A Novel Combined Model for Prediction of Daily Precipitation Data using Instantaneous Frequency Feature and Bidirectional Long Short Time Memory Networks

Levent Latifoğlu (latifoglu@erciyes.edu.tr)
Erciyes Üniversitesi: Erciyes Universitesi https://orcid.org/0000-0002-2837-3306

Research Article

Keywords: Bidirectional Long Short Time Memory Networks, biLSTM, Long Short Time Memory Networks, LSTM, Gated Recurrent Unit, GRU, Instantaneous frequency, Daily precipitation

Posted Date: June 10th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-525276/v1

Version of Record: A version of this preprint was published at Environmental Science and Pollution Research on January 29th, 2022. See the published version at https://doi.org/10.1007/s11356-022-18874-z.

A Novel Combined Model for Prediction of Daily Precipitation Data using Instantaneous Frequency Feature and Bidirectional Long Short Time Memory Networks

[1]Levent LATİFOĞLU

[1]Department of Civil Engineering, Faculty of Engineering, Erciyes University, Turkey

Corresponding Author

Assist. Prof. Dr. Levent LATIFOGLU

Erciyes University, Engineering Faculty

Civil Engineering Deptartment

38039, Kayseri / TURKEY

E-mail : latifoglu@erciyes.edu.tr

Phone : + 90 (352) 2076666 extension: 32375

ABSTRACT

In the developing world, to learn nature better, to get the maximum benefit from nature is being studied. Meteorological events constantly affect human life. The occurrence of excessive precipitation in a short time causes important events such as floods. However, in case of insufficient precipitation for a long time, drought occurs. In recent years, significant changes in precipitation regimes have been observed and these changes cause socioeconomic and ecological problems. Therefore, it is of great importance to correctly predict and analyze these variables.

In this study, reliable and accurate precipitation forecasting model is proposed. Ensemble of instantaneous frequency (IF) Bidirectional Long Short Time Memory Networks (biLSTM) model was employed for the aim of forecasting of daily precipitation data. To compare the performance of biLSTM model, Long Short Time Memory Networks (LSTM) and Gated Recurrent Unit (GRU) model was applied for forecasting of daily precipitation data. The performance of the proposed IF-biLSTM model was evaluated using Mean absolute error (MAE), Mean square Error (MSE), Correlation Coefficient (R) and Determination Coefficient (R^2) performance parameter. According to numerical results, IF-biLSTM model has the best forecasting performance for daily precipitation data. Especially six ahead precipitation forecasting is noteworthy.

Keywords: Bidirectional Long Short Time Memory Networks, biLSTM, Long Short Time Memory Networks, LSTM, Gated Recurrent Unit, GRU, Instantaneous frequency, Daily precipitation

INTRODUCTION

Water is the source of the life and the most important element for the survival of the life. Precipitation occurs when the moisture in the atmosphere condenses and returns to the earth in different conditions. There are many factors that affect precipitation. Factors such as pressure, temperature and wind in the atmosphere with the effect of global changes also affect the amount of precipitation over the years (Wei et al. 2005). Hydrology is known as water science a kind of an applied science. It examines the temporal and spatial distribution of water on the earth, its physical, chemical and biological properties, and its mutual relations with the environment and living things (Wu et al. 2012). Planning, projecting and operating water resources, which are of vital importance to human beings, are the leading hydrological studies (Price 2013). The economical use of water resources is of vital importance in many areas such as the continuation of vitality, the need for agricultural irrigation, and electricity generation. Therefore, the need for hydrology science is very important. While benefiting from water resources, which are the leading hydrological studies, many parameters (precipitation, stream flow, infiltration, evaporation, vegetation, etc.) should be analyzed truly and their effects should be examined in order to use the water resources correctly and to project the water structures to be built correctly. Precipitation is one of the most important of these parameters that generates the flow. Accurate measurement of this parameter is very important in the management and operation of water resources (Hamill et al. 2006; Price 2013; Wei et al. 2005; Wu et al. 2012). Precise estimation of precipitation is often difficult, as the exact mechanisms that affect its occurrence are not known. It is much more important to accurately forecast precipitation, especially on a small time scale such as daily or hourly. However, precipitation is a difficult subject to analyze due to its high complexity, non-stationary, non-linear and dynamic internal structure. Therefore, improving precision in precipitation prediction is a significant topic for hydrologists, and research on precipitation prediction methods has an increasing importance.

Studies for hydrometeorological parameter estimation were first based on linear approaches, and models such as the parametric Autoregressive Moving Average (ARMA) and the Autoregressive Integrated Moving Average (ARIMA) have been introduced since the 1970s to analyze time series data (Box et al. 2015). These techniques are basically linear models and model the time series assuming that it is stationary and linear. For this reason, forecasting of hydrometerological data studies are developed based on artificial intelligence techniques due to the non-stationary and non-linear character of this data in the last two decades. These approaches are based on machine learning techniques and as the basic classifier technique,

Artificial Neural Networks, Decision Trees and Decision Support Machines (Du et al. 2017; Retalis et al. 2017; Le et al. 2020; Nourani et al. 2027; Nourani et al. 2020). These methods have been found to obtain satisfactory forecasts capturing nonlinear property of hydrological and meteorological processes. In addition to traditional machine learning techniques, Deep Neural Networks (DNN) is recently introduced as an effective modeling method for mapping nonlinear relationships, and started to be applied in many different areas (Bashar 2019). Long Short-Term Memory Networks (LSTM) are DNN architectures designed by Hochreiter and Schmidhuber (1997) to learn the long-term dependencies of time series through gate and memory units (Hochreiter et al. 1997). This is a type of Deep Neural Networks and is based on Recurrent Neural Networks architecture. LSTM architecture is developed to eliminate some of the disadvantages that occur in RNN architectures, such as vanishing gradient problem and restricted the memory capabilities. LSTM was used for forecasting of river flood (Le 2019), air pollution forecasting [Freeman et al. 2018; Yu et al. 2019), fog (Miao et al. 2020), wind power (Shahid et al. 2021) and modeling of rainfall-runoff processes (Schuster et al. 1997). Recently, Gated Recurrent Unit (GRU) and Bidirectional LSTM (biLSTM) networks which are LSTM's variant are utilized analysis of long-term dependencies (Cho er al. 2014; Schuster et al. 1997). GRU neural network is a newly developed gating mechanism and is successfully employed in many fields, such as short-term electricity load forecasting method based on multilayered self-normalizing GRU network (Kuan et al. 2017) heat load forecast (Lu et al. 2018) and electricity generation and planning (Li et al. 2018). Also, biLSTM model is applied for the aim of time series forecasting (Siami-Namini et al. 2019), financial time series (Siami-Namini et al. 2019) and forecasting of trading area (Kim et al. 2019). In this study daily precipitation data was forecasted for one to six day ahead using biLSTM method. Also the effect of the instantaneous frequency feature on the forecasting performance is analyzed and forecasting performance of biLSTM model is compared with GRU and LSTM models. Novelty of this paper is that, it is the first study in the literature to predict daily precipitation using the instantaneous frequency feature and biLSTM. To the best of our knowledge there isn't any study in the literature, forecasting of daily precipitation data with high performance using proposed Instantaneous frequency and biLSTM (IF-biLSTM) model.

The rest of the paper is organized as follows. Section 2 gives information about study area and the data. Section 3 provides a brief review of the LSTM, GRU, biLSTM and instantaneous frequency methods for daily precipitation estimation. Section 4 describes the prediction results obtained by the LSTM, GRU, biLSTM and IF-biLSTM. Finally, Section 5 concludes the paper.

Application of proposed novel model using IF-biLSTM to the precipitation data for forecasting is seen in Figure 1 and the flow chart of the study is expressed in Materials and Methods section.

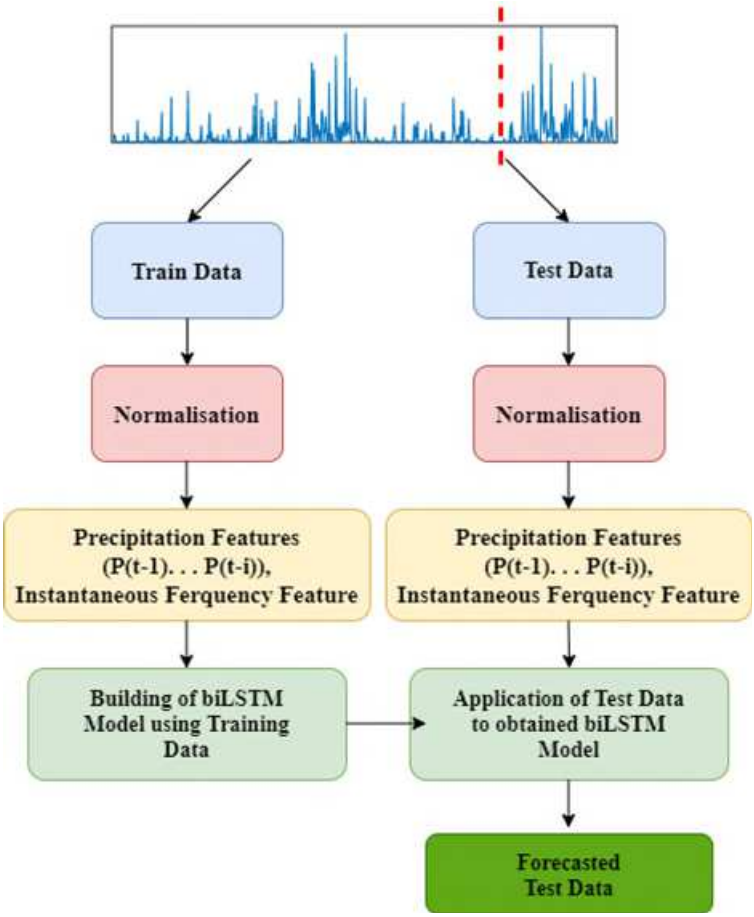


Figure 1. The proposed DNN forecasting model for daily precipitation data

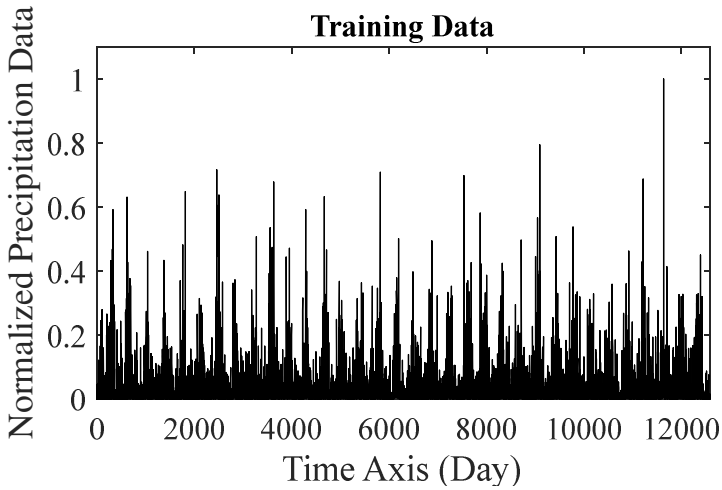
2. MATERIALS AND METHODS

2.1 STUDY AREA AND DATA

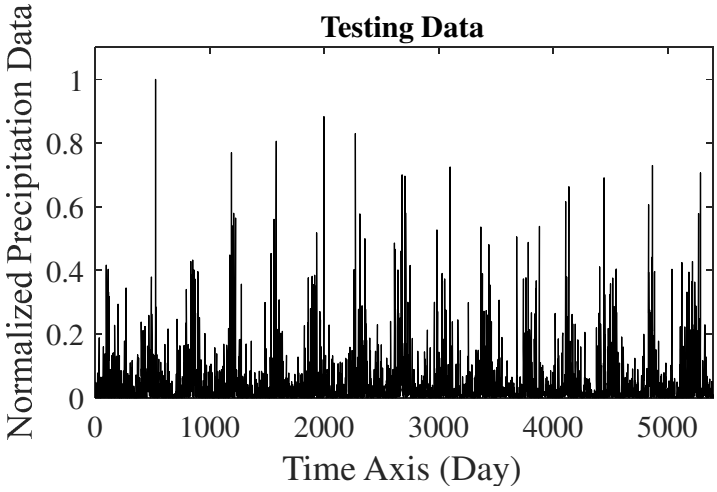
The daily precipitation (mm) data has been continuously gauged over the 49 year period of, between 1963 through 2012, hence consisting of 17990 number of data are used as the material of this study. Region where data was gauged is from the Curchill River at Churchill River above Otter Rapids basin lies between 55°38'51.0"N 104°44'09.0"W (latitude: 55.647499, longitude: -104.735832) in Saskatchewan Province in Canada. Drainage basin from which the data was recorded is shown in Figure 2. This data set is obtained from CANOPEX database (http://canopex.etsmtl.net/, Arsenault et al. 2016). The drainage area at this site is 114248 km^2.

The daily precipitation data was divided into a 70:30 ratio, where 70% of data was employed for training the model, and the remaining 30% data was utilized for testing the effectiveness of the model. Therefore, 12593 elements of the precipitation data was used for the training phase and the rest of 5397 elements of the precipitation data was used for the testing stage. Training and testing data are normalized as shown in Figure 3.

Figure 2: Map of Churchill River Above Otter Rapids Station



a



b

Figure 3. Daily Precipitation data measured on Churchill River above Otter Rapids basin a- Training Data, b- Testing Data

2.2. LONG SHORT-TERM MEMORY NETWORKS (LSTM)

LSTM Networks are a special type of Recurrent Neural Networks (RNN) that can learn long-term dependencies through special hidden units called memory cells, which are used to remember the previous input for a long time. It has been first introduced by Hochreiter & Schmidhuber (1997). It is developed in different versions in later times (Hochreiter et al. 1997). LSTM model can capture nonlinear trends in data and recall long-term information. Therefore, LSTM is successfully applied to many kinds of time series problems. As shown in Figure 4, there are three gates including input, forget and output gate in the LSTM unit composed of a cell.

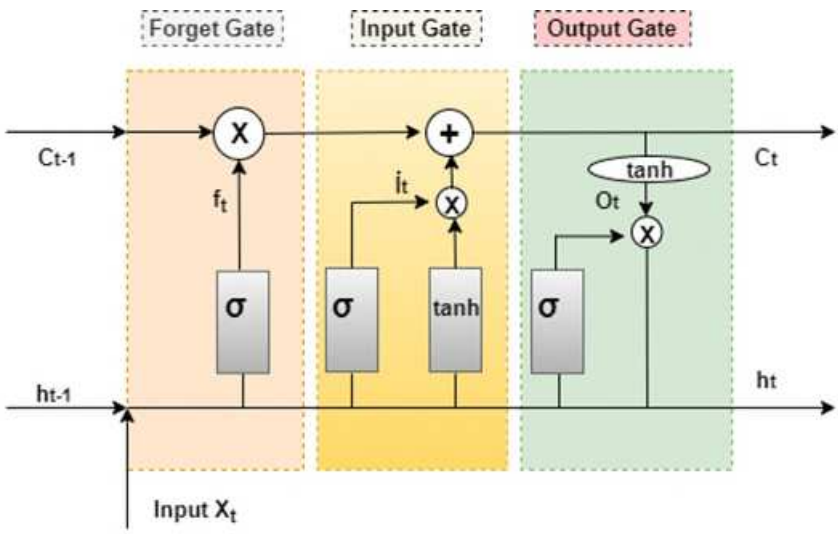


Figure 4. LSTM Unit

The main information flow of the LSTM memory cell (Figure 3) can be expressed mathematically. 'x' and '+' symbols indicate the addition and multiplication operations in the model. The flow direction of the information is shown by the arrow. The first layer of the memory unit decides to remove unnecessary information from the cell state. This decision is made with an operation denoted by forgetting gate and expressed by Equation 1. Here, C_{t-1} is assigned a value between 0 and 1 according to the cell status. The output of the forgetting gate is shown in f_t as shown in Equation 1 (Hochreiter et al. 1997). The forgetting gate in LSTM unit gives a certain weight (*W*) to long-term memory information.

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

The input gate layer is the layer where it is decided what new information to store in the cell state. This layer receives the input and learns new information along with the information previously learned from short-term memory. It consists of two parts. The entrance gate part of the input layer is a sigmoid layer called (i_t), which is the layer that decides which values to update. Equation gives the output expressed by Equation 2. The tanh layer of the input layer is the layer in which new candidate values (C_y) are formed and it is expressed by Equation 3. These two outputs are combined to create an update.

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

$$C_y = tanh(W_c.[h_{t-1}, x_t] + b_c) \quad (3)$$

C_t update status in LSTM networks, new values are created according to the information from other layers. Here, the long-term memory is updated by adding the new information learned to the parts coming from the long-term memory. The update status is determined by adding the forget gate layer and the input gate layer values.

$$C_t = f_t * C_{t-1} + i_t * C_y \quad (4)$$

In the last layer of LSTM networks, as shown in Equation 5, firstly, the inputs (x_t, h_{t-1}) are passed through a sigmoid layer that decides how much of the cell state will affect the output. The cell state is passed through the tanh activation function as shown in Equation 6 and multiplied by the output of the output gate.

$$O_t = \sigma(W_o.[h_{t-1}, x_t] + b_o) \quad (5)$$

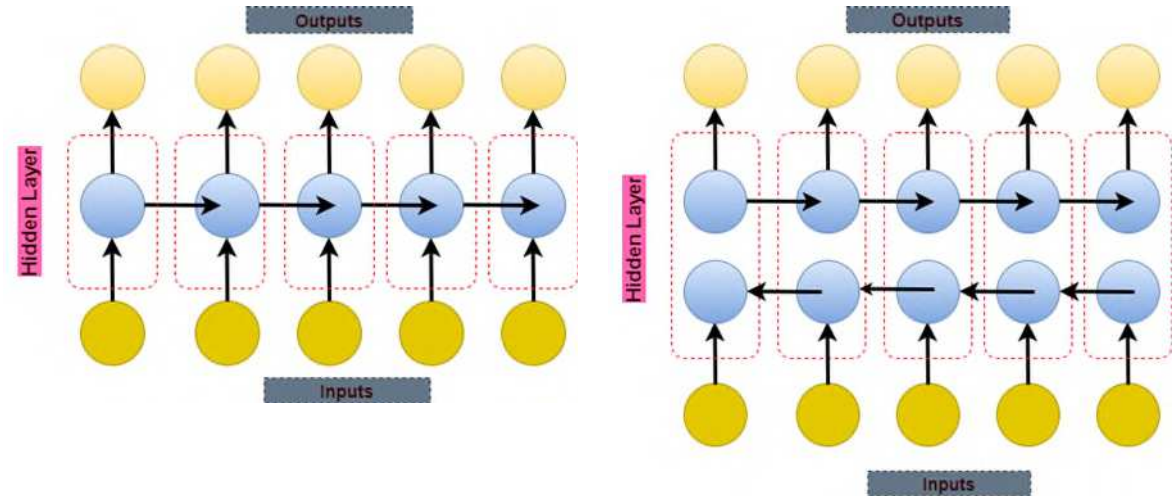
$$h_t = O_t * tanh(C_t) \quad (6)$$

The functions used in the LSTM unit (sigmoid (σ), hyperbolic tangent ($tanh$), product (x) and sum (+) are differentiable, as can be seen in Figure 3. Therefore, weights can be updated by derivatives in the backpropagation process.

2.3. BIDRECTIONAL LONG SHORT-TERM MEMORY NETWORKS (biLSTM)

A Bidirectional LSTM or biLSTM is a DNN model and the extended version of ordinary LSTM. This architecture has two LSTM units. In LSTM information flow is unidirectional while in biLSTM it is bidirectional as seen in Figure 5. One of them takes the input in the forward direction, the other in the back direction. biLSTMs improve the content available for the algorithm, effectively increasing the amount of information available on the network (Schuster et al. 1997).

biLSTM can simultaneously memorize long-term dependencies and process the information bidirectionally.



a **b**

Figure 5.a. LSTM architecture **Figure 5.b. biLSTM architecture**

2.4. GATED RECURRENT UNIT (GRU)

Today, different variations of LSTM architecture is introduced. The most commonly used of these is the Gated Recurrent Unit, or GRU. In the GRU architecture, forget gate and input gates are combined. It has less complexity compared to standard LSTM models. The GRU unit is given in the Figure 6 below. The main difference between a GRU and a LSTM is that a GRU has two gates (reset and update gates), an LSTM has three gates (input, forget and output gates). The output of h_t in the GRU unit is defined by the equations given below (Cho et al. 2014).

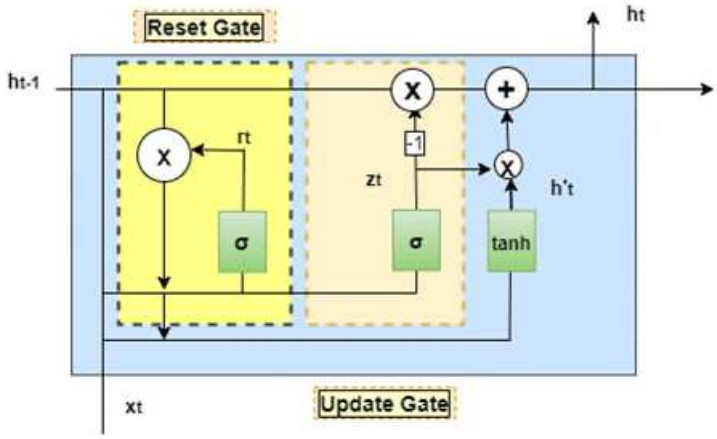


Figure 6. GRU Unit

It is assumed that r_t and z_t denote the reset and update gates. The update gate acts similar to the forget and input gate of an LSTM. It decides what information to throw away and what new information to add. The mathematical formula can be expressed as follows:

$$z_t = \sigma(W_z.[h_{t-1}, x_t]) \tag{7}$$

$$r_t = \sigma(W_r.[h_{t-1}, x_t]) \tag{8}$$

Where: σ represents the sigmoid function, x_t and h_t the variables represent the current input and the output of the GRU unit. Also h_{t-1} is hidden state at t-1 time.

$$h'_t = tanh(W.[r_t * h_{t-1}, x_t) \tag{9}$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * h'_t) \tag{10}$$

Where h'_t is the candidate state and tanh is the hyperbolic tangent function. If the r_t reset gate is closed, then the GRU will ignore the previous hidden state h_{t-1}. In this case output is affected only by the current input x_t. Also, the update gate z_t controls how much information of the past state h_{t-1} can be passed to the current state h_t.

It is important that the number of LSTM layer and the number of cells in each layer for forecasting performance and computation time. In the development of the DNN architecture, different number and order of layers was tried. DNN architecture giving the best performance and having the lowest processing load was developed. The architectures realized with LSTM, biLSTM and GRU in this study are shown in Figure 7.

The first layer in architecture consists of a sequence input layer, which is used to input the daily precipitation data into the network.

The second layer consists of LSTM, biLSTM and GRU layer for each model.

In this layer number of 32, 64, 128 and 256 memory unit was tried during development of the model and the best forecasting performance was obtained with number of 128 memory unit for LSTM, biLSTM and GRU model.

The third layer of the model is the Rectified Linear Unit (ReLU) layer. This activation function is the most commonly used function in Deep Neural Networks. This layer is also known as the activation layer. The effect it has on the input data is that it makes the negative values to zero. The fourth layer of the model is fully connected layer. In this layer, data from previous layers are combined by weighting and a loss function and the optimal weight to be given to neurons during training is found. In this layer 10-100 units were tried during construction of the model and the most reasonable unit value was obtained as 10.

The fifth layer of the model is dropout layer which is used to forget some neurons in order to prevent overfitting during training. In this study dropout was applied as 50%.

The sixth level of the model is fully connected layer. The output of this layer was defined as one.

The model is lasted with regression output layer.

In the training phase of the network, the maximum number of epochs was 100, the initial learning rate was 0.002, the learning rate drop range was 100, and the learning rate drop factor was 0.1. These values are decided by trial and error method. 'Adam' optimization algorithm was used for the training of the network.

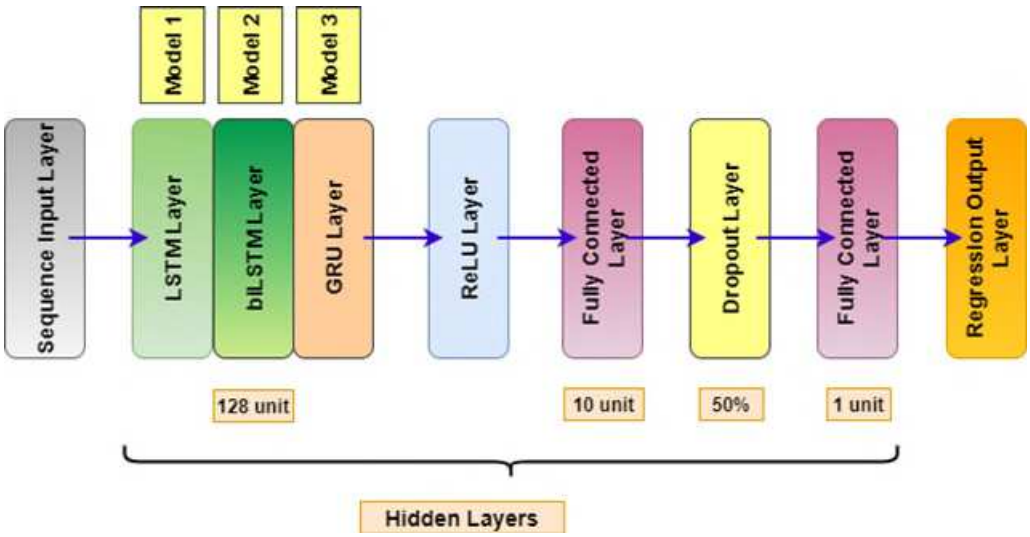


Figure 7. Forecasting models using LSTM Model, biLSTM and GRU Model

2.5. INPUTS FOR DEEP NEURAL NETWORKS

Deep neural networks are trained by input data and output data. In this study sequential data of previous rainfall and instantaneous frequency features are applied as input variables shown in Table 1. One to six day ahead forecasting using one-three inputs and after with instantaneous frequency feature.

Table 1. Input data, and output used in our experiments

	One Ahead Forecast			Four Ahead Forecast	
Forecasting with Daily Precipitation Inputs					
Input Number	**Inputs**	**Output**	**Input Number**	**Inputs**	**Output**
One Input	P(t-1)	P(t)	One Input	P(t-4)	P(t)
Two Input	P(t-2), P(t-1)	P(t)	Two Input	P(t-5), P(t-4)	P(t)
Three Input	P(t-3), P(t-2), P(t-1)	P(t)	Three Input	P(t-6), P(t-5), P(t-4)	P(t)
	Two Ahead Forecast			**Five Ahead Forecast**	
One Input	P(t-2)	P(t)	One Input	P(t-5)	P(t)
Two Input	P(t-3), P(t-2)	P(t)	Two Input	P(t-6), P(t-5)	P(t)
Three Input	P(t-4), P(t-3), P(t-2)	P(t)	Three Input	P(t-7), P(t-6), P(t-5)	P(t)
	Three Ahead Forecast			**Six Ahead Forecast**	
One Input	P(t-3)	P(t)	One Input	P(t-6)	P(t)
Two Input	P(t-4), P(t-3)	P(t)	Two Input	P(t-7), P(t-6)	P(t)
Three Input	P(t-5), P(t-4), P(t-3)	P(t)	Three Input	P(t-8), P(t-7), P(t-6)	P(t)
Forecasting with Daily Precipitation Inputs and Instantaneous Frequency					
Input Number	**Inputs**	**Output**	**Input Number**	**Inputs**	**Output**
	One Ahead Forecast			**Four Ahead Forecast**	
Two Input	P(t-2), P(t-1), IF	P(t)	Two Input	P(t-5), P(t-4), IF	P(t)
	Two Ahead Forecast			**Five Ahead Forecast**	
Two Input	P(t-3), P(t-2), IF	P(t)	Two Input	P(t-6), P(t-5), IF	P(t)
	Three Ahead Forecast			**Six Ahead Forecast**	
Two Input	P(t-4), P(t-3) , IF	P(t)	Two Input	P(t-7), P(t-6) , IF	P(t)

2.5.1. HILBERT TRANSFORM AND INSTANTANEOUS FREQUENCY

The instantaneous frequency is crucial in numerous signal processing applications and represents one of the most important parameters in the time-frequency analysis for modelling and classification of signals. Hilbert transform is applied to real valued x(t) signal to obtain instantaneous frequency. The Hilbert transform is a specific linear operator and defined as below equation:

$$H(x(t)) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{x(\tau)}{t-\tau} d\tau \qquad (11)$$

According to Equation 11, this linear operator is given by convolution with the function $\frac{1}{\pi t}$. Convolution operation with an x(t) signal in time domain imparts a phase shift of ±90° to every frequency component of the signal in frequency domain (Johansson 1999).

As a result of Hilbert transform, it is possible to obtain the analytic representation of a real valued x(t) signal. An analytic complex valued X(t) signal can be constructed from a real valued input signal x(t) as seen in Equation 12.

$$X(t) = x(t) + jh(t) \qquad (12)$$

Where, X(t) is the analytic signal obtained from x(t) and its Hilbert transform h(t). Also x(t) signal can be expressed in polar coordinates.

$$X(t) = A(t)e^{j\theta(t)} \quad (13)$$

Where A(t) is defined as the envelope or amplitude of the analytic signal and $\theta(t)$ is defined as the phase of analytic signal.

The derivative of the phase named as $\theta(t)$ of the analytic signal X(t) is called instantaneous frequency. Instantaneous frequency is defined as below equation.

$$w_i(t) = \frac{d\theta(t)}{dt} \quad (14)$$

In this study instantaneous frequency feature and previous two daily precipitation data were applied in the forecasting of daily precipitation data.

2.6. PERFORMANCE INDICATORS

In the proposed study, forecasting performance of IF-biLSTM model has been shown using Mean Absolute Error, The Mean Square Error,

The Mean Absolute Error (MAE):

Average absolute error is the measure of the difference between the observed time series data and the forecasted data by the proposed model. MAE is defined below equation:

$$MAE = \frac{1}{N} \sum_{i=1}^{N} |X_{observed,i} - Y_{estimated,i}| \quad (15)$$

The Mean Square Error (MSE): The Mean Square Error represents the difference between observed time series data and forecasted data by the proposed model extracted by squared the average difference over the data set. MSE is defined as below equation:

$$MSE = \frac{1}{N} \sum_{i=1}^{N} (X_{observed,i} - Y_{estimated,i})^2 \quad (16)$$

The Correlation Coefficient (R): The correlation coefficient reveals the degree, direction and importance of the relationship between observed time series data and forecasted data by the proposed model. The correlation coefficient is denoted by the R and takes a value between [−1, 1]. R value is defined as below equation:

$$R = \frac{1}{N-1} \sum_{i=1}^{N} \left(\frac{X_{observed,i} - \mu_X}{\sigma_X}\right) \left(\frac{Y_{estimated,i} - \mu_Y}{\sigma_Y}\right) \quad (17)$$

In this equation, $X_{observed,i}$ is the observed time series data, $\mu_X$the average and σ_X the standard deviation of the observed time series data$Y_{estimated,i}$, is estimated data μ_Y is the average and σ_y the standard deviation of the estimated data.

The Determination Coefficient (R^2):

R^2 is often used to evaluate the predictive power of used hydrological models. This statistical criterion takes a value between -∞ and 1. If the R^2 determination coefficient value is one between the actual and estimated data, it means that excellent results have been obtained. R^2 value is defined as below equation:

$$R^2 = 1 - \frac{\sum_{i=1}^{N}[X_{observed,i} - Y_{estimated,i}]^2}{\sum_{i=1}^{N}[X_{observed,i} - \mu_X]^2} \quad (18)$$

3. RESULTS and DISCUSSION

3.1. PERFORMANCE EVALUATION FOR LSTM, biLSTM and GRU FORECASTING MODEL

In the first stage of the study, to evaluate the forecasting performance of LSTM, biLSTM and GRU models, one to three previous daily precipitation data was applied to models as input variables and one day ahead forecasting was performed. As it can be seen in Table 2, the biLSTM model shows better forecasting performance compared to the LSTM and GRU models.

Table 2. One day ahead precipitation forecasting performance of LSTM, biLSTM and GRU models

ONE AHEAD FORECASTING				
LSTM MODEL				
	MSE (mm^2)	**MAE (mm)**	**R**	**R^2**
One Input	0.0054	0.0396	0.4011	0.1607
Two Inputs	0.0054	0.0389	0.3997	0.1596
Three Inputs	0.0054	0.0384	0.3993	0.1593
bİLSTM MODEL				
	MSE (mm^2)	**MAE (mm)**	**R**	**R^2**
One Input	**0.0000**	**0.0019**	**0.9994**	**0.9987**
Two Inputs	0.0001	0.0053	0.9968	0.9937
Three Inputs	0.0001	0.0073	0.9915	0.9831
GRU MODEL				
	MSE (mm^2)	**MAE (mm)**	**R**	**R^2**
One Input	0.0054	0.0373	0.4095	0.1677
Two Inputs	0.0054	0.0395	0.3984	0.1588
Three Inputs	0.0054	0.0398	0.3976	0.1581

It is seen that the performance of one input biLSTM model is the best for one ahead estimation of daily precipitation data. Observed and predicted results are shown for one day ahead forecasting model in Figure 8. As clearly observed from scatter plot of one ahead forecasted data and observed data, the linear trend from the biLSTM model is close to the line y=x.

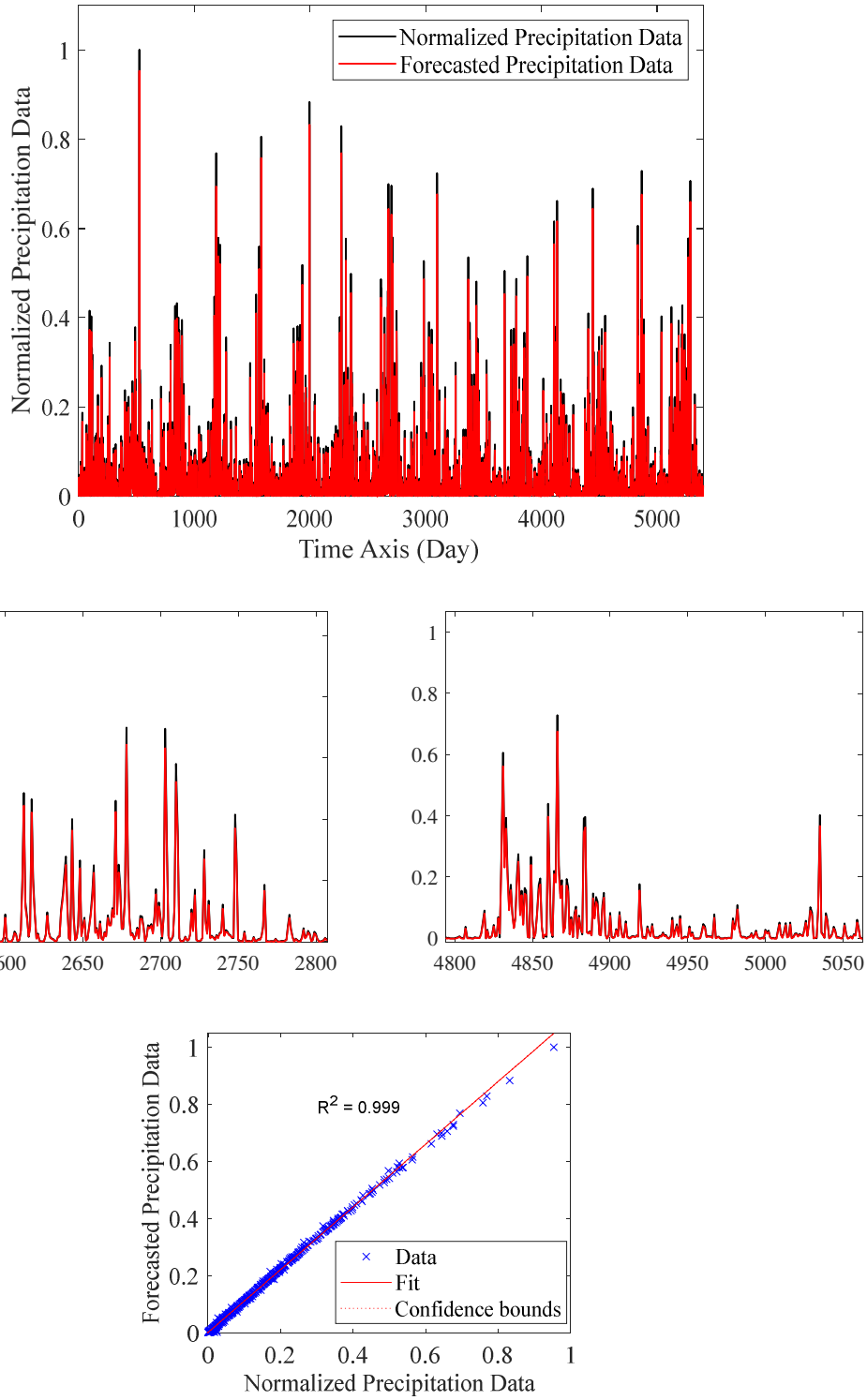


Figure 8. One-ahead forecasting of daily precipitation data using biLSTM Model

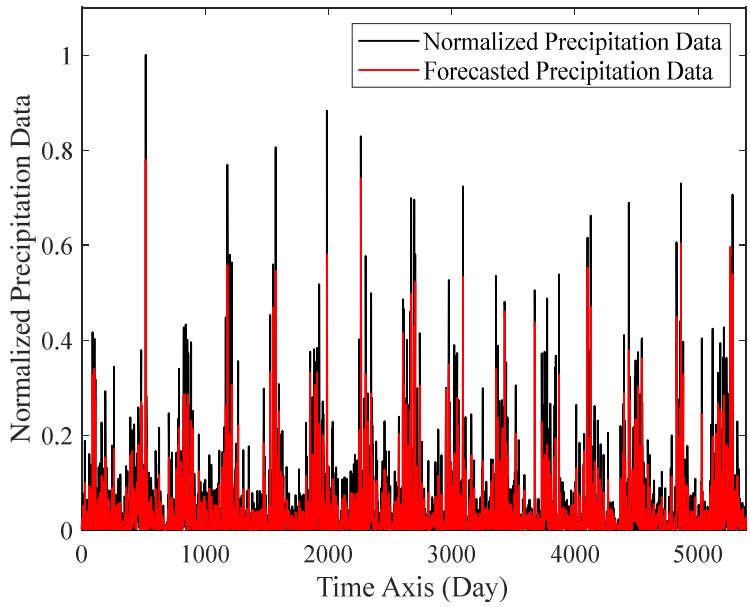
3.2. ONE to SIX AHEAD FORECASTING PERFORMANCE of biLSTM MODEL

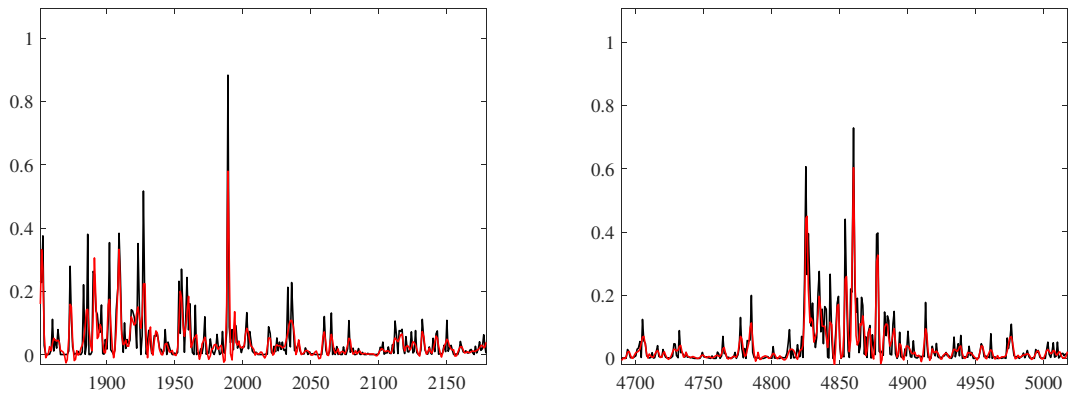
At this stage of the study, one to six day ahead forecasting model was performed using biLSTM method. As can be seen from Table 3, the best forecasting performance is obtained for one ahead forecasting using one input, two-four and six ahead forecasting using two inputs and five

ahead forecasting using three inputs. As an example, the model result of six ahead forecasting using two inputs is shown in Figure 9.

Table 3. One-six ahead precipitation forecasting performance of biLSTM model

ONE AHEAD FORECASTING				
	MSE (mm^2)	**MAE (mm)**	**R**	**R^2**
One Input	**0.0000**	**0.0019**	**0.9994**	**0.9987**
Two Inputs	0.0001	0.0053	0.9968	0.9937
Three Inputs	0.0001	0.0073	0.9915	0.9831
TWO AHEAD FORECASTING				
	MSE (mm^2)	**MAE (mm)**	**R**	**R^2**
One Input	0.0010	0.0188	0.9224	0.8508
Two Inputs	**0.0005**	**0.0152**	**0.9595**	**0.9206**
Three Inputs	0.0007	0.0168	0.9429	0.8890
THREE AHEAD FORECASTING				
	MSE (mm^2)	**MAE (mm)**	**R**	**R^2**
One Input	0.0010	0.0195	0.9161	0.8393
Two Inputs	**0.0008**	**0.0181**	**0.9324**	**0.8694**
Three Inputs	0.0009	0.0180	0.9307	0.8663
FOUR AHEAD FORECASTING				
	MSE (mm^2)	**MAE (mm)**	**R**	**R^2**
One Input	0.0017	0.0242	0.8624	0.7437
Two Inputs	**0.0014**	**0.0232**	**0.8817**	**0.7773**
Three Inputs	0.0014	0.0227	0.8798	0.7741
FIVE AHEAD FORECASTING				
	MSE (mm^2)	**MAE (mm)**	**R**	**R^2**
One Input	0.0019	0.0256	0.8404	0.7062
Two Inputs	0.0017	0.0249	0.8542	0.7297
Three Inputs	**0.0015**	**0.0232**	**0.8742**	**0.7642**
SIX AHEAD FORECASTING				
	MSE (mm^2)	**MAE (mm)**	**R**	**R^2**
One Input	0.0025	0.0294	0.7864	0.6184
Two Inputs	**0.0018**	**0.0254**	**0.8493**	**0.7214**
Three Inputs	0.0021	0.0282	0.8179	0.6690





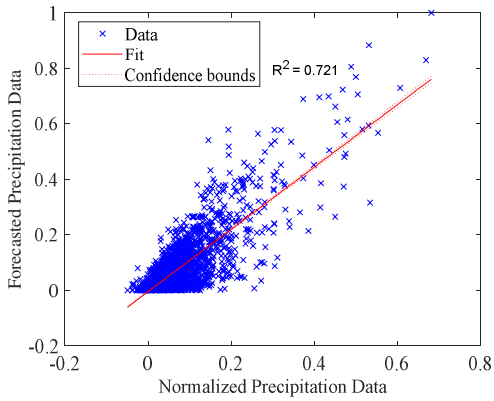


Figure 9. Six-ahead forecasting of daily precipitation data using three input biLSTM Model

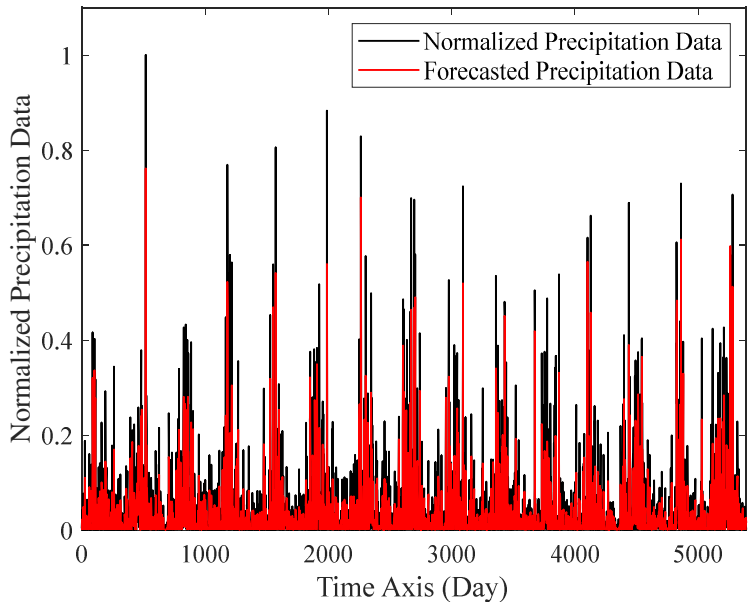
3.3. FORECASTING PERFORMANCE OF biLSTM MODEL USING INSTANTANEOUS FREQUENCY

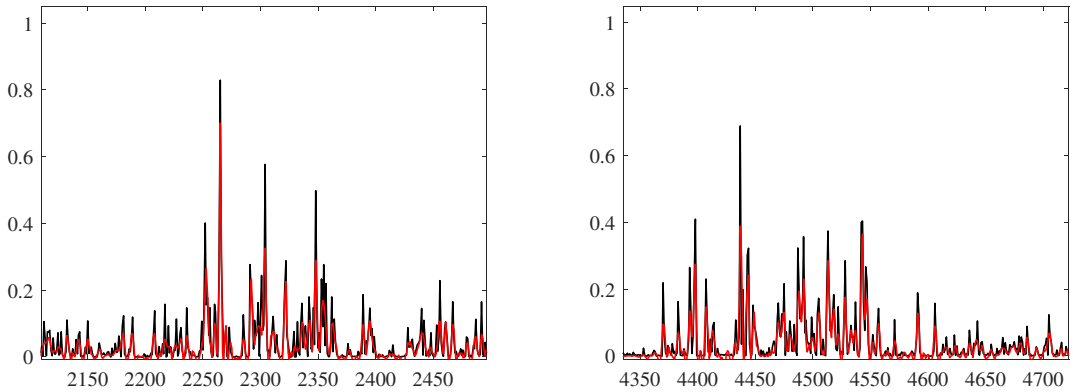
Instantaneous frequency feature is used for analysis of forecasting performance of daily precipitation data. Therefore, instantaneous frequency feature is applied besides two inputs to

biLSTM model. It is seen from Table 4 that, forecasting performance is improved with application of instantaneous features as input to biLSTM model. As an example, six ahead forecasting using IF-biLSTM model is shown in Figure 10.

Table 4. One-six ahead precipitation forecasting performance of IF-biLSTM model

N AHEAD FORECASTING WITH TWO PREVIOUS INPUTS				
	MSE (mm^2)	**MAE (mm)**	**R**	**R^2**
One ahead	0.0001	0.0053	0.9968	0.9937
Two ahead	0.0005	0.0152	0.9595	0.9206
Three ahead	0.0008	0.0181	0.9324	0.8694
Four ahead	0.0014	0.0232	0.8817	0.7773
Five ahead	0.0017	0.0249	0.8542	0.7297
Six ahead	0.0018	0.0254	0.8493	0.7214
N AHEAD FORECASTING WITH TWO PREVIOUS INPUTS+IF				
	MSE (mm^2)	**MAE (mm)**	**R**	**R^2**
One ahead	0.0001	0.0055	0.9992	0.9983
Two ahead	0.0002	0.0081	0.9923	0.9827
Three ahead	0.0006	0.0155	0.9535	0.9092
Four ahead	0.0010	0.0190	0.9224	0.8508
Five ahead	0.0014	0.0225	0.8836	0.7827
Six ahead	0.0016	0.0237	0.8697	0.7563





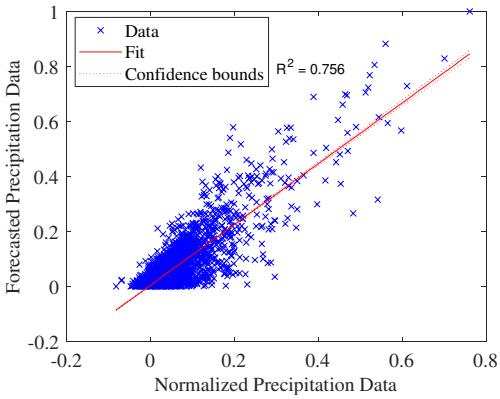


Figure 10. Six-ahead forecasting of daily precipitation data using IF- biLSTM Model

4. CONCLUSIONS

Prediction of daily precipitation is a challenging task because of having nonlinear and nonstationary property of the data. Recently, biLSTM model is used for forecasting aims (Siami-Namini S et al. 2019; Siami-Namini S et al. 2019; Kim et al 2019;Wu et al. 2020) in financial time series and trading area. In this study IF-biLSTM model was employed for daily precipitation estimation. For this aim, firstly, the daily precipitation data was splitted as training and the testing data. The testing data were completely unused (not applied the model) during the training stage of the model.

The LSTM and GRU models were used to compare the biLSTM model performance. The obtained performance parameters indicates that the forecasting performance of the biLSTM model is much better than LSTM and GRU model for one day ahead forecasting.

In this study, biLSTM model was applied for one to six day ahead forecasting. According to numerical results, the biLSTM model forecasting performance with two inputs is better for two, three, four and six day ahead forecasting. To analyze the forecasting performance of instantaneous frequency feature; IF with two previous precipitation data was applied to biLSTM model. It is seen that IF feature improves the forecasting performance of the proposed model.

As seen from Table 3, There are remarkable improvements as an example R^2 parameter starting with one ahead forecasting; R^2 values as 0.994 (without IF feature) to 0.998 (with IF feature), with two ahead forecasting; R^2 values as 0.921 (without IF feature) to 0.983 (with IF feature), with three ahead forecasting; R^2 values as 0.869 (without IF feature) to 0.909 (with IF feature), with four ahead forecasting; R^2 values as 0.777 (without IF feature) to 0.851 (with IF feature), with five ahead forecasting; R^2 values as 0.730 (without IF feature) to 0.783 (with IF feature), with six ahead forecasting; R^2 values as 0.721 (without IF feature) to 0.756 (with IF feature).

A new Deep Neural Network and instantaneous frequency based model called IF-biLSTM is proposed in this study. Thus, it is achieved high forecasting performance of precipitation data using reliable model and IF-biLSTM model is significantly outperformed according to LSTM and GRU models. Especially even for far forward forecastings (four, five, six) those have high standard deviations this model gives more accurate results according to the other models. It is resulted that, proposed model can be used for different forecasting studies confidently.

Author contribution Levent Latifoğlu conceived and designed the analysis, performed the analysis, and wrote the paper.

Funding Financing will be borne by the authors.

Data availability All data used are original and the method used in the article was applied for the first time.

Declarations

Ethics approval Ethical approval is no need for approval.

Consent to participate This article accepted by all authors.

Consent to publish For this article publishing approval has been given.

Competing interests The author declares no competing interests.

References

Arsenault R, Bazile R, Dallaire-Ouellet C, Brissette F (2016) CANOPEX: A Canadian hydrometeorological watershed database. Hydrological Processes, 30:2734-2736. doi:10.1002/hyp.10880.

Bashar A (2019) Survey on evolving deep learning neural network architectures. Journal of Artificial Intelligence, 1:73-82.

Box GE, Jenkins GM, Reinsel GC, Ljung GM (2015) Time series analysis: forecasting and control. John Wiley & Sons.

Cho K, Van Merriënboer B, Bahdanau D, Bengio Y (2014) On the properties of neural machine translation: Encoder-decoder approaches. arXiv preprint arXiv:1409.1259.

Du J, Liu Y, Yu Y, Yan W (2017) A prediction of precipitation data based on support vector machine and particle swarm optimization (PSO-SVM) algorithms. Algorithms, 10: 57.

Freeman BS., Taylor G, Gharabaghi B, Thé J (2018) Forecasting air quality time series using deep learning. Journal of the Air & Waste Management Association, 68:866-886.

Hamill TM, Whitaker JS (2006) Probabilistic quantitative precipitation forecasts based on reforecast analogs: Theory and application. Monthly Weather Review, 134:3209-3229.

Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural computation, 9: 1735-1780.

http://canopex.etsmtl.net/

Johansson M. (1999) The hilbert transform. Mathematics Master's Thesis. Växjö University, Suecia. Disponible en internet: http://w3. msi. vxu. se/exarb/mj_ex. pdf, consultado el, 19.

Kim J, Moon N (2019) BiLSTM model based on multivariate time series data in multiple field for forecasting trading area. Journal of Ambient Intelligence and Humanized Computing, 1-10.

Kuan L, Yan Z, Xin W, Yan C, Xiangkun P, Wenxue S, ... & Xin Z (2017) Short-term electricity load forecasting method based on multilayered self-normalizing GRU network. In 2017 IEEE Conference on Energy Internet and Energy System Integration (EI2) (pp. 1-5). IEEE.

Le XH, Ho HV, Lee G, Jung S (2019) Application of long short-term memory (LSTM) neural network for flood forecasting. Water, 11: 1387.

Le TT, Pham BT, Ly HB, Shirzadi A, Le LM (2020) Development of 48-hour precipitation forecasting model using nonlinear autoregressive neural network. In CIGOS 2019, Innovation for Sustainable Infrastructure (pp. 1191-1196). Springer, Singapore.

Li W, Logenthiran T, Woo WL (2018) Multi-GRU prediction system for electricity generation's planning and operation. IET Generation, Transmission & Distribution, 13:1630-1637.

Lu K, Meng XR, Sun WX, Zhang RG, Han YK, Gao S, Su D (2018) GRU-based encoder-decoder for short-term CHP heat load forecast. In IOP Conference Series: Materials Science and Engineering, 392:062173

Miao KC; Han TT, Yao YQ, Lu H, Chen P, Wang B, Zhang J (2020) Application of LSTM for short term fog forecasting based on meteorological elements. Neurocomputing, 408:285-291.

Nourani V, Sattari MT, Molajou A (2017) Threshold-based hybrid data mining method for long-term maximum precipitation forecasting. Water Resources Management, 31:2645-2658.

Nourani, V., Razzaghzadeh, Z., Baghanam, A. H., & Molajou, A. (2019). ANN-based statistical downscaling of climatic parameters using decision tree predictor screening method. Theoretical and Applied Climatology, 137(3), 1729-1746.

Price M (2013) Introducing groundwater. Routledge.

Retalis A, Tymvios F, Katsanos D., Michaelides S. (2017) Downscaling CHIRPS precipitation data: an artificial neural network modelling approach. International Journal of Remote Sensing, 38:3943-3959.

Schuster M, Paliwal KK (1997) Bidirectional recurrent neural networks. IEEE transactions on Signal Processing, 45:2673-2681.

Shahid F, Zameer A, Muneeb M (2021) A novel genetic LSTM model for wind power forecast, Energy, 223, 120069.

Siami-Namini S, Tavakoli N, Namin AS (2019) The performance of LSTM and BiLSTM in forecasting time series. In 2019 IEEE International Conference on Big Data (Big Data) 3285-3292.

Siami-Namini S, Tavakol, ., & Nami, AS (2019) A comparative analysis of forecasting financial time series using arima, lstm, and bilstm. arXiv preprint arXiv:1911.09512.

Wei H, Li JL, & Liang TG (2005) Study on the estimation of precipitation resources for rainwater harvesting agriculture in semi-arid land of China. Agricultural Water Management, 71:33-45.

Wu H, Adler RF, Hong Y, Tian Y, & Policelli, F (2012) Evaluation of global flood detection using satellite-based rainfall and a hydrologic model. Journal of Hydrometeorology, 13:1268-1284.

Wu X, Li J, Jin Y, Zheng S (2020) Modeling and analysis of tool wear prediction based on SVD and BiLSTM. The International Journal of Advanced Manufacturing Technology, 106: 4391-4399.

Yu Y, Cao J, Zhu J (2019) An LSTM short-term solar irradiance forecasting under complicated weather conditions. IEEE Access, 7:145651-145666.

Figures

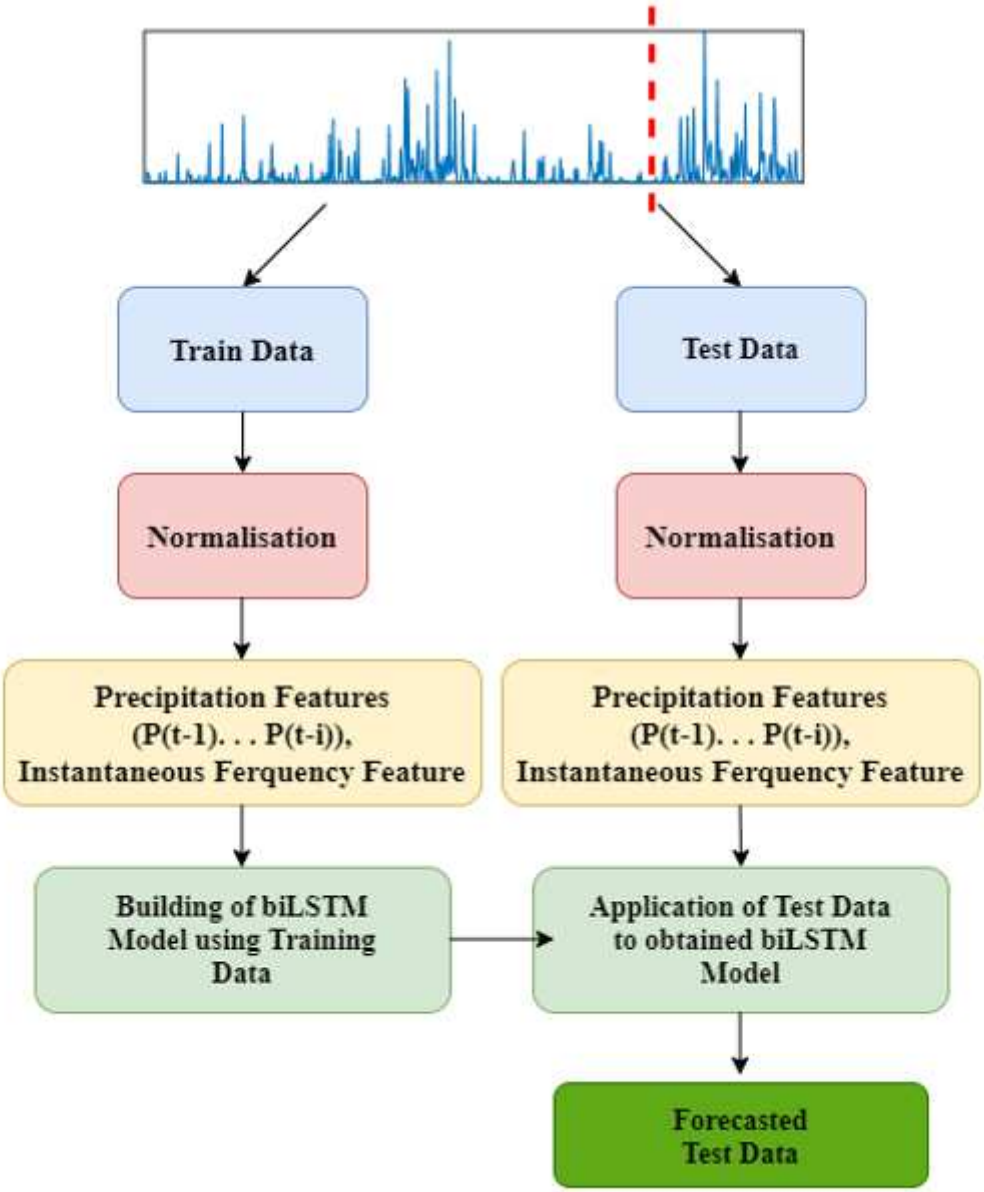


Figure 1

The proposed DNN forecasting model for daily precipitation data

Figure 2

Map of Churchill River Above Otter Rapids Station Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

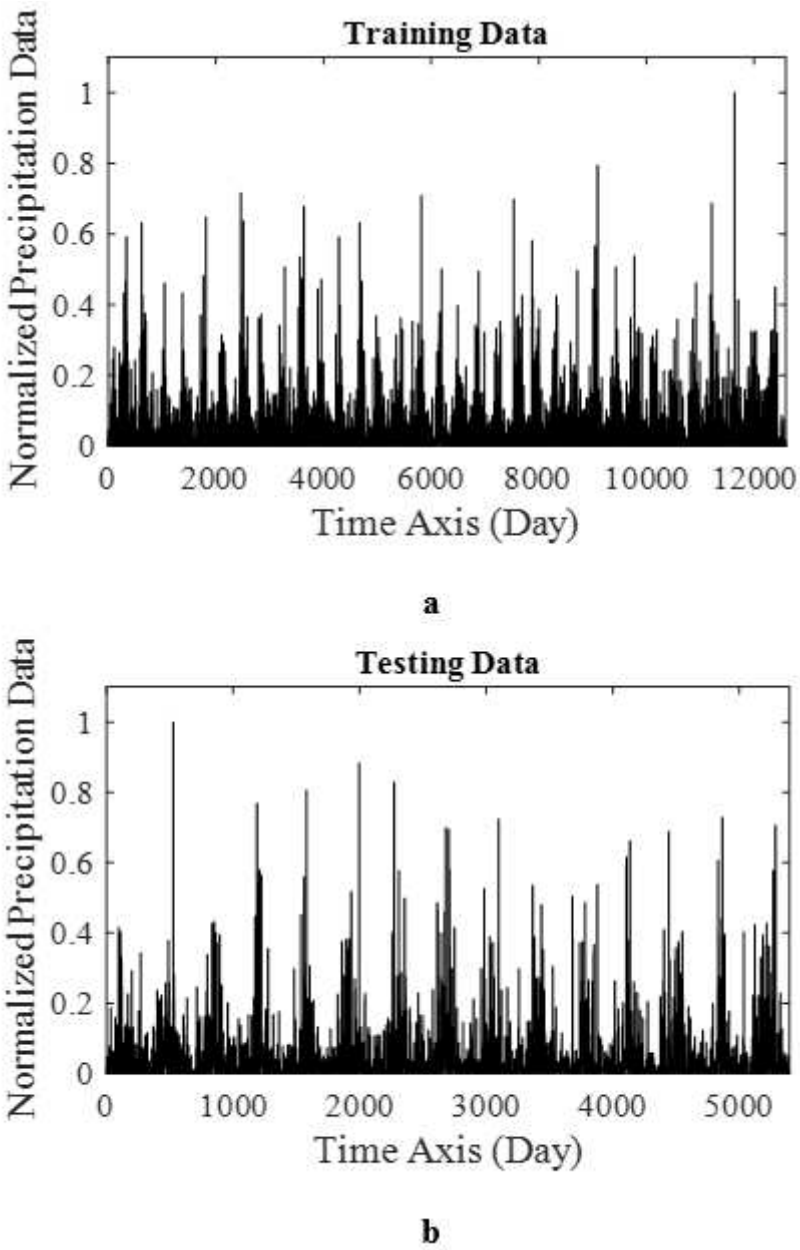


Figure 3

Daily Precipitation data measured on Churchill River above Otter Rapids basin a- Training Data, b- Testing Data

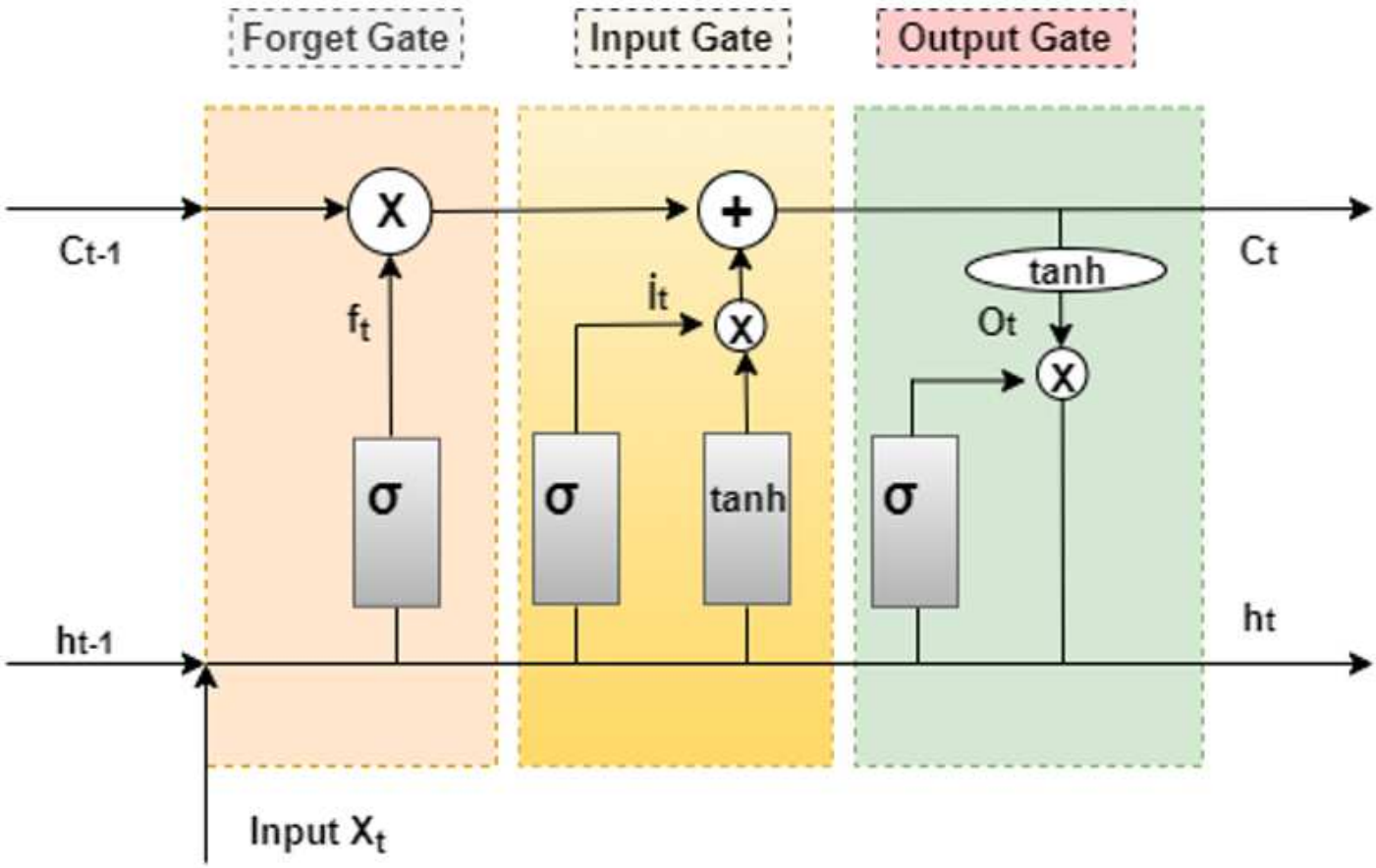


Figure 4

LSTM Unit

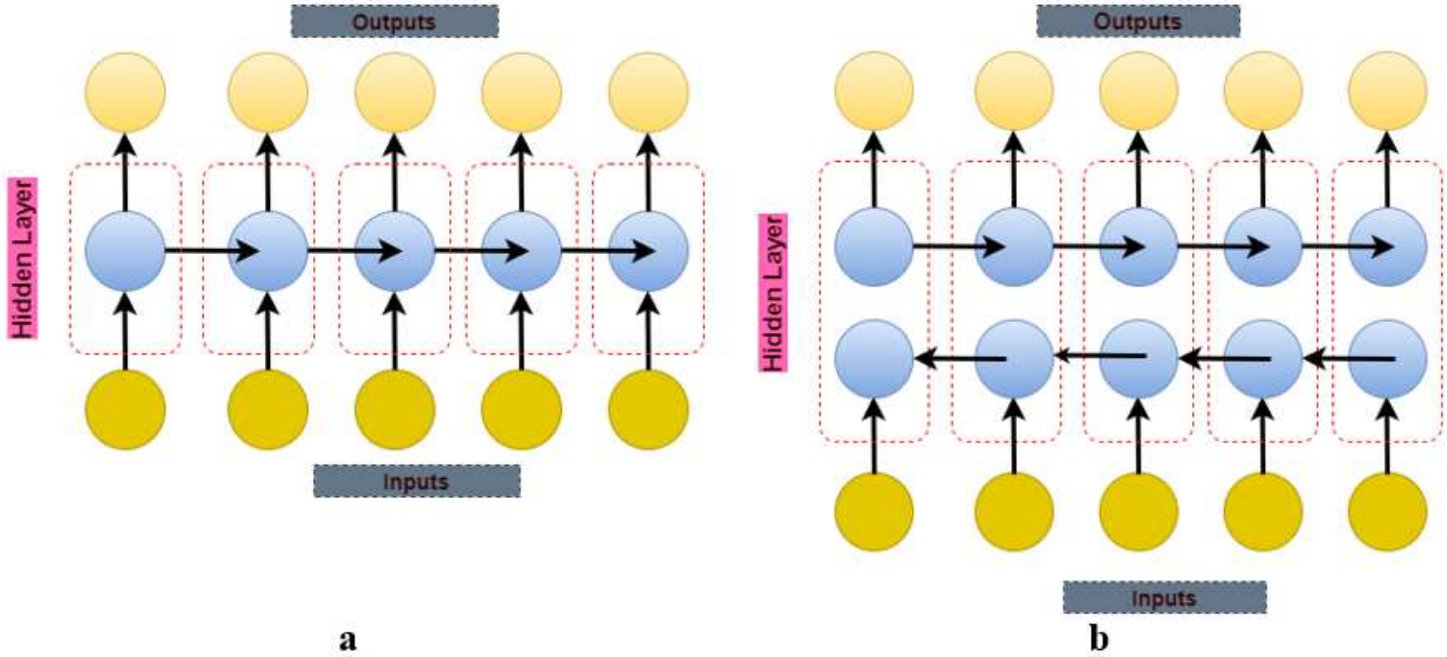


Figure 5

a. LSTM architecture b. biLSTM architecture

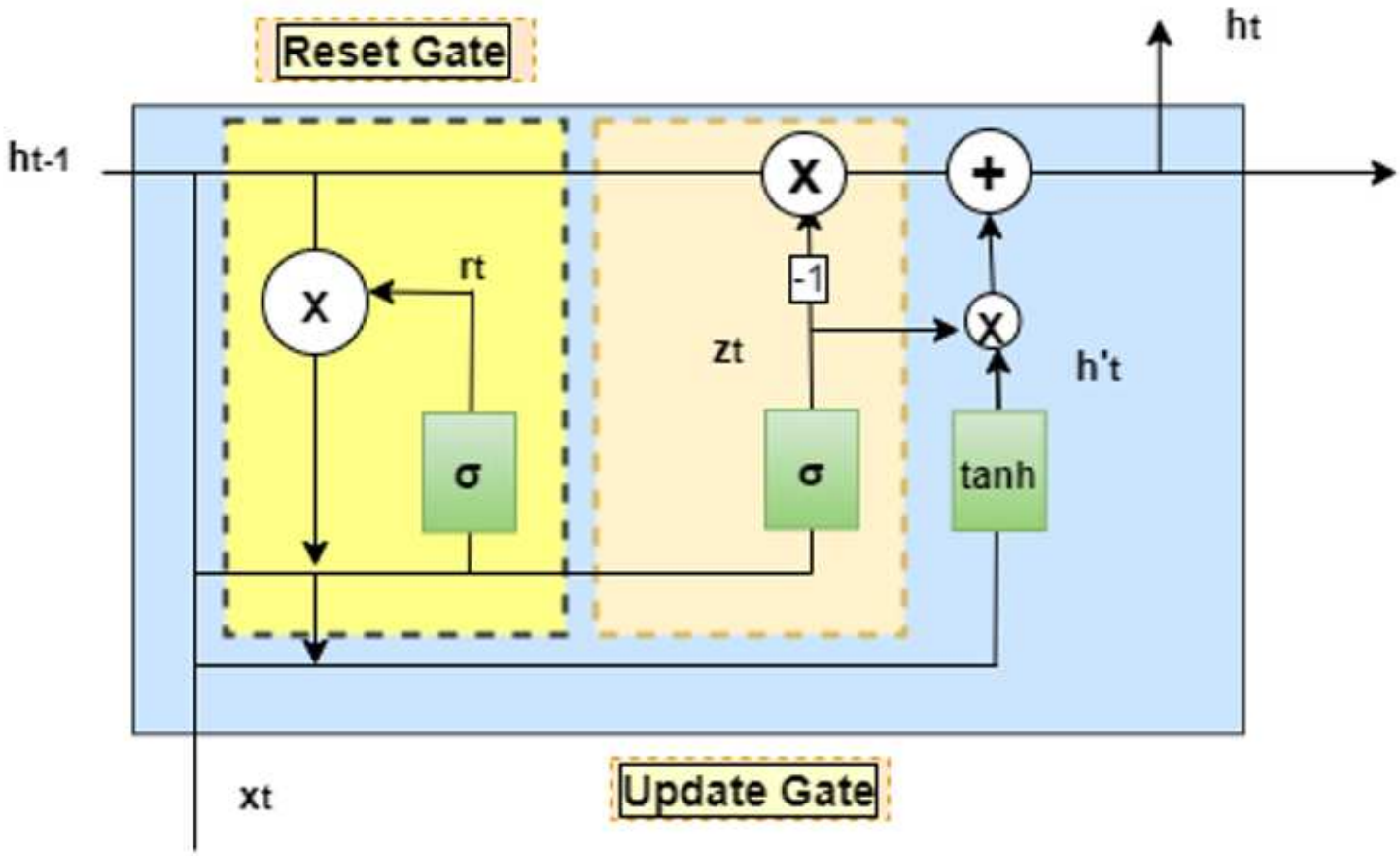


Figure 6

GRU Unit

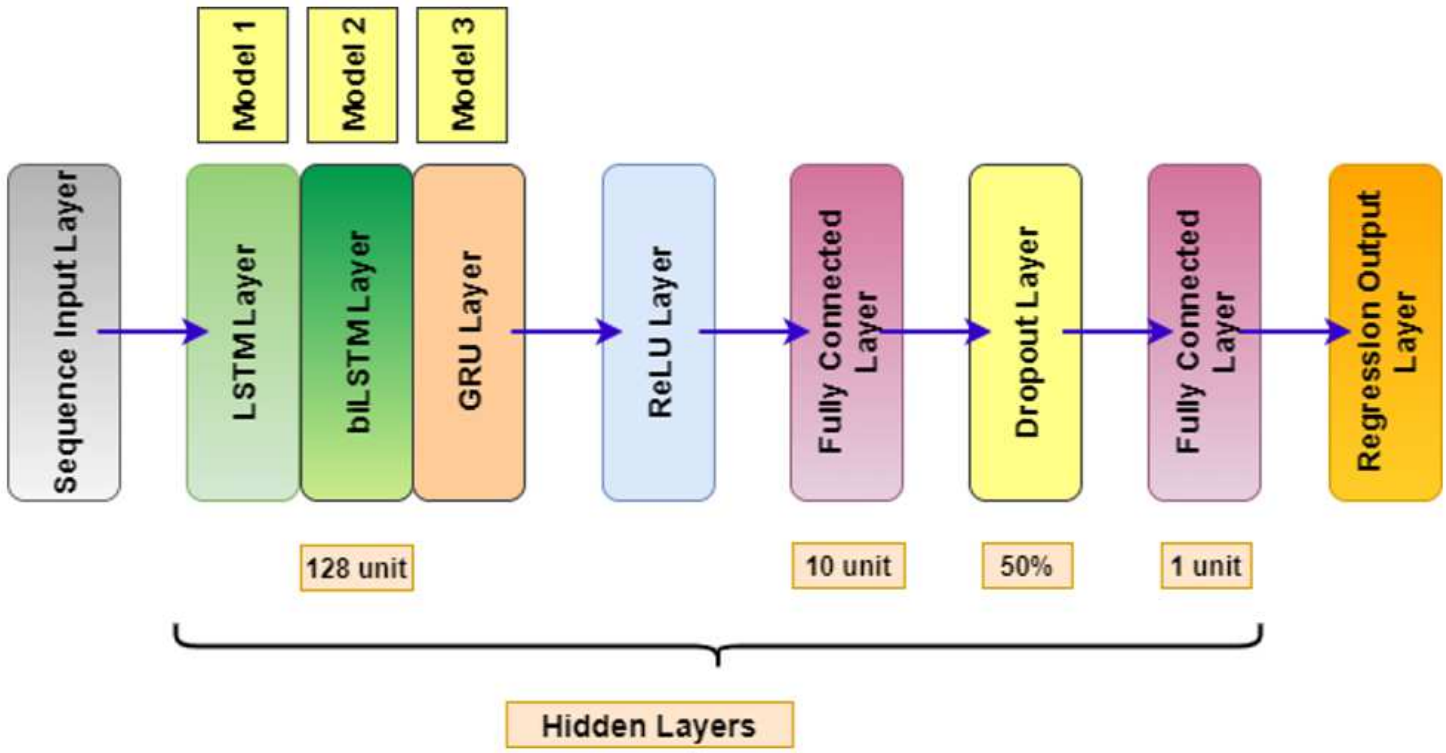


Figure 7

Forecasting models using LSTM Model, biLSTM and GRU Model

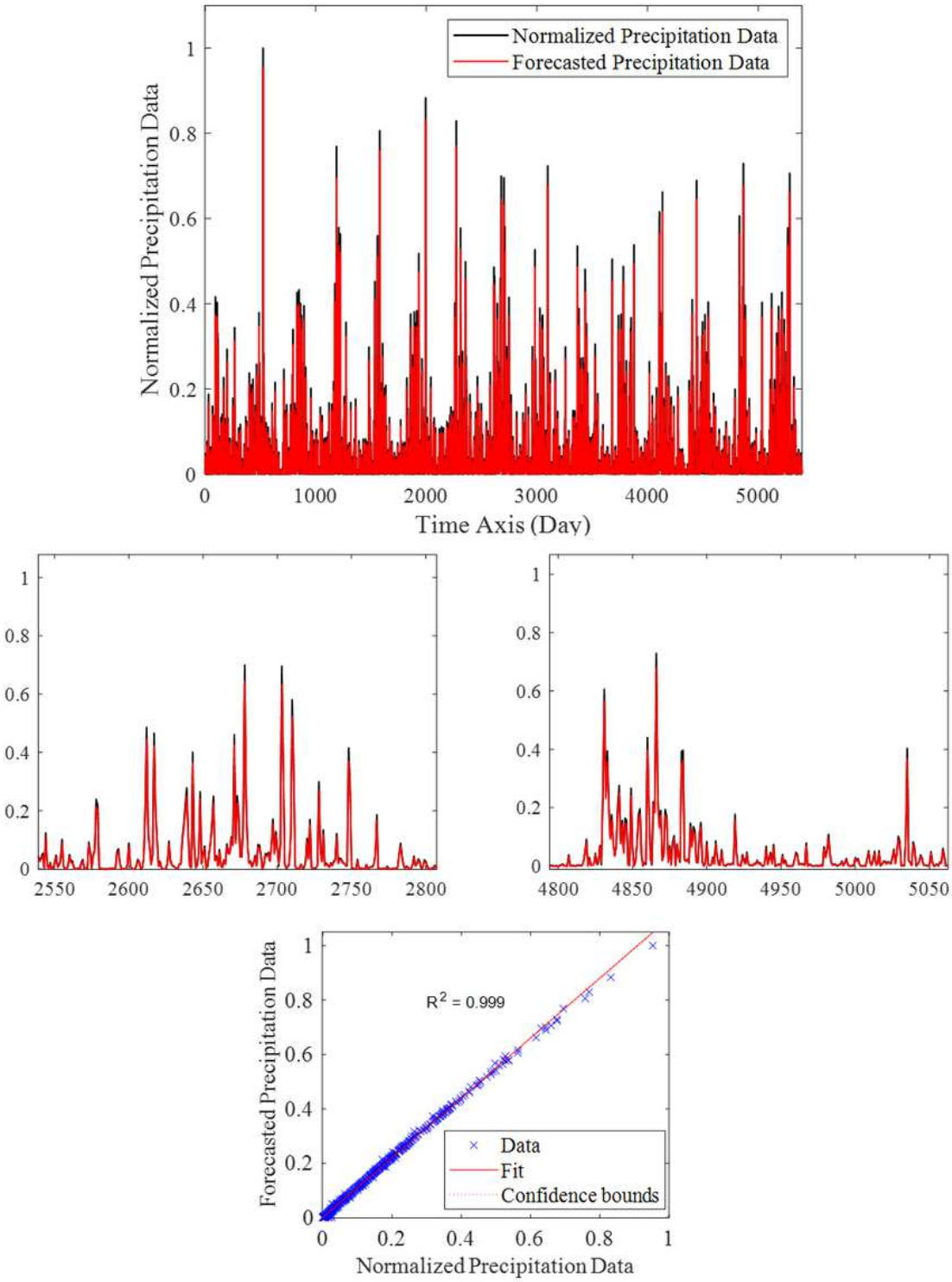


Figure 8

One-ahead forecasting of daily precipitation data using biLSTM Model

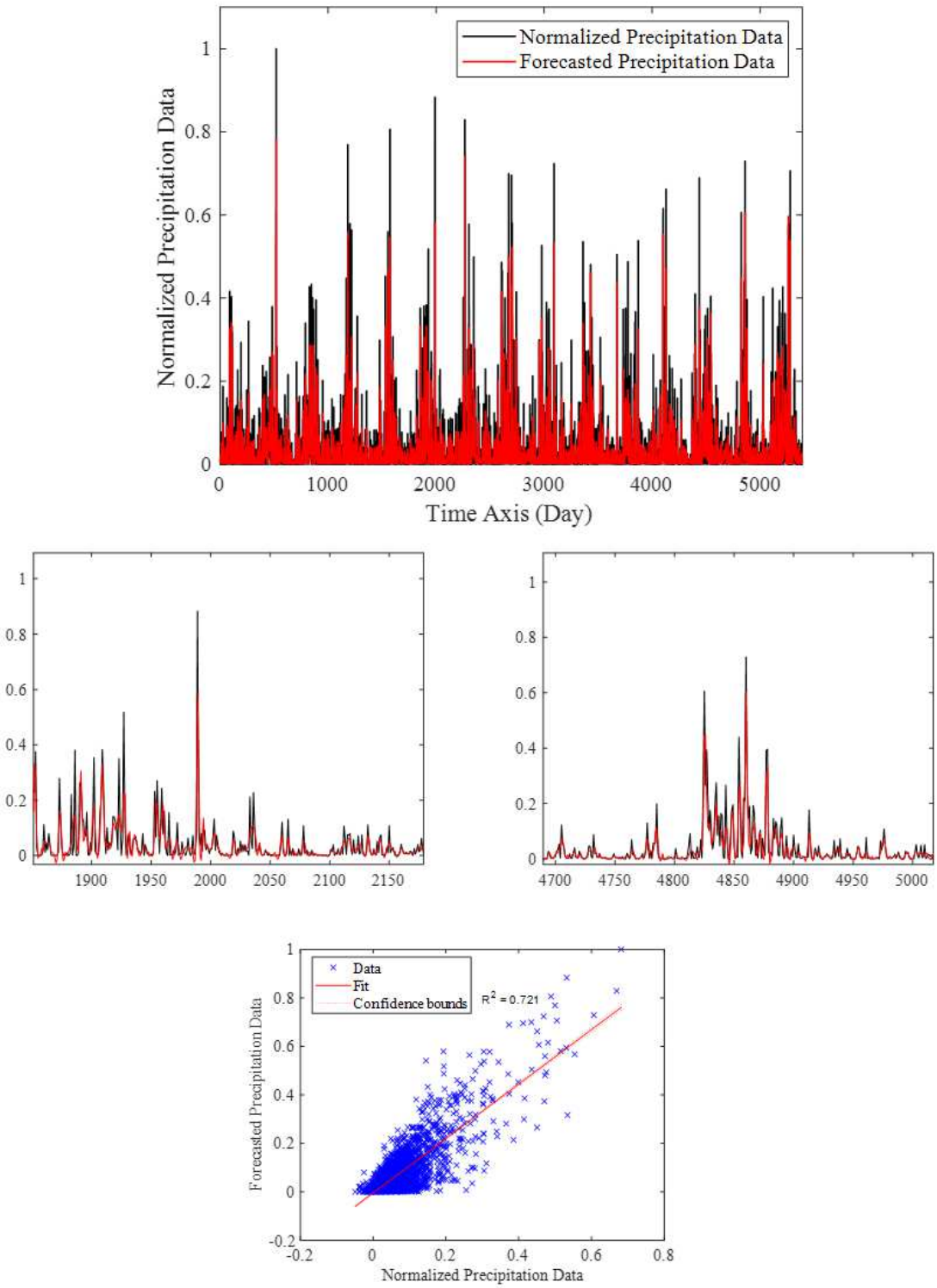


Figure 9

Six-ahead forecasting of daily precipitation data using three input biLSTM Model

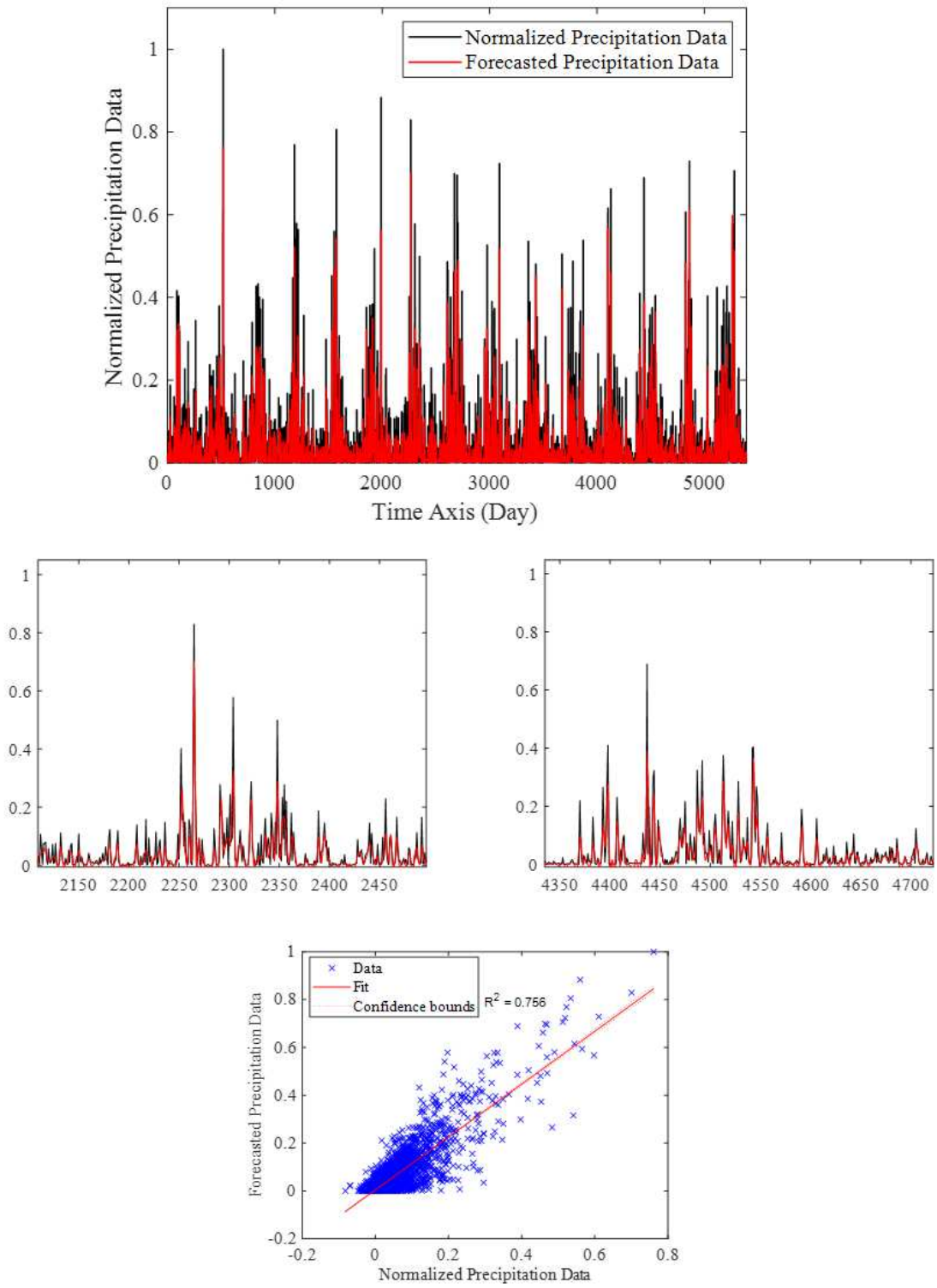


Figure 10

Six-ahead forecasting of daily precipitation data using IF- biLSTM Model