

---

# A Comparative Analysis of Statistical and Deep Learning Models for Global Temperature Anomalies Forecasting

---

Maryam Ibrahim Habadi\*, Shumukh AL-qahtani,  
Hadil Ibrahim Hariry and Mona Alshehri

*Department of Statistics, Faculty of Science, King Abdulaziz University, Jeddah  
21589, Saudi Arabia*

*E-mail: mhabadi@kau.edu.sa; mswd96775@gmail.com; hadil.hariry@gmail.com;  
Munaalshehri7@gmail.com*

*\*Corresponding Author*

Received 11 August 2025; Accepted 14 September 2025

## Abstract

Global warming is among the most pressing environmental challenges, mainly driven by human-induced greenhouse gas emissions. Accurate forecasting of global temperature anomalies is essential for understanding climate trends and planning effective interventions. This study utilizes historical temperature anomaly data from 1940 to 2023. Aiming to compare the forecasting performance of several statistical and machine learning models: Seasonal Autoregressive Integrated Moving Average, Triple Exponential Smoothing, Temporal Convolutional Networks, Long Short-Term Memory, and two hybrid models, SARIMA-LSTM and SARIMA-TCN. Forecast accuracy was evaluated using Mean Squared Error, Mean Absolute Error, and Root Mean Squared Error. The TCN model demonstrated superior forecasting performance, achieving the lowest error across all metrics, followed by the SARIMA-LSTM hybrid model. The results support the combination of

*Journal of Reliability and Statistical Studies, Vol. 18, Issue 2 (2025), 371–398.*

doi: 10.13052/jrss0974-8024.1825

© 2025 River Publishers

statistical and deep learning models for improved climate forecasting and offer valuable insights into future temperature trends amid global warming.

**Keywords:** SARIMA, machine learning, hybrid model, time series forecasting, temperature..

## 1 Introduction

Global warming is a critical environmental phenomenon characterized by a sustained increase in the Earth's average temperature. This trend, primarily driven by increased greenhouse gas emissions from human activity, poses severe risks to ecosystems, economies, and human health. Understanding and forecasting temperature anomalies is essential for assessing climate change's long-term impacts and supporting climate-related decision-making.

Traditional statistical models, such as the Seasonal Autoregressive Integrated Moving Average (SARIMA) and Triple Exponential Smoothing (TES), have been widely used in time series forecasting. However, with the rise of machine learning, more advanced models, such as Long Short-Term Memory (LSTM) and Temporal Convolutional Networks (TCN), have shown promise in capturing complex and nonlinear temporal dependencies.

In line with the growing literature that applies both statistical and deep learning models to climate forecasting, this study extends the discussion by offering a direct comparative analysis across multiple approaches.

This study comprehensively compares statistical and deep learning models, including the development of two hybrid models, SARIMA-LSTM and SARIMA-TCN, to forecast global temperature anomalies based on monthly data from 1940 to 2023. The primary aim is to identify the most accurate approach and to assess the potential of hybrid modeling strategies for improving climate forecasting. The study focuses on global temperature anomalies spanning 84 years, providing a clear context for evaluating patterns and trends in long-term climate behavior.

### 1.1 Literature Review

Several studies have investigated global temperature trends and developed statistical and machine learning forecasting techniques. For example, [1] aimed to analyze a five-time series to identify trends in global temperatures recorded from 1979 to 2010, including records of surface temperatures and lower atmospheric layer temperatures, while accounting for external

influences such as volcanic eruptions and solar activity. They employed multiple regression analysis to isolate the short-term effects of these phenomena on global warming. In addition, they conducted an annual cycle analysis to examine how variations in the annual cycle influence the calculation of thermal anomalies. The analysis indicated that the studied factors, such as solar changes, volcanic eruptions, and El Niño, significantly impact short-term changes, and their damage will be less. Overall, the rate of global warming remained steady throughout the study period, suggesting that human-induced warming is persistent and likely to continue for several decades.

[2] studied global temperature changes from 1970 to 2016 to determine whether there had been a slowdown or acceleration in the rate of global warming. With added white noise, they used unpaired Monte Carlo simulations to generate 10,000 synthetic time series based on the linear trend and standard deviation of the 1972–2000 reference period. These simulations aimed to assess whether the apparent slowdown in warming from 2001 to 2014 was due to natural variability or a genuine change in the underlying trend. The analysis, which compared the periods 1979–2000 and 2001–2014, found no statistically significant change in the rate of warming. However, the period from 2014 to 2016 showed a notable and steady increase in global temperatures, reinforcing the evidence of ongoing global warming.

Regarding machine learning approaches, [3] aimed to develop an accurate weather forecasting system using temporal convolutional networks (TCNs) and deep learning techniques based on analyzing time series data collected from six local meteorological stations. The data were recorded at 15-minute intervals and included various climatic variables to improve forecasting accuracy and support farmers and local communities. The study compared the performance of the TCN model with that of Long Short-Term Memory (LSTM) networks and several traditional regression methods to identify the most effective model. The results indicated that the TCN model outperformed the others in terms of forecasting accuracy and computational efficiency.

[4] aimed to evaluate the effects of global warming by analyzing global temperature anomalies and their temporal patterns. The researchers collected data from 1881 to 2020 to assess land and ocean surface temperature changes. They employed the Mann-Kendall test to detect trends in temperature. Specifically, whether they were increasing or decreasing over time. In addition, Sen's slope estimation method was used to quantify the temperature change rate over time, providing insight into the pace and consistency of global warming. The findings indicated a clear upward trend in temperature, attributed primarily to human activities, accompanied by rising sea levels and

melting ice. The study confirmed the reality of climate change, showing that the observed anomalies are strongly linked to anthropogenic factors rather than natural variability alone.

[5] conducted a time series analysis spanning the period from 1989 to 2021 to investigate the potential impacts of climate change on renewable energy sources, with a specific focus on monthly global temperature and wind speed data. The researchers applied the Mann-Kendall test to identify trends in temperature and wind speed. In addition, the Theil-Sen slope estimator was used to quantify the rate and direction of these changes. The study revealed significant monthly variations in both temperature and wind speed. Furthermore, it was found that the average global temperature increased by approximately  $0.34^{\circ}\text{C}$  every decade. These results suggest that climate change has a significant impact on wind patterns, which could, in turn, affect renewable energy production.

From a forecasting perspective, [6] analyzed global warming trends and forecasted future temperature anomalies using a statistical ARIMA model and the random walk with drift model. The study utilized historical climate data spanning from 1850 to 2021. Among the models tested, the ARIMA (1, 1, 3) model was identified as the most suitable, projecting a temperature anomaly of  $1.56^{\circ}\text{C}$  by the year 2050. This projection significantly exceeds the United Nations' target range of  $0.4\text{--}0.9^{\circ}\text{C}$ . The findings indicate that human activity is the primary driver of this increase. If current trends continue, severe and long-lasting environmental consequences will likely highlight the urgent need for coordinated international action.

Similarly, [7] utilized NASA GISTEMP data from 1880 to the present to forecast anomalies in global temperatures. The study employed two widely used time-series forecasting models, ARIMA and ETS (Error, Trend, Seasonal), to compare their effectiveness in forecasting future climate changes. By decomposing the time series into error, trend, and seasonal components, the study effectively captured long-term shifts and cyclical variations, which can aid in identifying climate trends. The results indicated that the ARIMA model is more reliable and accurate for long-term global temperature forecasting. In contrast, the ETS model is better suited for short-term predictions or datasets with significant seasonal variability.

Hybrid models have attracted attention in climate forecasting, as they combine the strengths of statistical methods in capturing linear patterns with deep learning models' ability in handling nonlinear patterns. Several

studies established the effectiveness of such approaches. For instance, [8] introduced an ARIMA–ML hybrid approach that combined ARIMA with gradient boosting for global temperature anomaly forecasting based on National Oceanic and Atmospheric Administration (NOAA) data and key ocean–atmosphere indices. The hybrid model achieved superior accuracy ( $R^2 = 0.8467$ , RMSE = 0.2697), confirmed through statistical validation, while Shapley Additive Explanations (SHAP) analysis underscored its interpretability by highlighting the role of major climate indices.

More recently, [9] examines the application of hybrid time series models to enhance the forecasting accuracy of global methane emissions. Integrating traditional statistical models, such as ARIMA, with machine learning methods, including Artificial Neural Networks (ANN) and Long Short-Term Memory (LSTM) networks. The results show that hybrid models significantly outperform single statistical time series and machine learning approaches, especially in capturing both linear and nonlinear trends.

In addition, [10] introduced a hybrid model that integrates the ARIMA model, the LSTM network, and Seasonal-Trend decomposition using Loess (STL). The study employed daily maximum temperature data from Rajshahi, and the results revealed that the hybrid model outperformed traditional statistical models, machine learning approaches, and other hybrid models by achieving the lowest error metrics. Moreover, it provided accurate forecasts for the period 2023–2026. The findings highlight that hybrid models, which combine deep learning and statistical methods, can be particularly effective for climate early warning systems and risk management.

In a recent study [11], hybrid models were developed for daily temperature data in Nairobi County, Kenya, by combining variational mode decomposition (VMD) with statistical (ARIMA) and deep learning (GRU, LSTM, Transformer) techniques. The results showed that hybrid models, particularly the VMD–ARIMA–GRU combination, achieved the lowest error metrics and high explanatory power ( $R^2 = 0.779$ ), highlighting the effectiveness of preprocessing techniques like VMD in improving feature representation and forecasting accuracy.

All reviewed studies consistently confirm that global temperatures have steadily increased in recent decades due to human activities. No evidence indicates a decline or stabilization in the overall warming trend. These findings provide strong and unequivocal support for the reality of global warming and underscore the urgent need for coordinated international action to mitigate its effects.

## 2 Data Description

Global Temperature Anomalies data was obtained from our World in Data website from 1940 to 2023 [12]. Table 1 shows that the minimum temperature anomaly was  $-1.06003$ , while the maximum temperature anomaly was  $0.93061$ . To further understand the distribution of the data, we note that 25% of the data was less than  $-0.63721$  in the first quartile. In the third quartile, 75% of the data was less than  $-0.33896$ . For the median, the results showed that 50% of the data was less than  $-0.43119$ .

**Table 1** Descriptive statistics for the global temperature anomalies.

Measures	min	Q1	Median	Mean	Q3	Max	Stander deviation
Value	$-1.0600$	$-0.6372$	$-0.4312$	$-0.3389$	$-0.0527$	$0.9306$	$0.3777$

## 3 Materials and Methods

Two categories of models, statistical and machine learning, were applied to forecast global temperature anomalies. We split the dataset into 80% training and 20% testing to evaluate forecasting performance. The training set was used solely for model development, while the testing set was for accurate assessment.

Seasonal Autoregressive Integrated Moving Average (SARIMA) and Triple Exponential Smoothing (TES) models were implemented using R (version 4.5.0). In contrast, the machine learning models, Long Short-term Memory (LSTM) and Temporal Convolutional Networks (TCN), and the hybrid models (SARIMA-LSTM and SARIMA-TCN) were developed using Python in the Google Colab environment. The Augmented Dickey-Fuller (ADF) test was used to examine the stationarity of the time series and the Kruskal Wallis to detect seasonality. All statistical tests were adopted at a significant level of 5%.

### 3.1 Triple Exponential Smoothing

Triple exponential smoothing (TES), also known as the Holt-Winters method, relies on forecasting future values found on historical data and is effective for analyzing data that contains trends and seasonality. It was developed as an extension of the double exponential smoothing method and the simple exponential smoothing method. It works on three components of the time

series: the level, which means the rate of change over time; the trend, which means the general trend of the time series; and seasonality, which is changes that occur periodically at regular intervals. There are two types of TES. The first is the additive model, which is used when the seasonal effect is constant over time, meaning that it is independent of the general trend. The research approves of this model. The second is the multiplicative model, which is employed when the seasonal effect changes with the level of the time series. The updated values of level, trend, and seasonality at each time point  $t$  are calculated using the following equations:

Level Update:

$$L_t = \alpha(Y_t - S_{t-m}) + (1 - \alpha)(L_{t-1} + T_{t-1}) \quad (1)$$

Trend Update:

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1} \quad (2)$$

Seasonality Update:

$$S_t = \gamma(Y_t - L_t) + (1 - \gamma)S_{t-m} \quad (3)$$

Forecast Calculation:

$$\hat{Y} = L_t + hT_t + S_{t+h-m(k+1)}, \quad (4)$$

where  $Y_t$  represents the actual observed value at time  $t$ ,  $L_t$  is the smoothed level of the time series. The trend component, denoted as  $T_t$ , captures the long-term direction of the series. The seasonal component  $S_t$ , accounts for periodic variations at time  $t$ ,  $m$  represents the number of periods in a complete seasonal cycle, while  $h$  is the forecast horizon, indicating the number of periods into the future forecast. The smoothing parameters  $\alpha$ ,  $\beta$ , and  $\gamma$  control the influence of past observations on the level, trend, and seasonality, respectively, with values between 0 and 1. Lastly,  $k$  represents the number of complete seasons between  $t$  and  $t + h$  [13].

### **3.2 Seasonal Autoregressive Integrated Moving Average**

Seasonal Autoregressive Integrated Moving Average (SARIMA) is a popular and flexible time series forecasting model. It is an extension of the non-seasonal ARIMA model, developed to handle data with seasonal patterns. It combines the autoregressive (AR), integrated (I), and moving average (MA) components of ARIMA along with seasonal terms.

The autoregressive component models the correlation between the current data point and past values. The integration denotes differencing, transforming non-stationary data into stationary data, as stationarity is essential for time series modeling. The moving average component models the relationship between past forecast errors and the current data point. The SARIMA model can be represented as SARIMA(p, d, q)(P, D, Q)<sub>s</sub> where p is the order of non-seasonal autoregressive, q is the order of non-seasonal moving average, I(d) integrated component of order d, P the order of seasonal autoregressive, Q the order of seasonal moving average component, I(D) seasonal integrated component of order D, and s is seasonal period [14]. Mathematically, SARIMA is described as follows:

$$(1 - \phi_1 B)(1 - \Phi_1 B^s)(1 - B)(1 - B^s)y_t = (1 + \theta_1 B)(1 + \Theta_1 B^s)\varepsilon_t, \quad (5)$$

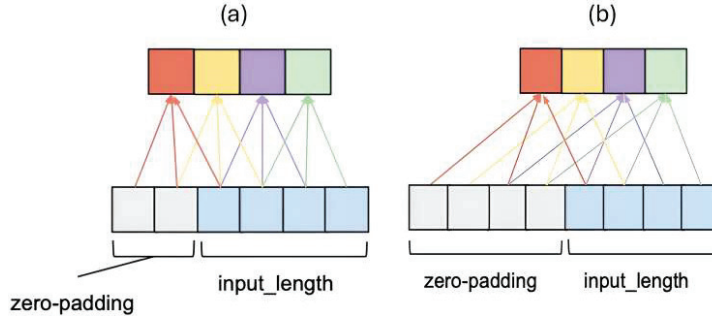
where  $y_t$  is the observed time series at time  $t$ ,  $B$  is the backward shift (or lag) operator,  $\phi$  is the non-seasonal autoregressive coefficient,  $\Phi$  is the seasonal autoregressive coefficient,  $\theta_1$  is the non-seasonal moving average coefficient,  $\Theta$  is the seasonal moving average coefficient,  $s$  is the seasonal period,  $\varepsilon_t$  is the white noise error term at time  $t$ .

## 4 Machine Learning Models

Machine learning models are creative instruments that analyze data and find patterns to forecast outcomes or make educated decisions based on the information they acquire from past data. Two types of machine learning models were used in this research: Temporal Convolutional Networks and Long Short-Term Memory Model.

### 4.1 Temporal Convolutional Networks

Deep neural networks designed to process temporal data are called temporal convolutional networks (TCN). Their design relies on convolutions instead of repetitive units, as in recurrent neural networks. This approach is helpful for parallel computation and achieves higher training stability. The TCN model uses convolutional layers to process temporal data. First, it applies a one-dimensional convolutional network that replaces repetitive structures, which helps improve training efficiency. Second, causal convolutions ensure that current inputs do not influence future values during training. Third, the model uses divergent convolutions to capture long-term dependencies



**Figure 1** Illustration of (a) causal convolutions and (b) dilated convolutions.

without increasing the network depth. Additionally, it uses residual connections to improve the flow of gradients across layers, which helps train deep networks and enhances model stability. Finally, the convolutional structure is designed to handle variable-length data, allowing the model to work flexibly with different sequence lengths [15].

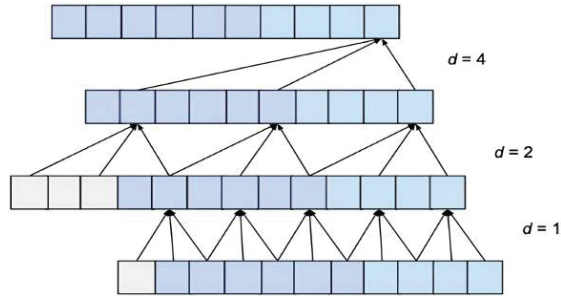
Figure 1 shows the causal convolution and dilated causal convolution mechanisms, which are the building blocks of TCN models. In part (a), causal convolution demonstrates how to rely only on the current and past values of the time series without considering the future, ensuring that the temporal sequence is respected and preserving the causality of the model. In part (b), dilated convolution demonstrates how inserting time intervals between input values expands the receptive field without increasing model parameters. This mechanism allows the model to capture long-range dependencies in time efficiently, a key advantage in analyzing complex time series data. The equation for causal convolution with dilation in a TCN is given by:

$$y(t) = \sum_{i=0}^{k-1} f(i) \cdot x(t - d \cdot i), \tag{6}$$

where  $y(t)$  is the output at time step  $t$ ,  $f(i)$  is a learned filter of size  $k$ , and  $x(t - d \cdot i)$  is the input at time step  $t$  with dilation factor  $d$ . The receptive field of the model is determined by:

$$W = 1 + (K - 1) \sum_{i=0}^{L-1} d_i, \tag{7}$$

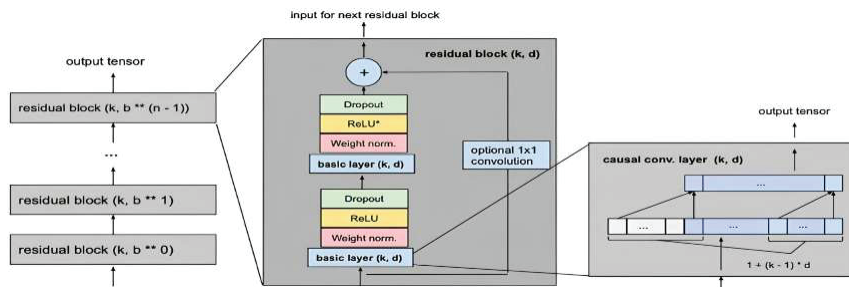
where  $W$  is the receptive field width,  $L$  is the number of layers,  $d_i$  is the dilation rate for each layer.



**Figure 2** Illustration of the receptive field.

Figure 2 illustrates how the receptive field expands using dilated convolutions. As the figure shows, increasing the divergence rates across layers enables the model to include a broader range of input priors without extending the network depth or number of parameters, enhancing the model’s ability to capture long-range dependencies effectively.

Forecasting with the TCN model is trained using (input, target) pairs, where the target is a shifted version of the input sequence. A sliding window approach generates multiple overlapping training samples from the same time series, ensuring the model learns to forecast future values based only on past observations. Figure 3 shows the complete architecture of the TCN model. The model begins with input data, which is passed into a series of modules called residual blocks. Each block consists of two layers of divergent causal convolutions, followed by activation and projection to enhance model accuracy and reduce overlearning. The model relies on intra-convolutional expansion, which helps it understand long-term relationships in temporal data. Finally, the outputs from the blocks are passed into final layers that are used to obtain the desired predictions [16].



**Figure 3** Illustration of the final TCN model.

### 4.2 Long Short-term Memory (LSTM)

The Long Short-Term Memory Model (LSTM) is a recurrent neural network (RNN) characterized by its ability to handle long-term relationships within time series data. It is effectively used in forecasting time series, recognizing speech, and detecting anomalies. Its design helps memorize information within the so-called memory cell for long periods. At each step, the state of the cell is updated, and the forecast depends on this condition. As shown in Figure 4, the memory cell consists of three main gates: the input gate, the forgetting gate, and the output gate, each of which controls how data enters or exits memory [17].

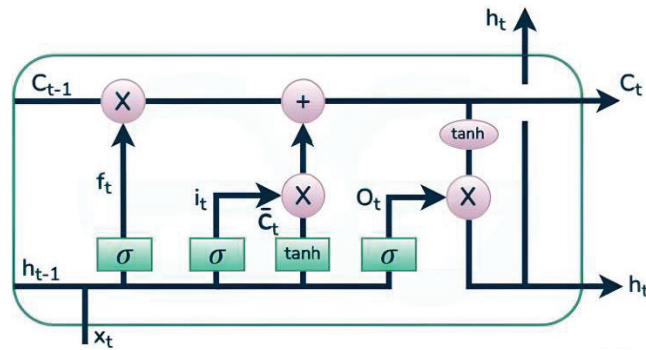


Figure 4 Long Short-Term Memory (LSTM) cell architecture.

The input gate, which has the following equation, controls the data that is added to the memory cell:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i), \tag{8}$$

where  $W_i$  refers to the weight matrices used in the input gate. The term  $[h_{t-1}, x_t]$  combines the previous hidden state and the current input.  $b_i$  stands for the bias vectors associated with the input gate. The function  $\sigma$  is the sigmoid activation, defined as  $\frac{1}{1+e^{-x}}$ . The following equation illustrates how the forgetting gate filters out unnecessary information from the cell's state.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f), \tag{9}$$

where the weight matrix  $W_f$  is associated with the forget gate. The notation  $[h_{t-1}, x_t]$  refers to the concatenation of the previously hidden  $h_{t-1}$  and the current input  $x_t$ . Additionally,  $b_f$  represents the bias term for the forget gate.

The sigmoid activation function  $\sigma$ . The output gate controls how information flows to the next cell state through a specific governing equation.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o), \quad (10)$$

where  $o_t$  is the weight matrix for the output gate and  $[h_{t-1}, x_t]$  refers to the combination of the previous hidden state and the current input.  $b_o$  is the bias term for the output gate, and  $\sigma$  is the sigmoid activation function. LSTM networks often have extra layers or cells below the cell state to improve data processing before forecasting the result. Preprocessing techniques are also crucial for optimizing data to facilitate successful learning and ensure that the model can effectively handle temporal dependencies and perform well across various applications [18].

### 4.3 Hybrid SARIMA-LSTM Model

The SARIMA-LSTM hybrid model combines the seasonal linear modeling capabilities of the Seasonal Autoregressive Integrated Moving Average (SARIMA) with the nonlinear learning capabilities of long short-term memory (LSTM) neural networks. The SARIMA model is first applied to the original time series to capture and model its seasonal and trend components. After training the SARIMA model, the residuals (the portion of the time series that SARIMA cannot explain) are extracted. These residuals, which contain, at best, the nonlinear structure, are then used to train an LSTM model. LSTMs are specifically used to learn the nonlinear dependencies that the SARIMA model fails to capture. Then, the final forecast is generated by combining the outputs of the two models [19].

$$\hat{y}_t = \hat{S}_t + \hat{N}_t, \quad (11)$$

where  $\hat{S}_t$  is the forecast from the SARIMA model (linear-seasonal component).  $\hat{N}_t$  is the forecast from the LSTM model (nonlinear component based on SARIMA residuals), and  $\hat{y}_t$  is the final hybrid forecast.

### 4.4 Hybrid SARIMA-TCN Model

The SARIMA model has been used to capture linear and seasonal patterns in climate data, but its ability to represent complex nonlinear relationships is limited. To address these limitations, hybrid models such as SARIMA-LSTM have been developed, combining the linear modelling of SARIMA with the nonlinear temporal dependencies learned by LSTMs. These hybrid models

have proven effective [19]. Temporal convolutional networks (TCNs) have also demonstrated high performance in time series modelling due to their ability to represent long dependencies. While both SARIMA and TCN have been successful individually, the combination of SARIMA modeling with TCN capacity for nonlinear temporal representation remains an unexplored direction, and this study seeks to investigate.

#### 4.5 Model Evaluation

Forecast accuracy measures are essential for assessing forecast accuracy. So, three scales were used: The Mean Absolute Error (MAE) shows the average size of the error, the Mean Squared Error (MSE) magnifies large errors, and the Root Mean Squared Error (RMSE) combines the two and gives a result in the original data unit that is easy to interpret. Using the three provides a more accurate picture of the model's performance.

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (12)$$

$$MAE = \frac{1}{n} \sum_{j=1}^n |y_i - \hat{y}_i| \quad (13)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (14)$$

The Akaike Information Criterion (AIC) is one of the basic tools used to choose the most appropriate forecast model, as it measures the quality of the model based on a balance between good data compatibility and the number of transactions used. The model that reduces the loss of information is preferred, that is, which gives the least value to AIC. It is defined as:

$$AIC = -2 \log(L) + 2k, \quad (15)$$

where  $L$  represents the likelihood of the model, and  $k$  denotes the number of parameters.

### 5 Results and Discussion

This section analyzes global temperature anomalies using statistical models, deep learning techniques, a hybrid models. It begins with a description of

the data, then presents an exploratory analysis, and then the application and evaluation of different models. Finally, the models are compared based on standard accuracy Metrics to evaluate their forecasting performance.

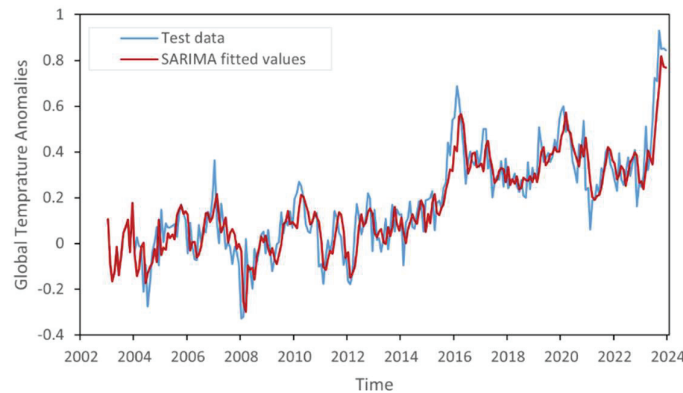
### 5.1 The SARIMA Model

Leading to the rejection of the null hypothesis that there is no seasonality. This result indicates that the data exhibit seasonal patterns. After differencing, the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF), respectively, show that the series is stationary, as most spikes fall within the significance range (blue lines). Moreover, the P-value of the Augmented Dickey-Fuller (ADF) test is 0.01, which means the series is stationary at a 95% significance level.

The best SARIMA model based on AIC is SARIMA (2,1,2) (0,1,1) [12], and it can be model equation is:

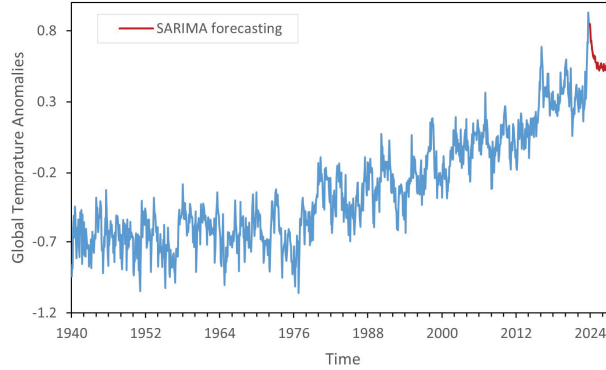
$$\begin{aligned} &(1 - 0.6204B - 0.2024B^2)(1 - B)(1 - B^{12})y_t \\ &= (1 + 1.1383B - 0.1499B^2)(1 - 0.9283B^{12})\varepsilon_t \end{aligned}$$

Figure 5 shows that the SARIMA model closely follows the actual trend of temperature anomalies, indicating reliable forecasting performance.



**Figure 5** Fitted and actual data of SARIMA model.

The forecast, as shown in Figure 6 and Table 2, indicates that temperature anomalies are expected to continue increasing over the coming years, resulting in the planet remaining warmer than usual.



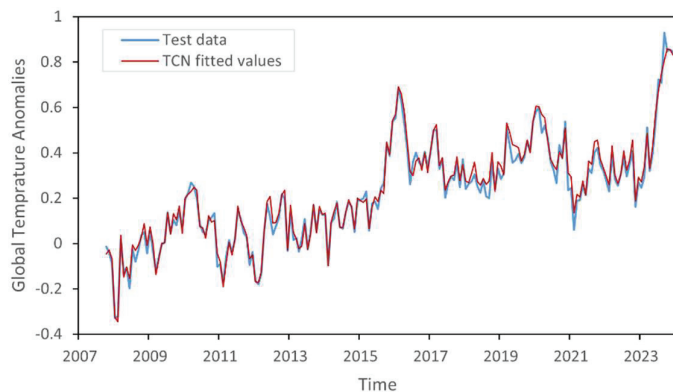
**Figure 6** Forecasting plot for the next three years of SARIMA model.

**Table 2** SARIMA model forecasts for the next three years

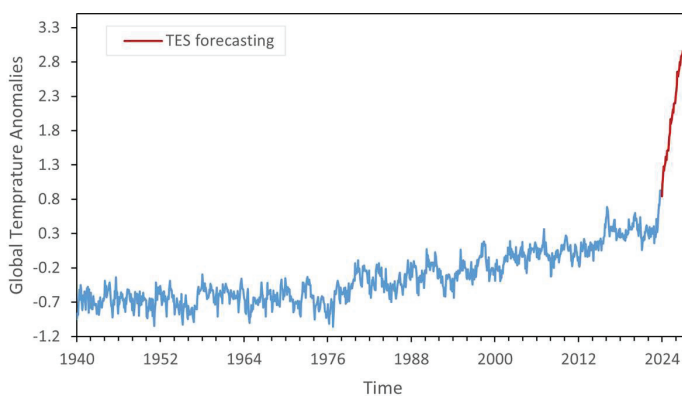
Month	Forecast 2024	Forecast 2025	Forecast 2026
January	0.77	0.56	0.54
February	0.71	0.53	0.52
March	0.73	0.57	0.57
April	0.66	0.54	0.54
May	0.65	0.54	0.55
June	0.61	0.52	0.53
July	0.62	0.55	0.56
August	0.59	0.54	0.55
September	0.62	0.57	0.58
October	0.61	0.58	0.59
November	0.57	0.55	0.56
December	0.57	0.55	0.57

### 5.2 Triple Exponential Smoothing (Holt-Winters)

The optimal TES model was selected with the following parameters ( $\alpha = 0.41$ ), which represents the degree of level smoothing; ( $\beta = 0.01$ ), which means the degree of trend smoothing; and ( $\gamma = 0.31$ ), which represents the degree of seasonal smoothing. They were determined using grid search, a method of trying all possible combinations of values to find the best combination that gives the lowest error.



**Figure 7** Fitted and actual data of TES model.



**Figure 8** Forecasting plot for the next three years of TES model.

Figure 7 shows that the predicted values follow the shape of the actual time series for most periods, indicating that the model captures both trend and seasonality, although there are some differences and slight variations, especially at sharp peaks or troughs.

Figure 8 and Table 3 illustrate that the TES model predicts a sharp rise in temperature anomalies beginning in 2024, suggesting an accelerated warming trend over the next three years.

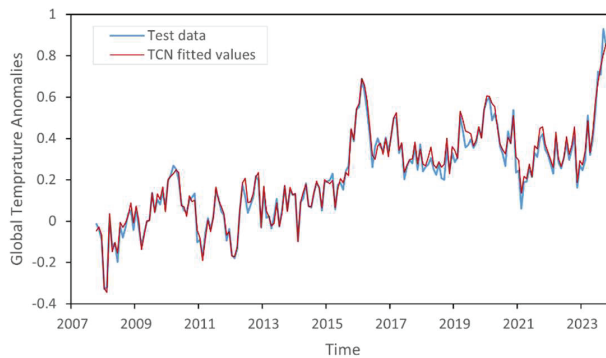
**Table 3** TES model forecasts for the next three years

Month	Forecast 2024	Forecast 2025	Forecast 2026
January	1.02	1.71	2.40
February	1.09	1.77	2.46
March	1.28	1.97	2.66
April	1.21	1.90	2.59
May	1.28	1.96	2.65
June	1.33	2.02	2.71
July	1.42	2.11	2.80
August	1.37	2.06	2.75
September	1.50	2.19	2.88
October	1.51	2.20	2.89
November	1.51	2.20	2.89
December	1.59	2.28	2.96

### 5.3 The TCN Model

The TCN model consists of an input layer that receives sequences of up to 36 months, followed by a single TCN layer designed to extract long-term temporal patterns. The output then passes through a dense layer containing a single cell with a tanh activation function. This activation function converts values to the range  $[-1, 1]$ , which helps stabilize learning and improve forecast accuracy in time series. This produces the predicted value of the temperature anomaly. The Adam optimizer was used to update the weights, and the model was trained over 108 epochs using an early stopping mechanism to maintain optimal performance on the validation data.

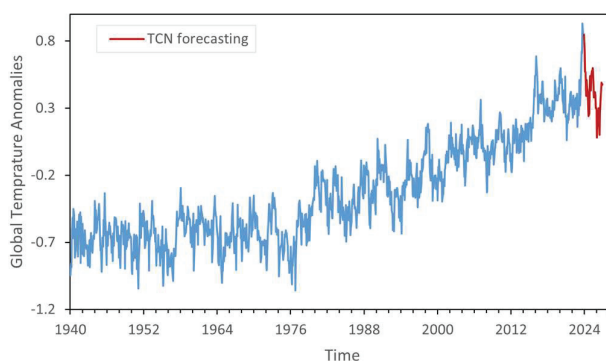
Figure 9 illustrates that the TCN model effectively predicts the actual behavior of the data. The strong alignment between the actual and forecast



**Figure 9** Fitted and actual data of TCN model.

**Table 4** TCN model forecasts for the next three years

Month	Forecast 2024	Forecast 2025	Forecast 2026
January	0.71	0.43	0.08
February	0.57	0.54	0.17
March	0.57	0.58	0.17
April	0.39	0.54	0.28
May	0.51	0.60	0.30
June	0.46	0.48	0.10
July	0.42	0.38	0.27
August	0.32	0.39	0.33
September	0.24	0.42	0.44
October	0.26	0.39	0.49
November	0.43	0.30	0.47
December	0.54	0.30	0.48

**Figure 10** Forecasting plot for the next three years of TCN model.

values demonstrates the model's ability to forecast reasonably accurately, making it a reliable tool for time series analysis. Table 4 and Figure 10 illustrate that the forecast values indicate a continued upward trend in temperature anomalies, accompanied by the expected seasonal fluctuations, suggesting that warming will persist in the coming years.

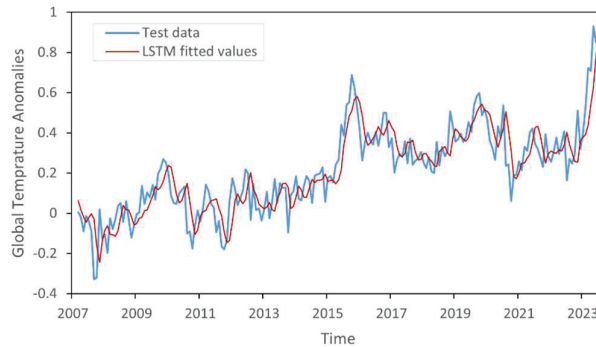
#### 5.4 The LSTM Model

The LSTM model consists of two consecutive layers, each containing 32 units. A 20% dropout is included after each layer to reduce the risk of overfitting. The outputs are then passed through a dense layer with 16 units

and a tanh activation function, followed by a single output layer that predicts the temperature anomaly value.

The Adam optimizer was used with a learning rate of 0.001, and the model was trained over 100 epochs using an early stopping mechanism to maintain optimal performance on the validation data. Figure 11 compares the actual and predicted temperature anomalies. The LSTM model demonstrates good forecasting performance, effectively capturing long-term trends and seasonal patterns in the data.

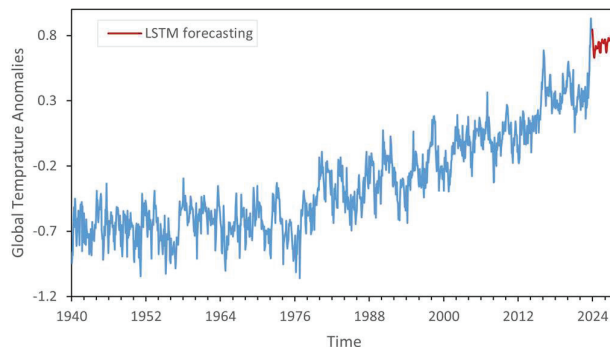
Table 5 displays the forecast values of the LSTM model, and Figure 12 illustrates that global temperature anomalies continue an upward trend, meaning temperatures remain above historical averages.



**Figure 11** Fitted and actual data of LSTM model.

**Table 5** LSTM model forecasts for the next three years

Month	Forecast 2024	Forecast 2025	Forecast 2026
January	0.79	0.75	0.77
February	0.68	0.68	0.69
March	0.64	0.67	0.67
April	0.63	0.67	0.68
May	0.69	0.74	0.75
June	0.69	0.75	0.75
July	0.72	0.77	0.78
August	0.71	0.76	0.77
September	0.70	0.74	0.76
October	0.70	0.74	0.76
November	0.71	0.74	0.76
December	0.75	0.77	0.80

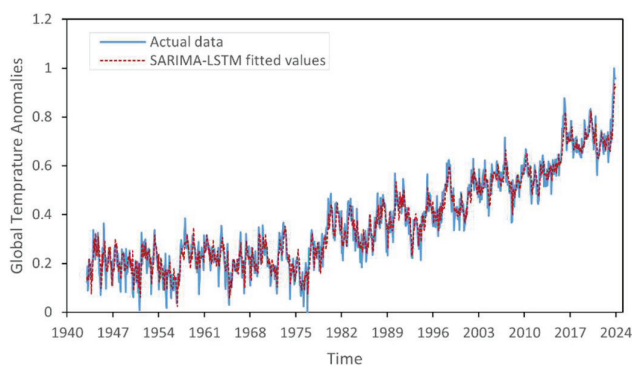


**Figure 12** Forecasting plot for the next three years of LSTM model.

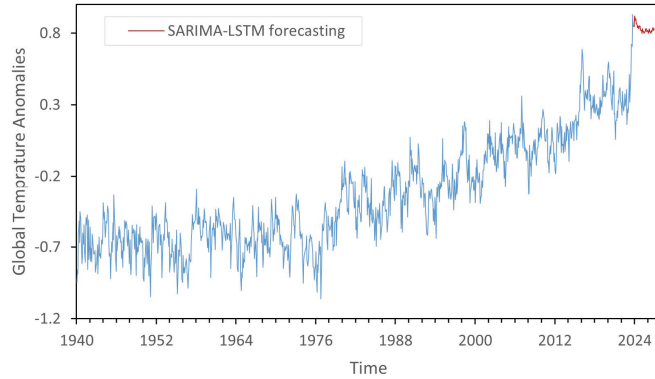
### 5.5 Hybrid SARIMA-LSTM Model

The hybrid SARIMA-LSTM model was constructed by first fitting the SARIMA (2, 1, 2) (0, 1, 1) [12] model to the time series data. The residuals from this model were then used as input to the same LSTM architecture described in Section 4.3.

Figure 13 shows that there is a strong alignment between the actual values and the model’s forecasting over most periods. It appears that the hybrid model was able to capture the general trends, seasonal fluctuations, and short-term changes with great accuracy. This result enhances the model’s confidence as an accurate tool for analyzing and forecasting time series in this case. Table 6 presents the forecast values of the hybrid SARIMA-LSTM model.



**Figure 13** Fitted and actual data of SARIMA-LSTM model.



**Figure 14** Forecasting plot for the next three years of SARIMA-LSTM model.

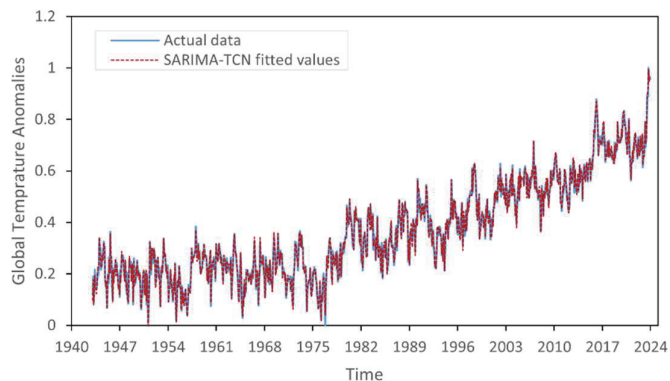
**Table 6** SARIMA-LSTM model forecasts for the next three years

Month	Forecast 2024	Forecast 2025	Forecast 2026
January	0.92	0.82	0.81
February	0.89	0.80	0.80
March	0.90	0.83	0.83
April	0.87	0.81	0.81
May	0.86	0.81	0.81
June	0.84	0.80	0.80
July	0.85	0.81	0.82
August	0.83	0.81	0.81
September	0.85	0.83	0.83
October	0.85	0.83	0.84
November	0.83	0.81	0.82
December	0.82	0.82	0.83

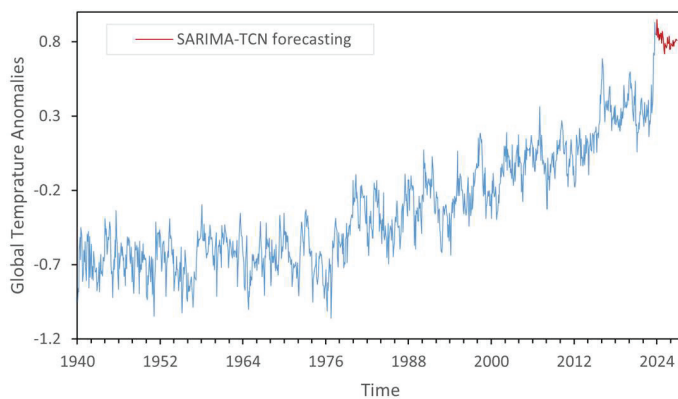
### 5.6 Hybrid SARIMA-TCN Model

The hybrid SARIMA-TCN model was built by fitting the SARIMA model  $(2, 1, 2) (0, 1, 1)$  [12] to time series data, aiming to capture linear and seasonal patterns. The residuals of the SARIMA model were then used as input to a TCN to capture the remaining nonlinear patterns and improve the final forecast accuracy.

Figure 15 shows that the hybrid model accurately predicts the actual behavior of the data, showing a close match between the actual values and



**Figure 15** Fitted and actual data of SARIMA-TCN model.



**Figure 16** Forecasting plot for the next three years of SARIMA-TCN model.

the predicted values across the entire period. This significant match reflects the model's ability to capture both the general trend and seasonality, making it a reliable model for time series analysis and forecasting in this context.

Figure 16 shows a continued upward trend in temperature anomalies over the coming years. Forecasts also indicate significant seasonal fluctuations. Table 7 presents the forecast values of the hybrid SARIMA-TCN model.

**Table 7** SARIMA-TCN model forecasts for the next three years

Month	Forecast 2024	Forecast 2025	Forecast 2026
January	0.95	0.74	0.79
February	0.83	0.72	0.78
March	0.89	0.79	0.78
April	0.81	0.77	0.77
May	0.85	0.76	0.81
June	0.82	0.78	0.77
July	0.86	0.83	0.80
August	0.86	0.79	0.80
September	0.79	0.79	0.81
October	0.87	0.84	0.82
November	0.84	0.75	0.81
December	0.82	0.75	0.81

### 5.7 Comparison of Forecasting Models

The forecasting performance of all models was compared using three standard error metrics: MSE, MAE, and RMSE. Table 8 summarizes the results for the test dataset.

Among all models, the TCN model has the lowest error values for all three metrics, making it the optimal choice for forecasting global temperature anomalies. The hybrid SARIMA-LSTM model ranked second in performance. On the other hand, the hybrid SARIMA-TCN model recorded the highest error rates, indicating the weakest predictive capability.

In addition to the standard error metrics, forecasted values were compared with the actual data reported for 2024 and the first four months of 2025.

The forecast values for 2024 to 2026 were generated using trained models and later compared with actual data that became available for 2024 and the

**Table 8** Performance evaluation of models

Models	MSE	MAE	RMSE
SARIMA	0.0105	0.0802	0.1032
TES	0.0155	0.0983	0.1244
TCN	0.0012	0.0248	0.0346
LSTM	0.0106	0.0796	0.1029
SARIMA-LSTM	0.0028	0.0416	0.0533
SARIMA-TCN	0.0185	0.1229	0.1361

**Table 9** Comparison of forecasted and newly observed values

Month	Forecast	Actual	Forecast	Actual	Forecast
	2024	Value	2025	Value	2026
January	0.71	0.70	0.43	0.79	0.08
February	0.57	0.81	0.54	0.63	0.17
March	0.57	0.73	0.58	0.65	0.17
April	0.39	0.67	0.54	0.60	0.28
May	0.51	0.65	0.60		0.30
June	0.46	0.67	0.48		0.10
July	0.42	0.68	0.38		0.27
August	0.32	0.71	0.39		0.33
September	0.24	0.73	0.42		0.44
October	0.26	0.80	0.39		0.49
November	0.43	0.73	0.30		0.47
December	0.54	0.76	0.30		0.48

first four months of 2025. The comparison, as shown in Table 9, indicates a reasonable alignment between forecasted and observed values, supporting the validity and robustness of the TCN forecasting model.

## 6 Conclusion

This study explored various statistical and deep learning models to forecast global temperature anomalies, including SARIMA, TES, TCN, LSTM, a SARIMA-LSTM hybrid, and a SARIMA-TCN hybrid model. The performance of each model was evaluated using three accuracy metrics: MSE, MAE, and RMSE. The findings showed that traditional models such as SARIMA effectively captured seasonal patterns and linear trends. In contrast, deep learning models such as TCN and LSTM demonstrated greater efficiency in handling complex nonlinear patterns in the data. These results highlight that the TCN model is the best model with the lowest values for all three metrics, supporting data-driven strategies for climate change adaptation efforts and developing more effective future strategies. To improve the accuracy of forecasting in future research, additional climate-related variables, such as ocean circulation indices or greenhouse gas concentrations, can be incorporated. Furthermore, expanding the analysis to regional or high-resolution datasets may provide more detailed insights into localized climate behavior.

## List of Symbols

ACF	Autocorrelation Function
ADF	Augmented Dickey-Fuller
LSTM	Long short-term memory
MAE	Mean Absolute Error
MSE	Mean square error
PACF	Partial Autocorrelation Function
Q-Q	Quantile-Quantile Plot
RMSE	Root Mean Square Error
RNN	Recurrent neural network
SARIMA	Seasonal autoregressive integrated moving average
TCN	Temporal Convolution Networks
TES	Triple Exponential Smoothing

## References

- [1] Foster, G., and S. Rahmstorf. 2011. Global temperature evolution 1979–2010. *Environmental Research Letters* 6: 044022.
- [2] Rahmstorf, S., G. Foster, and N. Cahill. 2017. Global temperature evolution: Recent trends and some pitfalls. *Environmental Research Letters* 12: 054001.
- [3] Hewage, P., A. Behera, M. Trovati, E. Pereira, M. Ghahremani, F. Palmieri, and Y. Liu. 2020. Temporal convolutional neural (TCN) network for an effective weather forecasting using time-series data from the local weather station. *Soft Computing* 24: 16453–16482.
- [4] Sadhukhan, B., S. Mukherjee, and R. K. Samanta. 2023. A study of global temperature anomalies and their changing trends due to global warming. In *Proceedings of the 2022 International Conference on Computational Intelligence and Communication Networks* 660–666.
- [5] Fei, Y., S. Leigang, and W. Juanle. 2023. Monthly variation and correlation analysis of global temperature and wind resources under climate change. *Energy Conversion and Management* 285: 1169925.
- [6] Wang, L. 2023. A century-long analysis of global warming and earth temperature using a random walk with drift approach. *Decision Analytics Journal* 7: 1–10.
- [7] Fan, C. 2024. Global temperature anomaly forecast: A comparative analysis of ARIMA and ETS models. In *Proceedings of the ICFTBA Workshop: Finance’s Role in the Just Transition*. 144: 70–76.

- [8] Nazlı, S. (2025). Enhancing global temperature anomaly forecasting with ARIMA–ML hybrid models with statistical validation. *SSRN*. <http://dx.doi.org/10.2139/ssrn.5261482>.
- [9] Habadi, M., M. Alshsehri, and I. Alsaggaf. 2025. Enhancing global methane emissions forecasting using hybrid time series models. *International Journal of Advanced Applied Sciences* 12: 34–41. doi:10.21833/ijaas.2025.04.005.
- [10] Qureshi, M. M. U., Ahmed, A. B., Dulmini, A., Khan, M. M. H., & Rois, R. (2025). Developing a seasonal-adjusted machine-learning-based hybrid time-series model to forecast heatwave warning. *Scientific Reports*, 15, 8699.
- [11] Mutinda, J. K., Langat, A. K., & Mwalili, S. M. (2025). Forecasting temperature time series data using combined statistical and deep learning methods: A case study of Nairobi County daily temperature. *International Journal of Mathematical Models and Methods in Applied Sciences*, 2025.
- [12] Ritchie, H. R. 2020. Temperature anomaly. *Our World in Data*. Accessed: May 10, 2025. <https://ourworldindata.org/temperature-anomaly>.
- [13] Hyndman RJ, Athanasopoulos G. 2018. Forecasting: Principles and Practice, 2nd ed. Otexts, Heathmont, Vic. <https://otexts.com/fpp2/>.
- [14] Chang, X., M. Gao, Y. Wang, and X. Hou. 2012. Seasonal autoregressive integrated moving average model for precipitation time series. *Journal of Mathematics and Statistics* 8: 500–505.
- [15] Bai, S., J. Z. Kolter, and V. Koltun. 2018. An empirical evaluation of generic convolutional and recurrent networks. *arXiv*. doi:arXiv:1803.01271v2.
- [16] Lea, C., M. D. Flynn, R. Vidal, A. Reiter, and G. D. Hager. 2016. Temporal convolutional networks for action segmentation and detection. *arXiv*. doi:arXiv:1608.05158.
- [17] Staudemeyer, R. C., and E. R. Morris. 2019. Understanding LSTM – A tutorial into long short-term memory recurrent neural networks. *arXiv*. <https://doi.org/10.48550/arxiv.1909.09586>.
- [18] Hochreiter, S., and J. Schmidhuber. 1997. Long short-term memory. *Neural Computation* 9: 1735–1780.
- [19] Adeyeye, J. S., and E. B. Nkemnole. 2023. Predicting malaria incident using hybrid SARIMA-LSTM model. *International Journal of Mathematical Sciences Optimization Theory and Applications* 9: 123–137.

## **Biographies**

**Maryam Ibrahim Habadi** is an Assistant Professor in the Department of Statistics at King Abdulaziz University, Jeddah, Saudi Arabia. She earned her Ph.D. in Statistics from the University of South Florida in 2019. Her research interests include statistical modeling, time series analysis, and applications of machine learning in environmental and health sciences. She has published several papers in international journals and conferences, focusing on climate change, Alzheimer's disease, and predictive modeling.

**Shumukh AL-qahtani** received her B.Sc. degree in Statistics from King Abdulaziz University, Jeddah, Saudi Arabia, in 2025. Her research interests include data analysis and time series forecasting, and she is currently working on publishing her first research paper in this field. She is also working at Baseera for Future Consultancy and Research.

**Hadil Ibrahim Hariry** received her B.Sc. degree in Statistics from King Abdulaziz University in Jeddah, Saudi Arabia, in 2025. and is currently employed at Basiera Consulting & Research Co., Ltd., Subcontracted within the National Center for Meteorology. Her research interests include statistical modeling and deep learning techniques, with applications in time series and climate analysis. She has practical experience with analytical tools such as Python and R and aspires to advance research in climate change forecasting.

**Mona Alshehri** received her M.Sc. degree in Statistics from King Abdulaziz University, Jeddah, Saudi Arabia, in 2025, with expertise in data analysis, statistical modeling, and machine learning. She has recently published two papers, and her research focuses on predictive analytics, time series forecasting, machine learning, and data-driven decision-making.

