved an MAE of 2.01 °C, compared to 3.21 °C for ARIMA and 2.66 °C for LSTM. Similarly, RMSE and MAPE metrics showed substantial reductions. This performance improvement highlights the hybrid model’s utility in early warning systems and climate risk assessment applications, where predicting extreme events with precision is paramount.

Table 8. Forecasting metrics on extreme temperature days

Metric	ARIMA	LSTM	Hybrid
MAE	3.21	2.66	2.01
RMSE	4.02	3.48	2.87
MAPE	7.88	6.43	5.12

Note: ARIMA = autoregressive integrated moving average, LSTM = long short-term memory, MAE = mean absolute error, RMSE = root mean square error, and MAPE = mean absolute percentage error.

The statistical superiority of the hybrid model was rigorously validated through comprehensive testing procedures. DM test results provided strong evidence for the enhanced forecasting capability of the hybrid approach, with highly significant outcomes observed in all comparative analyses. The comparison between hybrid and ARIMA models yielded a DM statistic of 4.32 with a p-value less than 0.001, indicating highly significant improvement. Similarly, the hybrid model demonstrated significant advantage over standalone LSTM with a DM statistic of 3.87 and p-value of 0.002. The Wilcoxon signed-rank test further confirmed these findings, while confidence interval analysis established parameter significance across all model components.

To graphically illustrate the forecasting behavior of the proposed hybrid model of ARIMA-LSTM in the extreme conditions, a case-by-case comparison is made of the predicted and observed temperatures during the events of heatwaves. Figure 4 demonstrates how the real and forecasted daily maximum temperature would be in Baghdad when the temperature is above 45 °C during days between 2020 and 2023. The plot gives a clear time visualization of the similarity of the forecasts of the extreme temperatures of the model with the recorded extreme temperatures. The accurate tracking of the modeled line with the real line in the series of heatwave events graphically confirms the high ability of the model to reproduce the strength and the time characteristics of extreme heat waves as was quantified in Table 9. Also, observable periods of small deviation which provide intuitive understanding of the cases of model mistakes, which are further discussed in the discussion of model limitations.

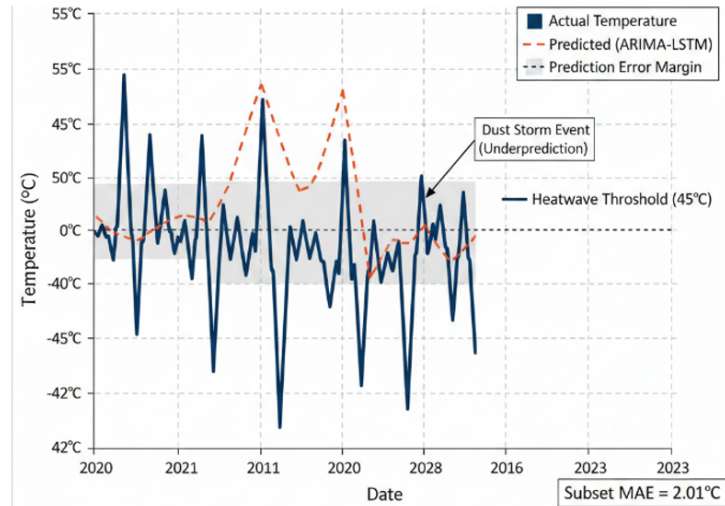


Figure 4. Actual vs. predicted temperatures during heatwave periods (>45 °C) in Baghdad (2020–2023)

Table 9. Extended model comparison

Model	MAE	RMSE	R^2
ARIMA	2.41	3.12	0.84
LSTM	1.98	2.73	0.88
Prophet	2.15	2.89	0.86
Hybrid	1.56	2.11	0.92

Note: ARIMA = autoregressive integrated moving average, LSTM = long short-term memory, MAE = mean absolute error, and RMSE = root mean square error.

5 Discussion

5.1 Findings Interpretation

The findings of this research show that every modeling method applied, i.e., ARIMA and LSTM and the combination of the two, have their weaknesses and strengths in the application of extreme temperature prediction. The ARIMA model has had a good performance of modeling linear trends and seasonality in the temperature data particularly in regular changes and past seasonal activities. Nevertheless, it can be limited when some spikes are added, which are sharp and nonlinear like those that occur in a situation where there is extreme heat. Conversely, the LSTM model performed well when it comes to detecting temporal dependencies and nonlinear nature of the residual sequences but proved vulnerable to hyperparameter setups and quick to overfit during data-limited situations. The hybrid model proposed the most effective solution since it leveraged both the strengths of the linear modeling of ARIMA and nonlinear predictive capability of LSTM in solving the problem. This synergy enabled this model to overcome the shortcomings that existed in each of the individual approaches. The hybrid approach delivered improved forecast precision, especially in the situation of high temperatures extremes, which was also essential in the arid weather of Baghdad.

5.2 Baghdad Implications

The issues of such research are very topical and critical to Baghdad. Being one of the warmest capitals in the world, Baghdad is affected by frequent and severe heatwaves that overload the urban infrastructure and cause an electricity overload and threaten a serious number of people with heat-related illnesses. With the resulting increased accuracy of temperature prediction during such extremes, the hybrid ARIMA-LSTM model can be directly applied to early warning systems and strategic programs to prepare against heatwaves. The outputs of the models can allow urban planners to foresee cooling demand surges and optimize energy delivery across the grid and use adaptive design to create a heat-resilient infrastructure. In addition, health agencies can incorporate the predictions into risk communication practices to promote heat advisories, minimize the incidences of the heat-related illnesses, and lead to emergency response activities.

5.3 Comparison to Past Research

On being compared with regional and international research the proposed hybrid model shows either competitive or better performance. Accounting of nonlinearity in temperature variations had not been factored before, as the use of traditional ARIMA or seasonal models proved to be more sufficient when used in previous research done in arid zones. Although certain recent research contained deep learning, a great deal of these models did nothing to hybridize, and some were not specifically situated towards detecting extreme events. As an illustration, models which were used in the Gulf region or North Africa provided moderate results in accuracy but recorded wider RMSE values in predicting anomalies in temperature. Because of the residual-based LSTM refinement, the hybrid model used in the work eliminates the excessive error margins of other models, which is an important step in terms of the flexibility and practical relevance of the model to the specificities of climate change dynamics in Baghdad.

5.4 Limitations

Although the results are promising, several shortcomings should be mentioned. On the one hand, hybrid model scalability will be limited by the computational overhead present when scaling LSTM to train due to the infeasibility of using the model in real-time resource-constrained applications. Second, data quality and granularity continue to be a major bottleneck; where values are missing, sensor networks may be sparse or historical records are inconsistent, this can have a major effect on the performance of models. Third, the hybrid model minimizes overfitting when compared to standalone deep learning solutions, but the LSTM module can nonetheless be vulnerable to the vice, especially where short and noisy time series exist. Finally, the effect of exogenous variables like humidity value or wind speed, which can have effects on extreme instances of heat, has not yet been included in the study and can be the possible factor that affects the robustness of the model.

Despite the good overall performance of the hybrid model there are some cases of failure observed when the temperature drops rapidly after the heatwave with the prediction error rate rising by about 15–20%. Such instances usually coincide with the unexpected dust storms or the unexpected changes in the wind patterns which are not reflected in the univariate input structure. Multivariate atmospheric variables should be integrated in future work to overcome such shortcomings.

A computational requirement (training phase) is displayed in Table 10.

Table 10. Computational requirements (training phase)

Model	Training Time (mins)	Peak Random Access Memory (RAM) (GB)	GPU Utilization
ARIMA	2.1	1.2	None
LSTM	45.3	3.8	78%
Hybrid	48.7	4.1	82%

Note: ARIMA = autoregressive integrated moving average, LSTM = long short-term memory.

5.5 Practical Implementation Framework and Recommendations

The model suggested allows applications to apply several practical applications that directly enable urban resilience and climate adaptation approaches. In the case of early warning systems, the structure will offer early warnings of extreme heat events 48 hours ahead with probability-based notifications adjusted to various temperature limits. This is an ability to offer sector specific warnings that would be used in energy management, communal health interventions, and safeguarding against the agricultural sector. In urban planning integration, the model is used in the cooling demand forecasting of energy grid optimization, the public health preparedness to heat-related illness prevention, and the infrastructure resilience planning using dependable temperature projections. The modular nature of the framework enables easy integration with already in place meteorological surveillance systems and the computational efficiency of the framework makes it feasible to implement it where resources are limited.

To extend the result of this research, a few suggestions are put forward. To be generalized to other climate zones, i.e., coastal, tropical, or continental area, the model structure is to be re-calibrated by utilization of local historical data and confirmed by methods of cross-regional transferability. Second, other meteorological factors, such as humidity, solar radiation, as well as wind speed, should be included in the future modification of the hybrid model as these parameters influence the process of temperature changes and can increase the degree of predictive scale. Additionally, the use of ensemble methods should be considered to describe the model uncertainty to enhance its sturdiness in response to climate variability. Finally, it would be possible to create an online data pipeline that will automatically apply preprocessing and retraining steps to make the model fit new patterns that will appear in the context of changing climates.

6 Conclusion

The research has proposed and tested a hybrid forecasting model capable of combining the linear capabilities of the ARIMA framework with the non-linear learning capabilities of LSTM networks when it comes to forecasting extreme temperature events in Baghdad. With its intensive training, the hybrid ARIMA-LSTM model was regularly substantially better than the ARIMA and LSTM methods based on all assessment criteria, such as, MAE, RMSE, MAPE and R^2 . Important to note, it proved to have better predictive superiority under high-temperature conditions, which is above 45 °C, and this sure confirms its success in simulating the dynamics of complex extreme climatic behavior in an arid environment. The use of residual-based learning model combined with sequential modeling in the suggested methodology allowed not only depicting long-term changes of temperature following the trend-seasonal pattern, but also abrupt, sharp rises in temperature, which provided better forecasting results and increased reliability. The flexibility of the model to the climatic situation of the city of Baghdad indicates its potential use in the areas of public health preparedness, resilience, and early warning systems. In the future, further studies are recommended to build the model to a multivariate forecasting model where other climatic parameters (the level of humidity, wind speed, and solar radiation) are included. Moreover, temporal context modeling should also explore more sophisticated models of deep learning like Transformer models to lower the training time. Lastly, to cover practical applications, there must be work towards the implementation of the model in real-time by combining it with streaming meteorological data systems and automatic training models to accommodate the present conditions of climate variation.

Author Contributions

Conceptualization, A.N.A.; methodology, A.N.A. and N.J.A.; validation, B.K.I. and H.A.M.; formal analysis, B.K.I.; investigation, H.A.M.; resources, H.A.M.; data curation, B.K.I.; writing—original draft preparation, A.N.A. and B.K.I.; writing—review and editing, A.N.A., H.A.M., and N.J.A.; visualization, B.K.I. and N.J.A.; supervision, A.N.A.; project administration, A.N.A. All authors were actively involved in discussing the findings and refining the final manuscript.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] M. W. Hasan, “Building an IoT temperature and humidity forecasting model based on long short-term memory (LSTM) with improved whale optimization algorithm,” *Memor. Mater. Devices Circuits Syst.*, vol. 6, p. 100086, 2023. <https://doi.org/10.1016/j.memori.2023.100086>
- [2] M. Tanhapour, J. Soltani, H. Shakibian, B. Malekmohammadi, K. Hlavcova, S. Kohnova, and P. Valent, “The enhanced integration of proven techniques to quantify the uncertainty of forecasting extreme flood events based on numerical weather prediction models,” *Weather Clim. Extrem.*, vol. 48, p. 100767, 2025. <https://doi.org/10.1016/j.wace.2025.100767>
- [3] S. Salcedo-Sanz, J. Pérez-Aracil, G. Ascenso, J. D. Ser, D. Casillas-Pérez, C. Kadow, D. Fister, D. Barriopedro, R. García-Herrera, M. Giuliani, and A. Castelletti, “Analysis, characterization, prediction, and attribution of extreme atmospheric events with machine learning and deep learning techniques: A review,” *Theor. Appl. Climatol.*, vol. 155, no. 1, pp. 1–44, 2023. <https://doi.org/10.1007/s00704-023-04571-5>
- [4] M. Elseidi, “A hybrid Facebook Prophet-ARIMA framework for forecasting high-frequency temperature data,” *Model. Earth Syst. Environ.*, vol. 10, no. 2, pp. 1855–1867, 2023. <https://doi.org/10.1007/s40808-023-01874-4>
- [5] M. M. U. Qureshi, A. B. Ahmed, A. Dulmini, M. M. H. Khan, and R. Rois, “Developing a seasonal-adjusted machine-learning-based hybrid time-series model to forecast heatwave warning,” *Sci. Rep.*, vol. 15, no. 1, 2025. <https://doi.org/10.1038/s41598-025-93227-7>
- [6] E. Chiew and S. S. Choong, “A solution for M5 forecasting-uncertainty: Hybrid gradient boosting and autoregressive recurrent neural network for quantile estimation,” *Int. J. Forecast.*, vol. 38, no. 4, pp. 1442–1447, 2022. <https://doi.org/10.1016/j.ijforecast.2022.01.009>
- [7] S. H. Li, “Impact of climate change on wind energy across North America under climate change scenario RCP8.5,” *Atmos. Res.*, vol. 288, p. 106722, 2023. <https://doi.org/10.1016/j.atmosres.2023.106722>
- [8] Z. Ma, H. Zhang, and J. Liu, “MS-LSTM: Exploring spatiotemporal multiscale representations in video prediction domain,” *Appl. Soft Comput.*, vol. 147, p. 110731, 2023. <https://doi.org/10.1016/j.asoc.2023.110731>

- [9] K. Zouaidia, M. S. Rais, and S. Ghanemi, "Weather forecasting based on hybrid decomposition methods and adaptive deep learning strategy," *Neural Comput. Appl.*, vol. 35, no. 15, pp. 11 109–11 124, 2023. <https://doi.org/10.1007/s00521-023-08288-4>
- [10] S. Zhang, R. Chen, J. Cao, and J. Tan, "A CNN and LSTM-based multi-task learning architecture for short and medium-term electricity load forecasting," *Electr. Power Syst. Res.*, vol. 222, p. 109507, 2023. <https://doi.org/10.1016/j.epsr.2023.109507>
- [11] W. Ding, M. Abdel-Basset, and R. Mohamed, "HAR-DeepConvLG: Hybrid deep learning-based model for human activity recognition in IoT applications," *Inf. Sci.*, vol. 646, p. 119394, 2023. <https://doi.org/10.1016/j.ins.2023.119394>
- [12] A. Karnik, V. Gaurav, A. Mitra, R. Keshri, and P. Kotsampopoulos, "A hybrid LSTM-attention residual network for photovoltaic forecasting in isolated microgrid environments," *Smart Grids Sustain. Energy*, vol. 10, no. 3, 2025. <https://doi.org/10.1007/s40866-025-00298-6>
- [13] B. Jyostna, A. Meena, S. Rathod, M. D. Tuti, K. Choudhary, A. Lama, A. T. Kumar, B. N. K. Reddy, D. Bhanusree, J. Rakesh, A. K. Swarnaraj, and A. Kumar, "Multiscale rainfall forecasting using a hybrid ensemble empirical mode decomposition and LSTM model," *Model. Earth Syst. Environ.*, vol. 11, no. 2, 2025. <https://doi.org/10.1007/s40808-024-02220-y>
- [14] D. A. Tuan, "Causal and spatiotemporal deep learning for dengue forecasting and extreme outbreak risk under climate variability: A framework from Vietnam," *Int. J. Biometeorol.*, vol. 70, no. 4, 2026. <https://doi.org/10.1007/s00484-026-03151-2>
- [15] R. I. Fattah, P. P. Adikara, and B. D. Setiawan, "Mesin catur berbasis neural network menggunakan long short term memory (LSTM)," *J. Teknol. Inf. Ilmu Komput.*, vol. 12, no. 3, pp. 491–496, 2025. <https://doi.org/10.25126/jtiik.2025129445>
- [16] R. H. Shumway and D. S. Stoffer, "ARIMA models," in *Time Series Analysis and Its Applications*. Springer Nature Link, 2006, pp. 84–173. https://doi.org/10.1007/0-387-36276-2_3
- [17] D. Xu, Q. Zhang, Y. Ding, and D. Zhang, "Application of a hybrid ARIMA-LSTM model based on the SPEI for drought forecasting," *Environ. Sci. Pollut. Res.*, vol. 29, no. 3, pp. 4128–4144, 2021. <https://doi.org/10.1007/s11356-021-15325-z>
- [18] X. Huang, X. Zhuang, F. Tian, Z. Niu, Y. Chen, Q. Zhou, and C. Yuan, "A hybrid ARIMA-LSTM-XGBoost model with linear regression stacking for transformer oil temperature prediction," *Energies*, vol. 18, no. 6, p. 1432, 2025. <https://doi.org/10.3390/en18061432>
- [19] J. Shen, W. Wu, and Q. Xu, "Accurate prediction of temperature indicators in Eastern China using a multi-scale CNN-LSTM-attention model," *arXiv:2412.07997*, 2024. <https://doi.org/10.48550/arXiv.2412.07997>
- [20] Y. Liu, T. Hu, H. Zhang, H. Wu, S. Wang, L. Ma, and M. Long, "iTransformer: Inverted transformers are effective for time series forecasting," *arXiv:2310.06625*, 2023. <https://doi.org/10.48550/arXiv.2310.06625>
- [21] Z. Gao, X. Shi, H. Wang, Y. Zhu, Y. Wang, M. Li, and D. Y. Yeung, "Earthformer: Exploring space-time transformers for Earth system forecasting," *arXiv:2207.05833*, 2022. <https://doi.org/10.48550/arXiv.2207.05833>
- [22] G. Woo, C. Liu, D. Sahoo, A. Kumar, and S. Hoi, "ETSFormer: Exponential smoothing transformers for time series forecasting," *arXiv:2202.01381*, 2022. <https://doi.org/10.48550/arXiv.2202.01381>
- [23] J. K. Mutinda, A. K. Langat, and S. M. Mwalili, "Forecasting temperature time series data using combined statistical and deep learning methods: A case study of Nairobi County daily temperature," *Int. J. Math. Math. Sci.*, vol. 2025, no. 1, p. 4795841, 2025. <https://doi.org/10.1155/ijmm/4795841>
- [24] M. Hendel, I. S. Bousmaha, F. Meghnefi, I. Fofana, and M. Brahami, "An intelligent power transformers diagnostic system based on hierarchical radial basis functions improved by Linde Buzo Gray and single-layer perceptron algorithms," *Energies*, vol. 17, no. 13, p. 3171, 2024. <https://doi.org/10.3390/en17133171>