# An Effective Forecasting Approach of Temperature Enabling Climate Change Analysis in Saudi Arabia

Sultan Noman Qasem[1], Samah M. Alzanin[2]*
Computer Science Department, College of Computer and Information Sciences,
Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh 11432, Saudi Arabia[1]
Department of Computer Science, College of Computer Engineering and Sciences,
Prince Sattam Bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia[2]

***Abstract*—Climate change is a global issue with far-reaching consequences, and understanding regional temperature patterns is critical for effective climate change analysis. In this context, accurate forecasting of temperature is critical for mitigating and understanding its impact. This study proposes an effective temperature forecasting approach in Saudi Arabia, a region highly vulnerable to climate change's effects, particularly rising temperatures. The approach uses advanced neural networks models such as the Long Short-Term Memory (LSTM), Gate Recurrent Unit (GRU), and Bidirectional LSTM (BiLSTM) model. A comparative analysis of these models is also introduced to determine the most effective model for forecasting the mean values of temperatures in the following years, understanding climate variability, and informing sustainable adaptation strategies. Several experiments are conducted to train and evaluate the models on a time series data of temperatures in Saudi Arabia, taken from a public dataset of countries' historical global average land temperatures. Performance metrics such as Mean Absolute Error (MAE), Mean Relative Error (MRE), Root Mean Squared Error (RMSE), and coefficient of determination (R-squared) are employed to measure the accuracy and reliability of each model. Experimental results show the models' ability to capture short-term fluctuations and long-term trends in temperature patterns. The findings contribute to the advancement of climate modeling methodologies and offer a basis for selecting a suitable model in similar environmental contexts.**

***Keywords—Climate change; Saudi Arabia; temperature; forecasting; recurrent neural network models***

## I. Introduction

The environment, human societies, and ecosystems are all significantly impacted by climate change, which is a major worldwide concern [1]. Being able to predict temperature changes with accuracy is a critical component of understanding and preventing climate change [2]. Temperature is a crucial indication of climate change, which is caused by complicated interactions between many environmental elements [2]. Climate change may have an influence on agriculture, change ecosystems, and cause more frequent and severe weather events [3]. To comprehend climate change consequences and develop practical adaptation and mitigation plans, an accurate temperature forecasting method is required, which is also crucial [4]. Numerical weather models that replicate atmospheric dynamics are the foundation of conventional temperature forecasting techniques [5]. Although these models have their uses, they cannot fully represent the intricate, nonlinear processes linked to climate change [6]. In contrast, machine learning (ML) is particularly worthy of dealing with big information [7], traffic recovery [8], social mobilization and migration prediction [9], seeing minute patterns [10], and responding to shifting circumstances [11]. As a potent instrument in climate research, ML provides advanced methods for examining past data, seeing trends, and forecasting outcomes. The use of machine learning for temperature forecasting has contributed greatly to climate change analysis. Researchers and scientists may obtain deeper insights into climate trends by utilizing ML algorithms [11]. This can help them make better-informed decisions and more accurate forecasts for reducing the effects of climate change. Temperature predictions may be made using linear regression and more sophisticated regression algorithms using historical data [12]. These models consider a number of variables, including location, season, and time of day. For the purpose of evaluating temperature data over time and identifying patterns, algorithms such as AutoRegressive Integrated Moving Average (ARIMA) and Seasonal ARIMA (SARIMA) work well [13]. By utilizing the advantages of various methods can improve prediction accuracy. However, there is a limitation in capturing temporal relationships of time series temperature data, which is the gap that the study is trying to fill to improve the performance of the forecasting process. The main goal of this study is to forecast Saudi Arabia's average temperature patterns using sophisticated forecasting techniques. It proposes an effective approach to produce accurate and dependable forecasts of future temperature trends by utilizing cutting-edge methods, including Gate Recurrent Unit (GRU), Bidirectional LSTM (BiLSTM), and Long Short-Term Memory (LSTM). These recurrent neural networks models have been chosen because they are excellent to find complex patterns in temperature data and capturing the intricate temporal correlations and relationships present in such time series data, which makes them useful for predicting applications. Thus, the main contributions of the work can be summarized in the following points:

- Proposing a forecasting average temperature approach to achieve an accurate analysis of climate change in Saudi Arabia.
- Developing effective neural networks models that are able to capture temporal relationships of time series temperature data.

*Corresponding Author.

- Evaluating the proposed approach on a time series data of temperatures in Saudi Arabia, taken from a public dataset of countries' historical global average land temperatures.
- A comparative analysis of the developed models will be introduced to determine the most effective model for forecasting the average values of temperatures in the next years.

The rest of the paper is organized into five sections. Section II gives a literature review. Section III explains the materials and methods in detail. Section IV presents the experimental results and discussions. Finally, section V summarizes the conclusion and future work.

## II. Literature Review

Climate change boosts temperatures and causes water scarcity [14]. Extreme weather events such as severe drought, heavy downpours, heat waves, and cold waves are becoming increasingly regular. Climate change has a wide-ranging impact on people's life, including agriculture and fisheries [15], mental health [16], physical health [17], and the economy [18]. Overall, the potential costs as a result of climate change outweighed the advantages. Communities with lower levels of socioeconomic development are more likely to endure the potential consequences. Many third-world nations are located in tropical climates, which are particularly vulnerable to climate change. Climate change has led to a significant influence on Southeast Asia, North and South India, Sub-Saharan Africa, West Africa, East and Southern Africa, Northern Latin America, and Central America [19]. As a result, these nations' food security is exposed and it is critical to predict and mitigate the effects of climate change. They are required to minimize the vulnerability of life in human-related sectors such as ecosystems, health, agriculture and fishing, economics, and culture. Monitoring temperature variations is one method for anticipating climate change and forecasting future temperatures can help humans prepare for future conditions. Consequently, the fast growth of statistical methodology, certain methods may now be used to forecast the future, including temperature. With a large amount of data from previous events, regression and statistical modeling tools are utilized for creating a relationship between variables [20].

Many researchers have employed regression models, particularly Autoregression (AR), to predict not only temperature but also other scientific variables. Yau et al. [21] applied AR and Support Vector Machine Regression (SVMR) Integrated Moving Average to predict the daily arrival of visitors in southwest China. Witaradya and Putranto [22] proposed to use AR for temperature predication and investigated its effectiveness as a regression mode for time-series data. Zakaria et al. [23] used the ARIMA model to analyze data from four weather stations in Iraq between 1990 and 2011. Chen et al. [24] examined monthly mean temperatures in Nanjing, China. They used monthly mean temperature data from 1951 to 2014 as the training set and data from 2015 to 2017 as the testing set to create an ARIMA model for their research. Murat et al. [25] introduced research on predicting and modeling daily temperature for four European sites in various climatic zones using data from 1980 to 2010. They employed the Seasonal Autoregressive Integrated Moving Average (SARIMA), and ARIMA with external regression method and demonstrated that the generated models could describe the data series and be used to estimate future daily temperatures. Dwivedi et al. [26] used the SARIMA model to forecast the average temperature for the city of Gujarat, India, using data from 1984 to 2015. They tested numerous models and chose the best SARIMA model for average temperature forecasting based on the Akaike Information Criterion (ACI). They examined the model's adequacy, and the diagnostics revealed that the model was reliable for projecting monthly average temperatures.

Also, Asha et al. [27] introduced an approach to forecasting daily maximum temperatures for four distinct locations in Kerala, India, using three different methods: ARIMA, SARIMA, and Autoregressive Fractional Integrated Moving Average (ARFIMA), utilizing data from January 2019 to December 2020. They then examined the performance of three techniques using measures such as Mean Squared Error (MSE), Mean Squared Error (MSE), and percentage accuracy (PA). According to the results, all of the models performed well, with the ARFIMA model outperforming the ARIMA and SARIMA models. Hennayake et al. [28] proposed a method using Long Short-Term Memory (LSTM) model for forecasting the most important meteorological variables, such as precipitation and temperature, for a weather station in Sri Lanka. They evaluated model performance using Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) measures. As a consequence, they demonstrated that both LSTM models designed for precipitation and temperature forecasting functioned well and could be used to forecast precipitation and temperature accurately.

Mitu and Hasan [29] presented a work based on SARIMA model for Memphis, Tennessee, using daily temperature data from 2016 to 2019. They examined temperature data from that time period to identify patterns and transitory fluctuations. They employed the Mann-Kendall (M-K) test as a nonparametric tool to discover time series analysis trends. They also used the SARIMA approach to anticipate the temperature for the following 50 days. The prognosis also indicated an upward tendency for the location. Dimri et al. [30] utilized monthly average maximum and minimum temperatures for the Bhagirathi River watershed in India using a seasonal ARIMA model based on data from 2001 to 2020. Their findings revealed that projected data is consistent with the data trend.

Gangshetty et al. [31] published a work for time series temperature forecasting in Pune, India, utilizing data from 2009 to 2020. They used SARIMA model and autocorrelation function with the partial autocorrelation function, as well as normalized residuals, to identify the best fit for the time series for their study. They discovered that the model performed well at predicting temperature values. Hoang et al. [32] implemented and developed a model using an LSTM on Amazon Web Services (AWS) machine learning platform. They discovered that the LSTM model produced significant and accurate results compared with other weather forecasting models.

Recently, Jaharabi et al. [33] investigated the use of machine and deep learning for temperature prediction of major cities in the world. Koçak [34] presented a time-series prediction approach of temperature based on LSTM and ARIMA models. Khokhar et al. [35] introduced a comparative analysis of ARIMA, LSTM, and BiLSTM for temperature and rainfall forecasting on Pakistan's time-series data of 116 years. Jafarian-Namin et al. [36] applied ARIMA and Artificial Neural Network (ANN) models for monthly temperature analysis and prediction on Tehran's time-series data. Topalova and Radoyska [37] proposed an automated change detection method to track the climate change of temperature in local geographic regions using a two-level structure of neural networks. However, the research gap of the previous work is that no study investigates, analyzes, and develops an effective machine learning model for climate change in Saudi Arabia's average temperature. This work explores the average temperature change in Saudi Arabia for the past 152 years (from 1861 to 2013). Furthermore, we show the effectiveness of Gate Recurrent Unit-based Neural Network (GRU-NN) model for average temperature forecasting and compare it with other RNN variants and other common models in the previous studies.

## III. Materials and Methods

### A. Earth Surface Temperature Dataset

The Earth Surface Temperature (EST) dataset is received from the KAGGLE platform [38]. It is collected by the National Oceanic and Atmospheric Administration (NOAA) Merged Land-Ocean Surface Temperature Analysis (MLOST), NASA GISTEMP, and UK HadCrut organizations. This collected data is repackaged or put together by Berkeley Earth and affiliated with Lawrence Berkeley National Laboratory. The EST dataset has several CSV files, including global ocean-and-land temperatures, global land temperatures by city, global average land temperature by country, global land temperatures by major city, and global average land temperature by state. Each file in the EST dataset comprises certain types of data that are required for climate data analysis and finding long-term trends and patterns in temperature and climate variables. The date, nation, average temperature, longitude, and latitude columns give critical information that allows researchers to obtain insight into the environmental impact of climate change and develop mitigation solutions. The study focuses on the monthly global land temperatures by city file, which contains 8599212 instances and seven attributes. Table I presents the attribute types of the selected dataset file. From this file, the data instances related to Saudi Arabia and its cities are filtered to form a dataset of 12795 instances. It consists of the average temperature in Saudi Arabia from January 1st, 1861 to September 1st, 2013. Table II gives the first and the last five rows of the dataset used in this study.

The DT column gives the date of collected temperature data as a time series. The average temperature column gives data on temperatures for the location in which the data was gathered. This column is commonly represented as a numerical data type, with the temperature measured in degrees Celsius. The average temperature column is critical for assessing climate data to determine temperature trends over time and identify changes in temperature patterns caused by climate change. The longitude column describes a point's east-west location on the Earth's surface. The latitude indicates the north-south position of a specific location wherever temperature data was recorded. The longitude and latitude columns are commonly represented as a numeric value expressed in degrees with the longitude or latitude letter.

TABLE I. Types of Attributes for the Selected Dataset File

| No. | Column | Data type |
|---|---|---|
| 1 | DT | Date |
| 2 | AverageTemperature | float64 |
| 3 | AverageTemperatureUncertainty | float64 |
| 4 | City | String |
| 5 | Country | String |
| 6 | Latitude | String |
| 7 | Longitude | String |

TABLE II. First and Last Five Rows of the Dataset used in this Study

| DT | AverageTemperature | AverageTemperatureUncertainty | City | Country | Latitude | Longitude |
|---|---|---|---|---|---|---|
| 1861-01-01 | 17.429 | 1.834 | Abha | Saudi Arabia | 18.48N | 42.25E |
| 1861-02-01 | 19.162 | 1.810 | Abha | Saudi Arabia | 18.48N | 42.25E |
| 1861-03-01 | 21.228 | 1.610 | Abha | Saudi Arabia | 18.48N | 42.25E |
| 1861-04-01 | 23.592 | 1.711 | Abha | Saudi Arabia | 18.48N | 42.25E |
| 1861-05-01 | 25.909 | 1.676 | Abha | Saudi Arabia | 18.48N | 42.25E |
| ... | ... | ... | ... | ... | ... | ... |
| 2012-12-01 | 13.012 | 0.423 | Tabuk | Saudi Arabia | 28.13N | 37.27E |
| 2013-01-01 | 12.134 | 0.328 | Tabuk | Saudi Arabia | 28.13N | 37.27E |
| 2013-02-01 | 14.880 | 0.232 | Tabuk | Saudi Arabia | 28.13N | 37.27E |
| 2013-03-01 | 18.676 | 1.919 | Tabuk | Saudi Arabia | 28.13N | 37.27E |
| 2013-04-01 | 21.375 | 0.612 | Tabuk | Saudi Arabia | 28.13N | 37.27E |

## B. Recurrent Neural Networks Models

Recurrent neural networks (RNNs) are a type of artificial neural network (ANN) developed to process data from sequential activities. The RNNs are able to maintain the hidden state or memory of past inputs compared with standard feed-forward neural networks because of their connections forming directed cycles. The use of RNN's internal state makes it suitable for processing sequences of input, especially time series data of some applications, such as speech recognition, natural language processing, and cloud service forecasting, in which temporal dependencies or context are crucial.

The key feature of RNNs is their ability to maintain a hidden state that captures information about previous inputs in the sequence. This hidden state is updated at each time step and influences the network's output at the current time step. The basic formula for updating the hidden state $h_t$ at time $t$ in an RNN is:

$$h_t = f(W_{hh}h_{t-1} + W_{xh}x_t + b_h) \quad (1)$$

where, $h_t$ is the hidden state at time $t$, $x_t$ is the input at time $t$, $W_{hh}$ is the weight matrix for the hidden state, $W_{xh}$ is the weight matrix for the input, $b_h$ is the bias vector, and $f$ is the activation function.

However, the vanishing gradient problem, in which the gradients decrease exponentially as they are transmitted back in time during training, is one issue facing the classic RNNs. The RNNs have difficulty capturing long-term dependencies in sequences because of this restriction [19]. To overcome this problem, effective variations of RNNs have been created, including Long Short-Term Memory networks (LSTMs), Bidirectional LSTM (BiLSTM), and Gated Recurrent Units (GRUs). These architectures are able to capture long-term dependencies in sequential data because they have the capabilities to selectively preserve or forget information. The following subsections explain the effective neural networks models proposed to forecast the average values of temperatures from historical time-series average temperature data.

*1) Long Short-Term Memory-based Neural Network (LSTM-NN) Model:* This neural network model depends on long short-term memory (LSTM) cells, which is a popular type of RNNs. It is most frequently utilized in sequential data issues. The design of an LSTM cell consists of four primary parts: input gate, cell state, forget gate, and output gate. The gates determine which data should be retained or discarded for each cell, which represents a time step. In order to create a forecast, it only does that by passing pertinent data down a lengthy chain of sequences. For that reason, LSTM-NN may learn long-term dependencies more effectively than traditional RNNs. To regulate when memory material is given to other cells, LSTM makes use of cell state. Fig. 1 describes the LSTM cell design [39].

*2) Gated Recurrent Unit-based Neural Network (GRU-NN) Model:* It is developed based on the gated recurrent unit (GRU). It is another type of RNN, presented with the intention of preserving significant information in a sequence, much like LSTM-NN. On the other hand, the architecture of GRU-NN has fewer parameters and is less complex, making it computationally less expensive and faster to train. The GRU has just two gates: the reset gate and the update gate, which can eliminate the cell state and make the full memory accessible to other units. The update gate specifies which data to retain, whereas the reset gate merges the fresh input with the prior memory cell. Fig. 2 depicts the architecture of the GRU cell design [39].

*3) Bidirectional LSTM-based Neural Network (BiLSTM-NN) Model:* Based on the concept of Bidirectional RNNs [20], it is an extension of conventional LSTM-based neural networks that can enhance model performance. BiLSTM takes into account sequences in both forward and backward order as it is the result of combining several LSTMs on input in several opposed orientations. The essential components of a BiLSTM-NN are forward LSTM, backward LSTM, and combination. The forward LSTM can process the sequence from beginning to end. The backward LSTM processes the sequence from end to beginning. The outputs from both directions are combined or merged before being sent on to the next layer or job through a combination component. This may offer more network information to help with forecast accuracy. The main advantage of utilizing a BiLSTM is its capacity to gather information from both the past and the future at each time step. This is especially valuable for jobs that require knowing the context from both sides. In average temperature forecasting, information about an average temperature can be influenced by both previous and subsequent average temperatures. Fig. 3 illustrates the design of BiLSTM cell architecture [39].

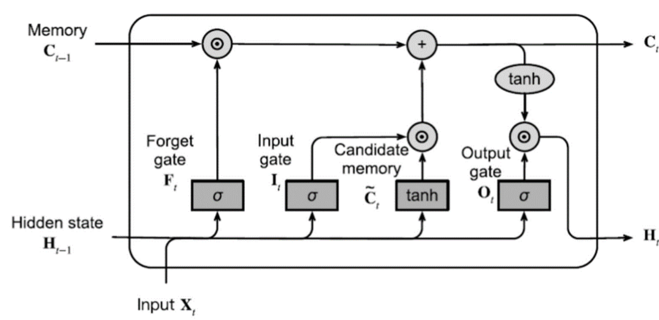


Fig. 1. The architecture of LSTM cell design.

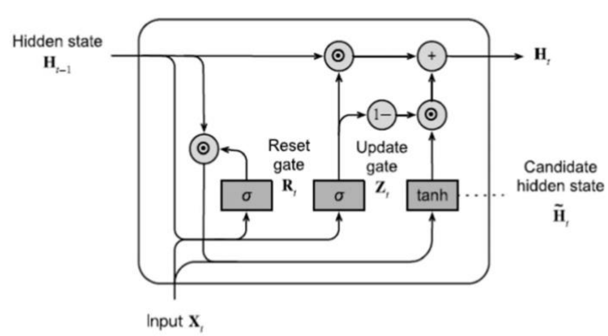


Fig. 2. The architecture of GRU cell design.

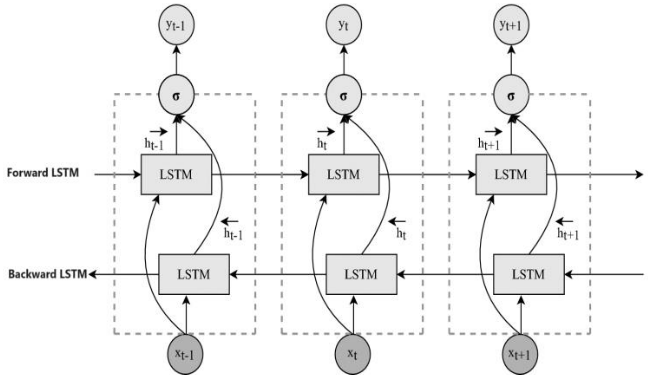


Fig. 3. Design of BiLSTM cell architecture.

### C. Earth Surface Temperature Dataset

The flowchart of the proposed approach for this study is shown in Fig. 4. It contains four core steps, including data preprocessing, data analysis and splitting, model building and training, and model comparison and evaluation. Explaining these steps in detail is given the following subsections.

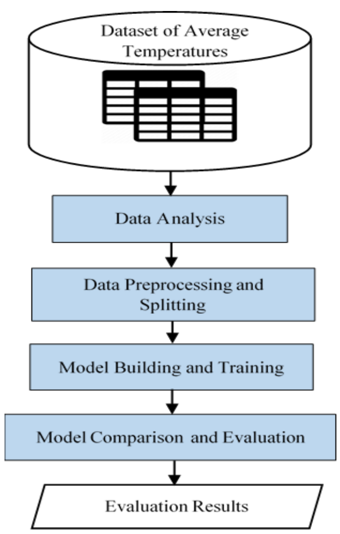


Fig. 4. Flowchart of the proposed approach.

*1) Data analysis:* Data analysis is a critical step in understanding data and making a decision about which models of machine learning are appropriate for prediction or forecasting duty. Data analysis of average temperature in Saudi Arabia involves loading the average temperature data in its structured formats, exploring the distribution of average temperatures over time, visualizing trends, and extracting meaningful insights. First, the step checks if the data is stationary. The Augmented Dickey-Fuller (ADF) statistical test is usually used to determine whether a particular time series data is stationary or not. The null hypothesis of the ADF test is not stationary. To reject the null hypothesis, the p-value should be less than 0.05.

Other tasks of the data analysis step include calculating the basic statistics such as mean, median, and standard deviation, comparing average temperatures across different years, months, and cities within the country, visualizing their trends using relevant visualization charts, and looking for seasonality or patterns in the data.

*2) Data Pre-processing:* In data mining applications, data pre-processing is the most critical and time-consuming step. Because the temperature data of this study is the models' input, it is reasonable that the more accurate the input, the more accurate the output. The obtained data are the monthly temperature data from previous years. Temperature readings may not be recorded or have no value for a variety of reasons. In this study, we employed the interpolation method to prevent data bias. The interpolation method uses two known data points to estimate unknown data values. It is most commonly used to fill in missing values in a data record or series during data pre-processing.

An interpolation method is used to fill missing values with the aid of their neighbors. Filling missing time-series data with average values does not work well. Therefore, interpolation is suitable for time-series data to fill missing values with the preceding one or two values. For time-series data of average temperatures, it is preferred to fill the month's average temperature with the mean of the past two months rather than the months' mean.

The second method in data preprocessing step is data normalization. In data normalization, the average values of temperatures are scaled into a small and specific domain between 0 and 1 to prevent the neural network models from biasing the results. In this step, the min–max normalization method is used to convert the data to the range of 0 to 1. Computing min–max value for each average temperature $t_i$ is done by the following equation, in which$Max_{t_n}$and $Min_{t_n}$are the maximum and minimum values of average temperatures.

$$t_i = \frac{t_i - Min_{t_n}}{Max_{t_n} - Min_{t_n}} \quad (2)$$

The third method is data splitting. In data splitting, we use a train-validation-test split technique to divide the dataset into training, validation, and test sets with a ratio of 60%, 30%, and 10%, respectively. First, we take 10% of the dataset for the unseen test set. Then, from the 90%, we take 30% for the validation set, and the remaining 60% is as a training set. Because we deal with time series data, temporal aspects are considered when splitting to ensure that the test set represents future data.

*3) Model building and training:* The model building and training step is an iterative process of evaluation and refinement to produce a model that performs effectively on the specified task. The model's effectiveness depends on carefully selecting architectures, hyper-parameters, and optimization algorithms at this step. Based on our experience with deep learning and the size of data, we build each of the three recurrent neural network models to have one hidden layer with $x$ units, where $x \in [50, 100, 150]$ These three values are enough for search on the best number of hidden layer units, achieving an accurate forecasting of average temperatures from time-serious data. These built models are trained for 200 epochs, and the best values of the parameters set are preserved for evaluation. It is worth noting that a model's gates interact with data using a set of weights and biases known as parameters or hyper-parameters. During training, the back-propagation method is used to update these parameters. The final parameters set is referred to as the trained model and is used to make forecasting. The more parameters a model contains, the more computation time and resources it requires. As a result, the total number of parameters indicates a model's complexity and efficiency. Table III shows the number of parameters in each model.

TABLE III. MODEL'S TOTAL NUMBER OF PARAMETERS

| **Model** | **Total Number of Parameters** | | |
|---|---|---|---|
| | ***50 units*** | ***100 units*** | ***150 units*** |
| LSTM-NN | 10,451 | 40,901 | 91,351 |
| GRU-NN | 8,001 | 31,001 | 69,001 |
| BiLSTM-NN | 20,901 | 81,801 | 182,701 |

To successfully train the models, we feed the training set into the model and adjust the model's weights based on the loss values between forecasted and actual values. The models are iteratively updated their parameters through a series of pre-defined epochs. The models' performance is also monitored on the validation set to detect the under-fitting or over-fitting in the training progress.

*4) Model evaluation and comparison:* Once the training process is complete, the test set's model evaluation and comparison step is started to assess the generalization performance of trained models on the unseen data. The performance measures utilized to evaluate the proposed temperature forecasting models are statistical measurements. They are used to assess the models' ability to fit the data and include the Mean Absolute Error (MAE), Mean Relative Error (MRE), Root Mean Squared Error (RMSE), and coefficient of determination (R-squared). Models work well on the test set, obtaining the lowest value of all error measures. These lowest values of errors imply that the discrepancies between the actual and forecasted values are relatively small and unbiased. A higher R-squared value indicates that the models can accurately fit the data. In other words, error metrics evaluate the models' capacity to correctly estimate average temperatures based on the error values. The R-squared statistic simply indicates the relationship between actual and forecasted average temperatures. The following equations are used to compute all of the used performance measures:

$$MAE = \frac{1}{N}\sum_{k=1}^{N}|v_k - \widehat{v_k}| \tag{3}$$

$$MRE = \frac{1}{N}\sum_{k=1}^{N}\frac{|v_k - \widehat{v_k}|}{v_k} \tag{4}$$

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{k=1}^{N}(v_k - \widehat{v_k})^2} \tag{5}$$

$$\text{R-squared } = 1 - \frac{\sum_{k=1}^{N}(v_k - \widehat{v_k})^2}{\sum_{k=1}^{N}(v_k - \bar{v})^2} \tag{6}$$

The actual values of average temperatures are denoted by $v_k$, the forecasted values are represented by $\widehat{v_k}$, and the mean value of actual average temperatures is denoted by $\bar{v}$.

## IV. EXPERIMENTAL RESULTS AND DISCUSSIONS

This section conducts several experiments on the dataset using Python programming language. In the first experiment, we applied the ADF test of the data analysis step to check whether the average temperature data in Saudi Arabia was stationary. This is required for Table IV, which presents the results of the ADF test on the dataset.

Table IV shows that the test statistics are lower than the Critical Value of 5%, and the p-value (0.0193) is less than 0.05. This means we reject the null hypothesis (data is not stationary). Therefore, the time series data of Saudi Arabia's temperature seems stationary. In stationary time-series data, there is no notable trend or variation in the mean, which makes it a critical quality for trustworthy analysis and modeling. Constant variance is another essential attribute, suggesting that data point dispersion remains constant across time.

The second experiment is applied to show the average temperature trend in Saudi Arabia per year based on time-series data. We resample the mean of average temperatures yearly instead of monthly for the last 152 years, as shown in Fig. 5.

TABLE IV. RESULTS OF AUGMENTED DICKEY-FULLER (ADF) TEST

| **Metric** | **Value** |
|---|---|
| Test Statistic | -3.2128 |
| p-value | 0.0193 |
| Lags Used | 40.0000 |
| Number of Observations Used | 12754.0000 |
| Critical Value (1%) | -3.4309 |
| Critical Value (5%) | -2.8618 |
| Critical Value (10%) | -2.5669 |

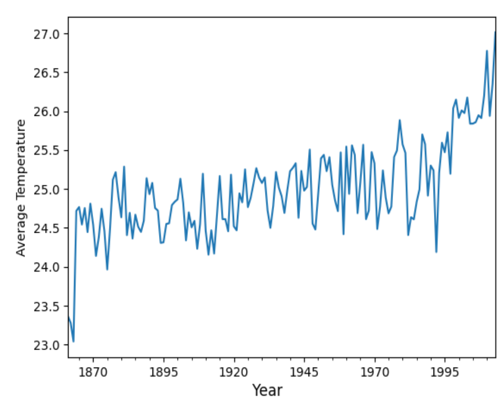


Fig. 5. The average temperature of Saudi Arabia in the last 152 years

From Fig. 5, we can see that there is a significant increase in the average temperature after the year 1995. Fig. 6 illustrates the average temperature curve of Saudi Arabia colored orange. It shows the growth of average temperature per year.

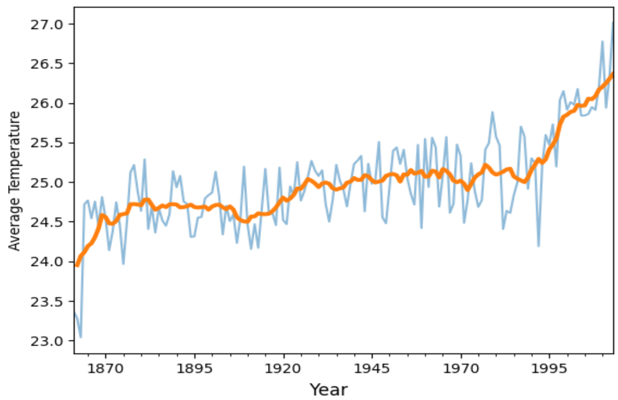


Fig. 6. The average temperature curve of Saudi Arabia in the last 152 years.

The trend of actual and forecast temperature regarding the growth of average temperatures is given in Fig. 7. As seen in this TA, the trend of temperature forecasting in the year 2045 will reach approximately 28 °C.

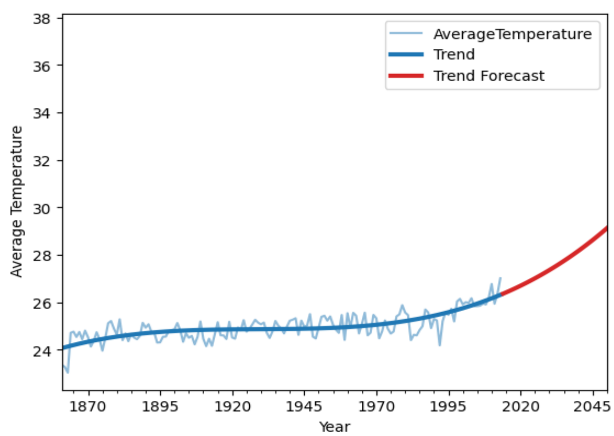


Fig. 7. Trend of actual and forecast temperature regarding to the average temperatures.

Through the experiment, we also visualize the monthly average temperature of Saudi Arabia from January 1st, 1861 to September 1st, 2013, as shown in Fig. 8. We can see a fluctuation in the temperature distribution with a little increase until 1995. This fluctuation is due to temperature variations throughout the months of each year. Fig. 9 illustrates the major cities of Saudi Arabia with the highest average temperatures. It shows that Buraydah and Riyadh have the highest average temperatures compared to other cities. Fig. 10 uses a box plot to visualize the monthly distribution and range of numerical data using the average daily temperature values. The box plot shows that February, March, May, and December have the most extended boxes and whiskers. However, February has the most extensive distribution of average temperatures.

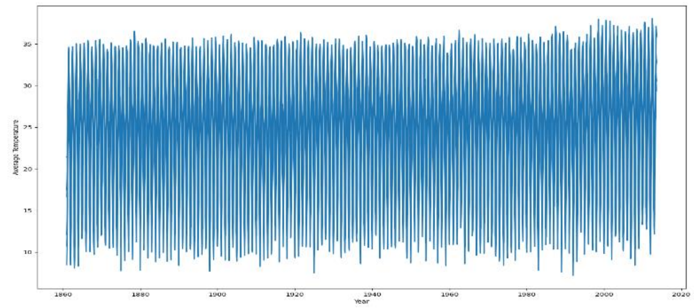

Fig. 8. Distribution of monthly average temperature time-series data

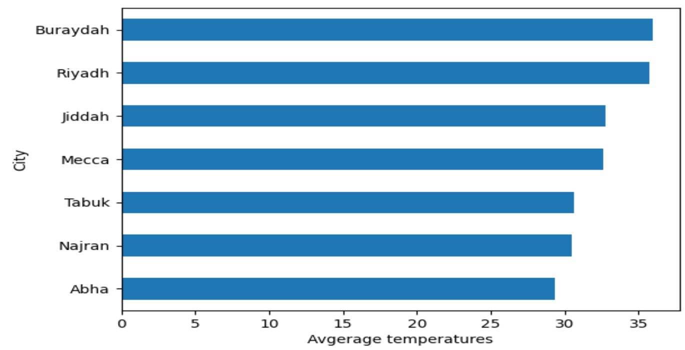


Fig. 9. Cities of Saudi Arabia with highest average temperature

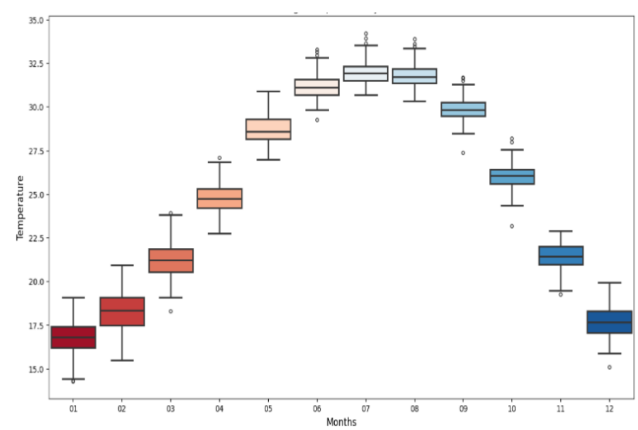


Fig. 10. Average temperature by months

Fig. 11 shows the average temperature season-wise over the years. The average temperature seasonally across the years can have a variety of effects depending on geographical location, climatic patterns, and local ecosystems of Saudi Arabia.

The previous figures give a picture for understanding the long-term trends of average temperatures and their season-wise over the years, which are crucial for mitigating and addressing the impacts of climate change on human society and the environment. It enables informed decision-making and the creation of adaptation strategies to climate change.

After performing a data analysis experiment, we conduct another experiment for average temperature forecasting using developed neural network models. We first apply the data pre-processing step. We check the number of null values in the dataset. We found that 146 records have null values for the number of Average Temperature and Average Temperature Uncertainty columns, as presented in Table V. For these null values, we fill them using an interpolation technique, which replaces them with the mean of the past two months. Then, we normalize the average temperature values to be in the range between 0 and 1. The normalization of average temperature values is required for the gradient descent of the neural network. Next, the dataset is divided into training and test sets using the data splitting step. Fig. 12 visualizes the distribution of training and test sets.

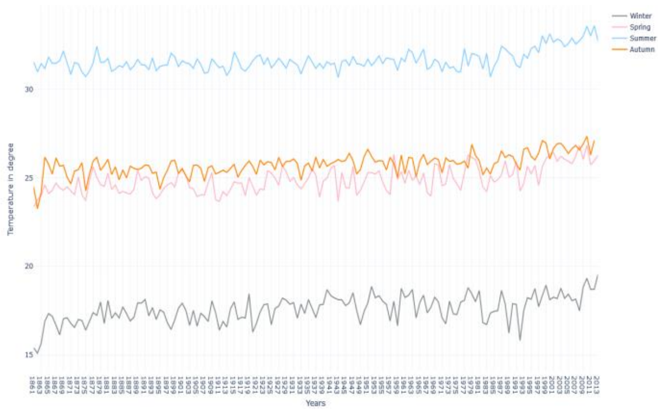

Fig. 11. Average temperature season-wise over the years.

TABLE V. CHECKING THE NUMBER OF NULL VALUES

| Column Name | No. of Null Values |
|---|---|
| DT | 0 |
| AverageTemperature | 146 |
| AverageTemperatureUncertainty | 146 |
| City | 0 |
| Country | 0 |
| Latitude | 0 |
| Longitude | 0 |

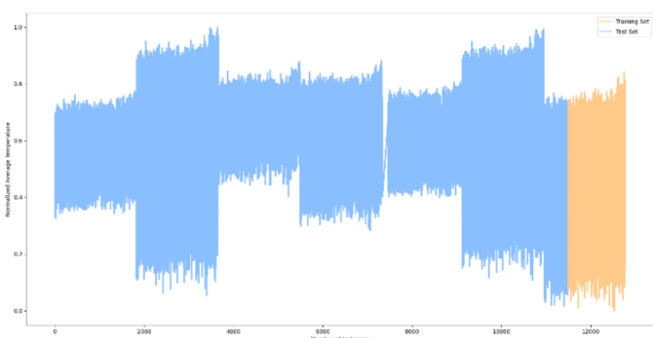

Fig. 12. Distribution of training and test sets.

After that, we train the LSTM-NN, BiLSTM-NN, and GRU-NN models on the training set using different numbers of hidden layer's units, which are 50, 100, and 150. In the training process, 30% of the training set is used for validation. Fig. 13 to 23 show the training and validation loss during the learning progress for the models at 50 units, 100 units, and 150 units. As shown in these figures, we can see that the gap between training and validation loss for the models with 50 units is very small compared to using other numbers of hidden layer's units. This means that there is no over-fitting in the training of the models. However, the gap between training and validation loss for the GRU-NN with 50 hidden layer's units is the smallest, indicating that its performance is better than the other models.

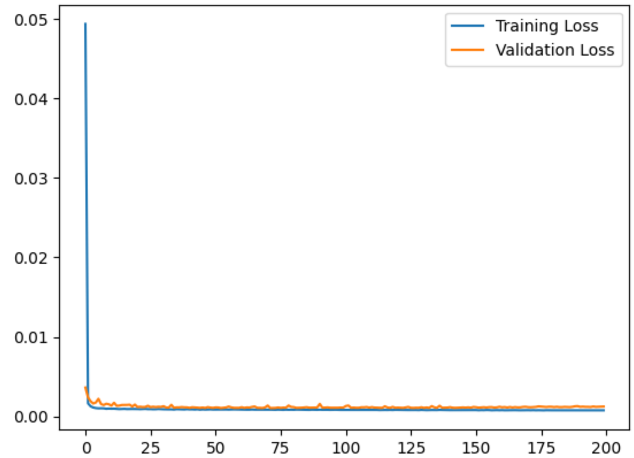


Fig. 13. Training and validation loss of LSTM-NN with 50 hidden layer's units

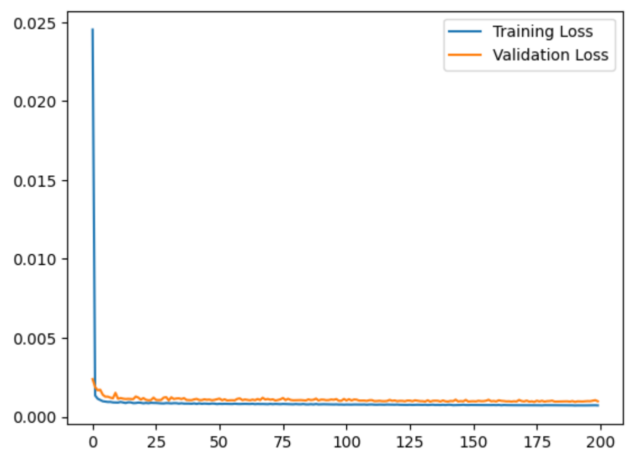


Fig. 14. Training and validation loss of BiLSTM-NN with 50 hidden layer's units

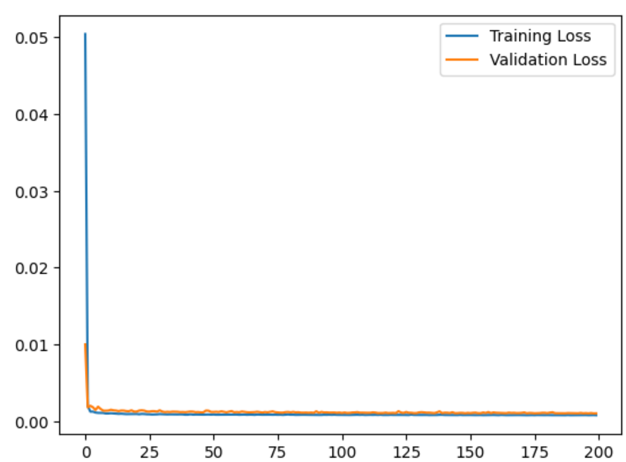


Fig. 15. Training and validation loss of GRU-NN with 50 hidden layer's units

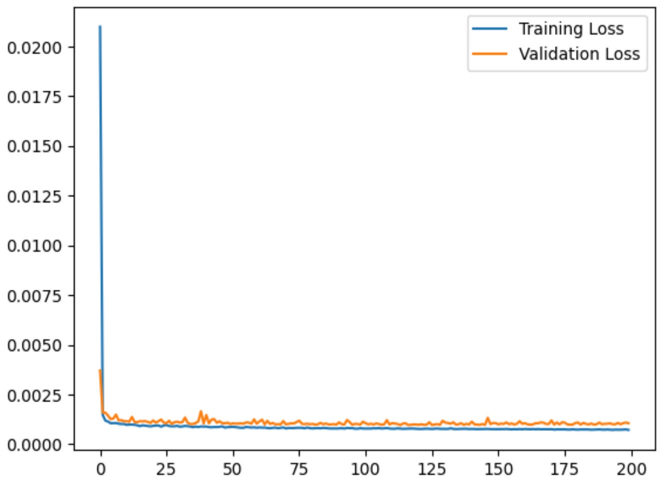


Fig. 16. Training and validation loss of LSTM-NN with 100 hidden layer's units.

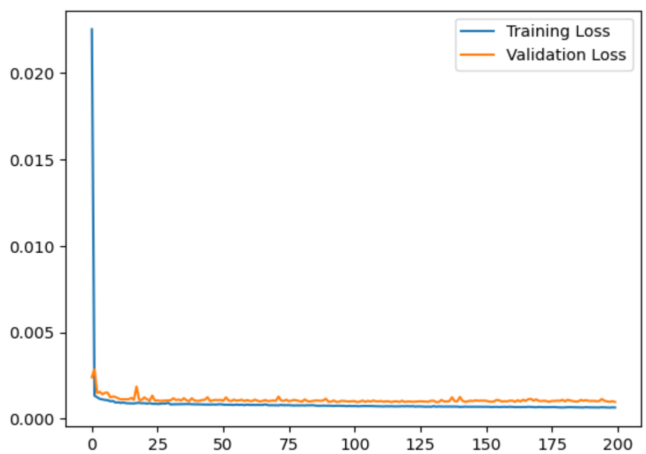


Fig. 17. Training and validation loss of BiLSTM-NN with 100 hidden layer's units.

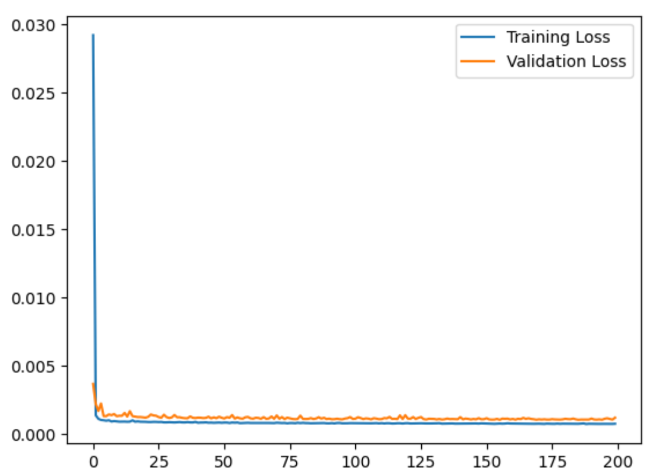


Fig. 18. Training and validation loss of GRU-NN with 100 hidden layer's units.

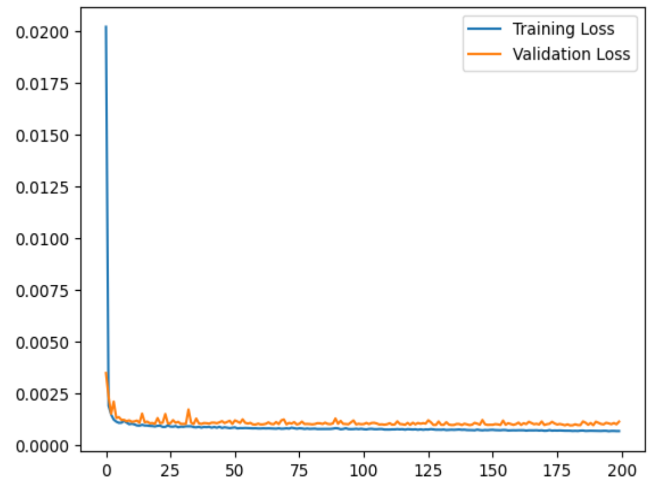


Fig. 19. Training and validation loss of LSTM-NN with 150 hidden layer's units.

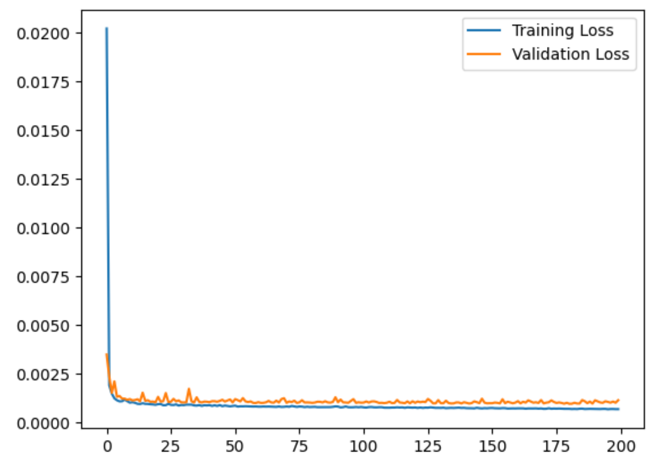


Fig. 20. Training and validation loss of BiLSTM-NN with 150 hidden layer's units.

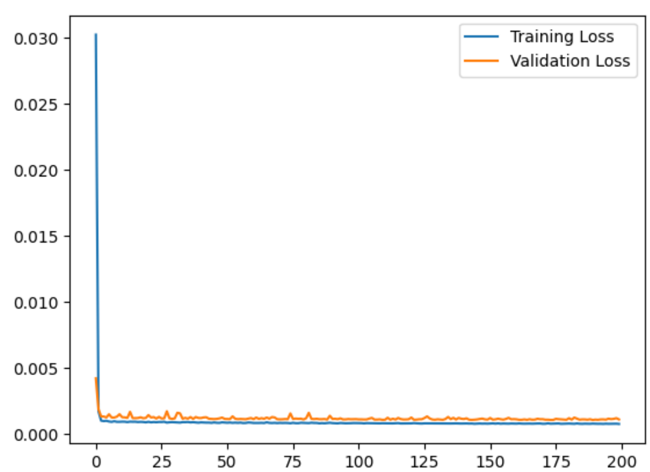


Fig. 21. Training and validation loss of GRU-NN with 150 hidden layer's units.

Table VI gives the performance results regarding MAE, MRE, RMSE, and R-squared for the LSTM-NN, BiLSTM-NN, and GRU-NN models on the test set using a different number of hidden layer's units.

TABLE VI. PERFORMANCE RESULTS OF DEVELOPED MODELS

| Model | Evaluation Measure | Number of Units in Hidden Layer | | |
|---|---|---|---|---|
| | | *50 Units* | *100 Units* | *150 Units* |
| LSTM-NN | MAE | 0.04213 | 0.03561 | 0.03811 |
| | MRE | 0.09291 | 0.07855 | 0.08406 |
| | RMSE | 0.05252 | 0.04468 | 0.04787 |
| | R-squared | 94.808% | 96.242% | 95.685% |
| BiLSTM-NN | MAE | 0.03309 | 0.03251 | 0.03062 |
| | MRE | 0.07298 | 0.07172 | 0.06754 |
| | RMSE | 0.04172 | 0.04264 | 0.03931 |
| | R-squared | 96.723% | 96.578% | 97.091% |
| GRU-NN | MAE | **0.02976** | 0.03072 | 0.03121 |
| | MRE | **0.06565** | 0.06777 | 0.06883 |
| | RMSE | **0.03889** | 0.04015 | 0.04046 |
| | R-squared | **97.152%** | 96.965% | 96.919% |

As listed in Table VI, we can see that the GRU-NN with 50 hidden layer's units achieves the best performance result on the test set for all evaluation measures, as highlighted in bold font. Figs. 22-24 display the distributions of ground-truth average temperatures of the test set and forecasted average temperatures generated by the three models with 50 hidden layer's units. We can see in Fig. 24 that the forecasted average temperatures generated by the GRU-NN model are more fitted with ground-truth average temperatures of the test set than the two other models.

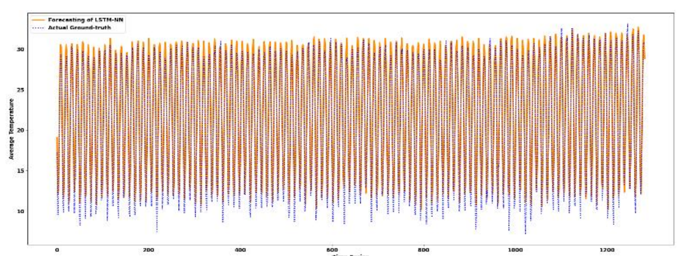

Fig. 22. Distribution of ground truth and forecasted average temperatures for the LSTM-NN model with 50 hidden layer's units.

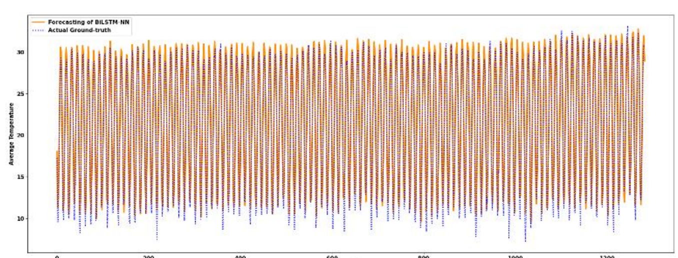

Fig. 23. Distribution of ground-truth and forecasted average temperatures for BiLSTM-NN model with 50 hidden layer's units.

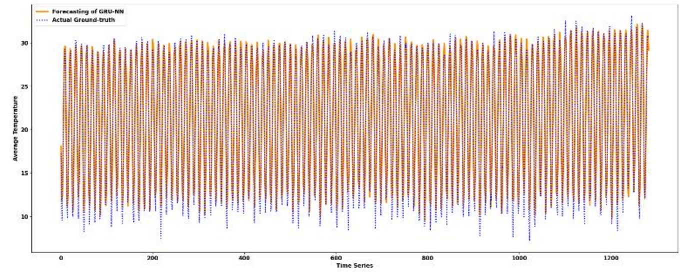

Fig. 24. Distribution of ground truth and forecasted average temperatures for GRU-NN model with 50 hidden layer's units.

For visualizing the performance of the three models, Fig. 25 compares the results of RMSE, showing that the GRU-NN model with 50 hidden layer's units has a lower value than the other models. Moreover, we compare the GRU-NN model with 50 hidden layer's units with two common regression models used widely in the literature review, which are ARIMA [21] and SVMR [21]. Fig. 26 and Fig. 27 present the distribution of ground truth and forecasted average temperatures for the ARIMA [21] and SVMR [21] models, respectively. We can see that the ground truth and forecasted average temperatures are not more fitted like the GRU-NN model. Finally, we compare the results of RMSE for ARIMA and SVMR models with the GRU-NN model, as shown in Fig. 28. Clearly, we can see that the GRU-NN model achieves the lowest RMSE value compared with ARIMA and SVMR models. This confirms the ability of the GRU-NN model with 50 hidden layer's units for accurate temperature forecasting and its suitability for the nature of Saudi Arabia's time-series temperature data.

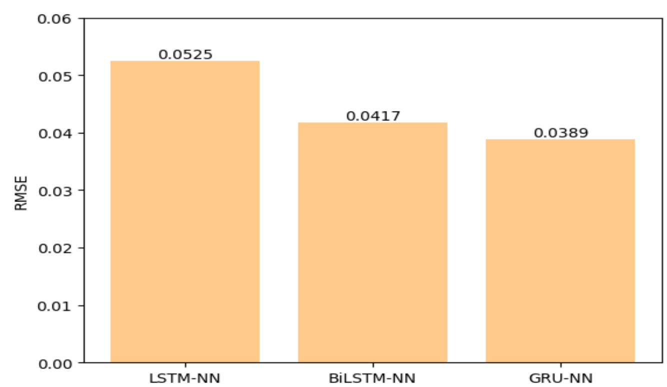


Fig. 25. Results of RMSE for the three models with 50 hidden layer's units.

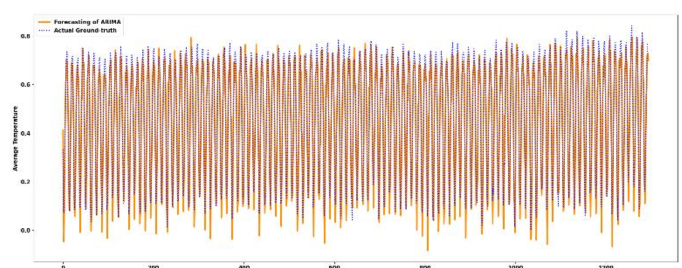

Fig. 26. Distribution of ground truth and forecasted average temperatures for ARIMA model.

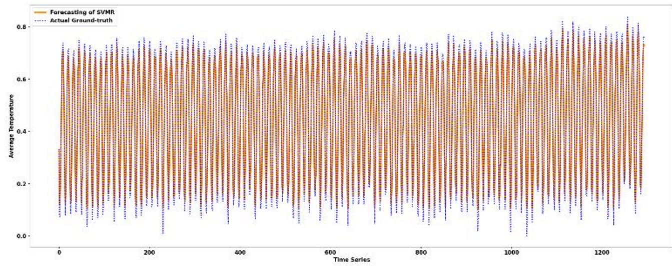


Fig. 27. Distribution of ground truth and forecasted average temperatures for SVMR model.

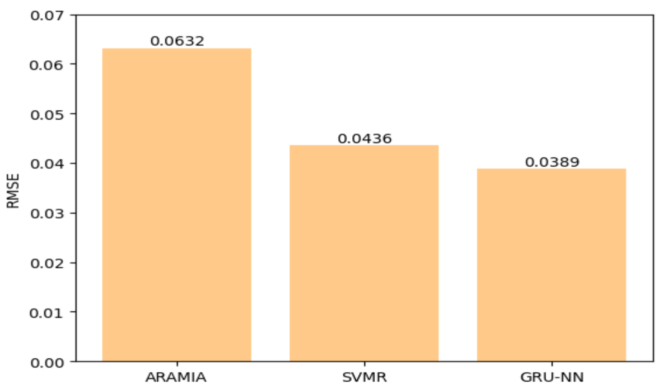


Fig. 28. Comparison results of RMSE for ARIMA, SVMR, and GRU-NN with 50 hidden layer's units models.

## V. Conclusion and Future Work

Analyzing and forecasting the temperature of Saudi Arabia region using historical time-series data can give valuable insights for climate change mitigation and adaptation plans. Decision-makers can use the analysis outcomes and forecasts to plan and execute actions to mitigate the anticipated effects of climate change, such as water scarcity, severe temperatures, and changes in agricultural methods. Average temperature forecasting utilizing recurrent neural networks plays a vital role in accurate climate change analysis.

The use of sophisticated neural network architectures, such as LSTM-NN, BiLSTM-NN, and GRU-NN, has shown great promise in capturing the complicated patterns and temporal correlations seen in temperature time-series data. They have been exposed to be successful in capturing the temporal dependencies seen in Saudi Arabia's historical temperature data. Their capacity to understand long-term relationships allows for more accurate representations of climatic trends and fluctuations across time. The experimental results showed that the GRU-NN model has improved accuracy in temperature forecasting for Saudi Arabia compared with other models. The model has demonstrated its ability to handle the non-linear and complex nature of temperature fluctuations, making it a valuable tool for climate change analysis.

In future work, we plan to implement a strategy for real-time temperature forecasting-based climate change monitoring and deployment of proposed models in operational settings. Moreover, we will investigate the transferability of LSTM-NN, BiLSTM-NN, and GRU-NN models to other domains beyond temperature forecasting, such as energy consumption or environmental monitoring. These directions of future work make the field of temperature forecasting using LSTM-NN, BiLSTM-NN, and GRU-NN models continue to evolve, providing more reliable and accurate predictions or forecasts for a wide range of applications.

## Acknowledgment

The authors extend their appreciation to Prince Sattam bin Abdulaziz University for funding this research work through the project number (PSAU/2023/01/ 27466).

## References

[1] D. P. Loucks, "Impacts of climate change on economies, ecosystems, energy, environments, and human equity: A systems perspective," in The impacts of climate change: Elsevier, 2021, pp. 19-50.

[2] T. M. Lee, E. M. Markowitz, P. D. Howe, C.-Y. Ko, and A. A. Leiserowitz, "Predictors of public climate change awareness and risk perception around the world," Nature climate change, vol. 5, no. 11, pp. 1014-1020, 2015.

[3] K. Furtak and A. Wolińska, "The impact of extreme weather events as a consequence of climate change on the soil moisture and on the quality of the soil environment and agriculture–A review," Catena, vol. 231, p. 107378, 2023.

[4] S. Zia, "Climate Change Forecasting Using Machine Learning SARIMA Model," iRASD Journal of Computer Science Information Technology, vol. 2, no. 1, pp. 01-12, 2021.

[5] [5] P. Lynch, "The origins of computer weather prediction and climate modeling," Journal of computational physics, vol. 227, no. 7, pp. 3431-3444, 2008.

[6] D. McNeall, P. R. Halloran, P. Good, and R. A. Betts, "Analyzing abrupt and nonlinear climate changes and their impacts," Wiley Interdisciplinary Reviews: Climate Change, vol. 2, no. 5, pp. 663-686, 2011.

[7] O. Y. Mohammed, H. I. Abed, and N. A. Sultan, "Design and Implementation of Machine Learning and Big Data Analytics models for Cloud Computing platforms," International Journal of Intelligent Systems Applications in Engineering, vol. 11, no. 6s, pp. 185-192, 2023.

[8] F. Kiani and Ö. F. Saraç, "A novel intelligent traffic recovery model for emergency vehicles based on context-aware reinforcement learning," Information Sciences, vol. 619, pp. 288-309, 2023.

[9] F. D. Khangahi and F. Kiani, "Social Mobilization and Migration Predictions by Machine Learning Methods: A study case on Lake Urmia," International Journal of Innovative Technology Exploring Engineering, vol. 10, no. 6, pp. 123-127, 2021.

[10] J. Sun, K. Xiao, C. Liu, W. Zhou, and H. Xiong, "Exploiting intra-day patterns for market shock prediction: A machine learning approach," Expert Systems with Applications, vol. 127, pp. 272-281, 2019.

[11] A. L'heureux, K. Grolinger, H. F. Elyamany, and M. A. Capretz, "Machine learning with big data: Challenges and approaches," IEEE Access, vol. 5, pp. 7776-7797, 2017.

[12] Y. Radhika and M. Shashi, "Atmospheric temperature prediction using support vector machines," International Journal of Computer Theory Engineering, vol. 1, no. 1, p. 55, 2009.

[13] P. Kabbilawsh, D. Sathish Kumar, and N. Chithra, "Trend analysis and SARIMA forecasting of mean maximum and mean minimum monthly temperature for the state of Kerala, India," Acta Geophysica, vol. 68, no. 4, pp. 1161-1174, 2020.

[14] I. P. Pais, F. H. Reboredo, J. C. Ramalho, M. F. Pessoa, F. C. Lidon, and M. M. Silva, "Potential impacts of climate change on agriculture-A review," Emirates Journal of Food Agriculture, pp. 397-407, 2020.

[15] J. E. Cinner, I. R. Caldwell, L. Thiault, J. Ben, J. L. Blanchard, M. Coll, A. Diedrich, T. D. Eddy, J. D. Everett, and C. Folberth, "Potential impacts of climate change on agriculture and fisheries production in 72 tropical coastal communities," Nature communications, vol. 13, no. 1, p. 3530, 2022.

[16] R. Patrick, T. Snell, H. Gunasiri, R. Garad, G. Meadows, and J. Enticott, "Prevalence and determinants of mental health related to climate change

in Australia," Australian New Zealand Journal of Psychiatry, vol. 57, no. 5, pp. 710-724, 2023.

[17] K. L. Ebi, J. Vanos, J. W. Baldwin, J. E. Bell, D. M. Hondula, N. A. Errett, K. Hayes, C. E. Reid, S. Saha, and J. Spector, "Extreme weather and climate change: population health and health system implications," Annual review of public health, vol. 42, no. 1, pp. 293-315, 2021.

[18] J. A. Rising, C. Taylor, M. C. Ives, and R. E. Ward, "Challenges and innovations in the economic evaluation of the risks of climate change," Ecological Economics, vol. 197, p. 107437, 2022.

[19] S. Sen Roy and S. Sen Roy, "Climate change in the global south: trends and spatial patterns," Linking gender to climate change impacts in the Global South, pp. 1-25, 2018.

[20] R. Kršmanc, A. Š. Slak, and J. Demšar, "Statistical approach for forecasting road surface temperature," Meteorological applications, vol. 20, no. 4, pp. 439-446, 2013.

[21] L. Yao, R. Ma, and H. Wang, "Baidu index-based forecast of daily tourist arrivals through rescaled range analysis, support vector regression, and autoregressive integrated moving average," Alexandria Engineering Journal, vol. 60, no. 1, pp. 365-372, 2021.

[22] G. P. Witaradya and Y. T. Putranto, "The Effectiveness of Autoregression to Predict Temperature," in 2023 International Seminar on Application for Technology of Information and Communication (iSemantic), 2023, pp. 276-280: IEEE.

[23] S. Zakaria, N. Al-Ansari, S. Knutsson, and T. Al-Badrany, "ARIMA Models for weekly rainfall in the semi-arid Sinjar District at Iraq," Journal of Earth Sciences Geotechnical Engineering, vol. 2, no. 3, 2012.

[24] P. Chen, A. Niu, D. Liu, W. Jiang, and B. Ma, "Time series forecasting of temperatures using SARIMA: An example from Nanjing," in IOP Conference Series: Materials Science and Engineering, 2018, vol. 394, p. 052024: IOP Publishing.

[25] M. Murat, I. Malinowska, M. Gos, and J. Krzyszczak, "Forecasting daily meteorological time series using ARIMA and regression models," International agrophysics, vol. 32, no. 2, 2018.

[26] D. Dwivedi, G. Sharma, and S. Wandre, "Forecasting mean temperature using SARIMA Model for Junagadh City of Gujarat," IJASR, vol. 7, no. 4, pp. 183-194, 2017.

[27] J. Asha and S. Rishidas, "Forecasting performance comparison of daily maximum temperature using ARMA based methods," in Journal of Physics: Conference Series, 2021, vol. 1921, no. 1, p. 012041: IOP Publishing.

[28] K. Hennayake, R. Dinalankara, and D. Y. Mudunkotuwa, "Machine learning based weather prediction model for short term weather prediction in Sri Lanka," in 2021 10th International Conference on Information and Automation for Sustainability (ICIAfS), 2021, pp. 299-304: IEEE.

[29] K. N. Mitu and K. Hasan, "Modeling and Forecasting Daily Temperature Time Series in the Memphis, Tennessee," International Journal of Environmental Monitoring and Analysis, vol. 9, no. 6, pp. 214-221, 2021.

[30] T. Dimri, S. Ahmad, and M. Sharif, "Time series analysis of climate variables using seasonal ARIMA approach," Journal of Earth System Science, vol. 129, pp. 1-16, 2020.

[31] A. Gangshetty, G. Kaur, and U. Malunje, "Time Series Prediction of Temperature in Pune using Seasonal ARIMA Model," International Journal of Engineering Research Technology, vol. 10, no. 11, 2021.

[32] D. Hoang, P. L. Yang, L. Cuong, P. Trung, N. Tu, L. Truong, T. Hien, and V. Nha, "Weather prediction based on LSTM model implemented AWS Machine Learning Platform," International Journal for Research in Applied Science Engineering Technology, vol. 8, no. 5, pp. 283-290, 2020.

[33] W. Jaharabi, M. Hossain, R. Tahmid, M. Z. Islam, and T. Rayhan, "Predicting Temperature of Major Cities Using Machine Learning and Deep Learning," arXiv preprint arXiv:.13330, 2023.

[34] H. Koçak, "Time Series Prediction of Temperature Using Seasonal ARIMA and LSTM Models," Gazi Mühendislik Bilimleri Dergisi, vol. 9, no. 3, pp. 574-584, 2023.

[35] A. Khokhar, S. Talpur, and M. Memon, "Comparative Analysis of LSTM, BILSTM and ARIMA for Time Series Forecasting on 116 years of Temperature and Rainfall Data from Pakistan," International Journal of Scientific Research in Computer Science, Engineering Information Technology, pp. 350-357, 2023.

[36] S. Jafarian-Namin, D. Shishebori, and A. Goli, "Analyzing and Predicting the Monthly Temperature of Tehran using ARIMA Model, Artificial Neural Network, and Its Improved Variant," Journal of Applied Research on Industrial Engineering, 2023.

[37] I. H. Topalova and P. G. Radoyska, "Neural Network Structure for Tracking the Climate Temperature Change," Preprint, 2023.

[38] B. Earth, "Climate change: earth surface temperature data," ed: Kaggle, 2019.

[39] [39] G. P. Zhang, "An investigation of neural networks for linear time-series forecasting," Computers Operations Research, vol. 28, no. 12, pp. 1183-1202, 2001.