IJMS 2022 9(1): 65 - 87

International Journal of Multidisciplinary Studies (IJMS)
Volume 9, Issue I, 2022

Machine Learning Based Weather Prediction Model for Short Term Weather Prediction in Sri Lanka

Hennayake K.M.S.A.[1], Dinalankara R.[2] and Mudunkotuwa D.Y.[3]

[1]Faculty of Engineering, University of Sri Jayewardenepura
[2]Department of Computer Engineering, Faculty of Engineering, University of Sri Jayewardenepura
[3]Department of Mechanical Engineering, Faculty of Engineering, University of Sri Jayewardenepura

ABSTRACT

Weather forecasting is the field of making predictions of the future state of the atmosphere of a certain location by analyzing initial values of relevant atmospheric characteristics which are obtained by meteorological observations. Since weather prediction has substantial effect in economic sectors such as agriculture, health, aviation, hydro power generation and even in daily lives of people, issuing accurate weather forecasts is a major responsibility of meteorological authorities across the world. Even though forecasting weather in mid-latitudes is uncomplicated and reliable, weather prediction in a tropical country like Sri Lanka is notoriously difficult as sudden changes of convective tropical weather phenomena are quite difficult to be predicted by prevailing Numerical Weather Prediction (NWP) methods. Therefore, the current research aims to present machine learning based weather prediction models for Sri Lanka for making short term forecasts for the most significant weather attributes such as temperature and precipitation. This paper discusses on implementing two multivariate Long Short-Term Memory Network models (LSTM) to make predictions on temperature and precipitation separately for a selected weather station in Sri Lanka and review the applicability of machine learning to solve highly nonlinear and complex weather problems. The prediction performances of the implemented LSTM models are evaluated using standard evaluation techniques such as Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE). The results show that two LSTM models have made predictions with least RMSE and MAE values, evidencing the successful applicability of machine learning for solving complex and nonlinear patterns of past observational weather data and making accurate weather forecasts.

KEYWORDS: *Weather Prediction, Neumerical Weather Prediction, Artificial Neural Network, LSTM*

Corresponding author: Hennayake K.M.S.A., Email: anuradhahennayake@gmail.com

1 INTRODUCTION

Weather forecasting has a long history as humans have started forecasting weather by observing periodic atmospheric occurrences, signs of nature and behavioral patterns of the animals etc. Weather predictions have a major impact on the daily lives of people across the world and the economic sectors such as, agriculture, healthcare, aviation, hydropower generation etc. Presently, with the technological advances, the most used weather prediction method is numerical weather prediction. NWP uses mathematical models to make weather predictions by solving a set of complex mathematical equations that describe the behavior of the atmosphere. These mathematical equations are derived from basic principles of conservation of mass, Momentum, and energy. Additionally, water vapor equation and the ideal gas law are also applied in NWP (Holmstrom, Liu & Christopher 2016, Coiffier 2011, Larraondo, Inza & Lazano 2017, Warner 2011).

Even though NWP is widely used for weather forecasting and offers reliable outputs for the mid latitude region, the applicability of NWP for making reliable weather forecasts for tropical weather is debatable since tropical weather systems are spontaneously changing and tropical atmospheric dynamics are quite complicated for making predictions. Furthermore, there are some own limitations of NWP which lead to unrealistic weather forecasts. The Physical models used in NWP are normally suitable for predicting higher spatial resolutions which scale between 1km and 10km (Coiffier 2011). This can exclude more localized weather details and make forecasts less accurate. Capturing precise initial conditions and boundary conditions of the atmosphere is essential since these conditions have a significant effect on the predictions made by NWP (Krasnopolsky, Fox-Rabinovitz & Belochitski 2008). However, the initial conditions of the atmosphere can be uncertain and difficult to be captured. The approximations in the mathematical equations that describe the dynamics of the atmosphere may fail to interpret the unstable and chaotic nature of the atmosphere precisely and may lead to make erroneous forecasts. Moreover, Lack of human expertise (Elhoseiny, Huang, Elgammal 2015), limitations of the processing power and the usage of expensive instruments and sensors for collecting weather data, also discourage the use of NWP for predicting weather.

Even though traditional weather predicting methods possess above drawbacks, most of the countries including Sri Lanka, rely on NWP for weather predicting. The economy of Sri Lanka mainly depends on agriculture and volatile tropical weather has a major impact on agricultural productivity of the country. Therefore, having a reliable weather prediction model will be quite advantageous for Sri Lanka.

Considering above pitfalls of NWP, researchers' attention has been focused on finding alternative weather prediction methods to address these problems and give a reliable solution to tremendously difficult weather prediction problem. Machine Learning is such trendy technology that has given promising results by solving a wide range of outrageously complex and nonlinear problems involving huge datasets which are difficult to be solved by traditional programming or by humans. Machine Learning allows computer algorithms to learn themselves and improve their ability of performing a given task through experience without human intervention. By the current research, we aim to present a machine learning based weather prediction model for Sri Lanka for making short term weather forecasts for main weather parameters such as temperature and precipitation. This section includes literature review of multiple previous work that has been conducted on weather forecasting using machine learning techniques.

Mohammed et al (2020) have presented three machine learning models for predicting rainfall for several locations in eastern India. The machine learning models implemented are multiple linear regression, support vector regression and lasso regression. The authors have used a rainfall dataset from 1901 to 2015 for their research. Mean absolute error is used to measure the performance and support vector regression has shown the minimum error of 4.35 of MAE and linear regression model has shown 11.71 of MAE, lasso regression model has shown 0.95 of MAE for the rainfall predictions.

Abdul Keder et al (2020) have proposed a machine learning related hybrid technique for forecasting rainfall by combining Particle Swarm Optimization (PSO) and Multi-Layer Perceptron (MLP) neural Network and compared with a back propagation algorithm namely Levenberg-Marquardt (LM). The RMSE for MLP based PSO is 0.14 while RMSE for MLP based LM is 0.18.

Zaytar & Amrani (2016) have presented a deep neural network architecture for time series prediction. The authors have proposed two types of multi stacked long short-term memory network models per city for nine cities in Morocco to make 24 and 72 forecasts for three weather parameters. The selected dataset contains 15 years of hourly meteorological data. The selected weather attributes are temperature, humidity, and wind speed. The Mean Squared Error (MSE) is used to evaluate the predictions and the results have shown that LSTM network can forecast weather variables with a good accuracy.

(Aswin, Geetha & Vinayakumar 2018) have proposed deep learning models for rainfall prediction. The implemented deep learning models include a LSTM and a ConvNet (Convolution Neural Network). These models have been used for making predictions for the global monthly average rainfall data for 10368 geographic locations around the world. The time series dataset has been collected from NCEP center. The results of the study have shown that they have achieved 2.55 of RMSE and 1.6897 of MAPE value for the precipitation predictions made by LSTM and 2.44 of RMSE and 1.7281 of MAPE value for the precipitation predictions made by ConvNet. The authors have stated that the accuracy can be further increased by increasing the number of hidden layers of deep learning models.

As in reference (Salman et al, 2018), the authors have proposed single layer and multi-layer LSTM models with an intermediate variable for weather forecasting. In this study the models are utilized to investigate the effect of an intermediate weather variable on the prediction accuracy. The weather dataset is collected from Weather Underground at Hang Nadim Indonesia airport domain. The LSTM model is implemented by adding intermediate variable signal to memory block of the LSTM. The output variable is visibility, and temperature, pressure, humidity and dew point are used as intermediate variables. The best result is given by multi-layer LSTM model with the intermediate pressure variable with high validation accuracy and minimum RMSE error.

Singh et al, (2019) have presented three machine learning models for weather forecasting. They have analyzed and compared the performances of the models using the RMSE error on test dataset. The machine learning models implemented were Support Vector Machine (SVM), Artificial Neural Network (ANN) and a Time Series based Recurrent Neural Network (RNN). The forecasts have been made for a period of eight weeks. Datasets including Previous 12 years of weather data have been collected from different airport weather stations in India and these datasets include several weather attributes. For support vector machine, selected input features were time, rain, snow and humidity. The output feature was temperature. For artificial neural network and for recurrent neural network, the input attributes were temperature, pressure and humidity. The output attribute was temperature. According to the comparison of above three models it can be concluded that time series RNN has shown minimum RMSE error compared to SVM and ANN models when forecasting temperature.

Thilakarathne & Premachandra (2017) have proposed a flood prediction hybrid model using machine learning and data mining methods for predicting floods in north central province of Sri Lanka. The hybrid model is a combination of time series based ARIMA model for predicting future weather attributes and binary classifier ANN for predicting flood occurrence probability. A dataset including historical weather data from 1976 to 2015 is collected from the department of meteorology Sri Lanka for Anuradhapura district in north central province and different datasets including 51 flood type disaster records from 1976 to 2015 are collected from disaster information management system of Sri Lanka for same domain. The predictions made by ARIMA model and flood type disaster dataset were inputs to the binary classifier. The model has shown 91.7% of accuracy of predicting flood probability and this research clearly shows the capability of machine learning in making weather related forecasts.

As in (Narvekar, & Fargose 2015), the authors have implemented a neural network with back propagation technique for daily weather forecasts. In this study weather dataset is collected for one of the meteorological stations in India. The ANN model predicts minimum and maximum temperature, relative humidity and rain fall of the day and the accuracy of the model has been measured using mean squared error (MSE) function and this research has recommended ANN with back propagation for weather forecasting.

Reference (Subashini, Thamarai & Meyyappan 2019) has suggested a Deep Learning Model for weather forecasting. The authors have implemented a LSTM network and have made predictions for different weather parameters such as temperature, cloud cover and wind speed with different combinations of weather attributes. The dataset used in the study was collected from national climate data center for about 12 years. The results of the study have shown that they have achieved a fine accuracy using the proposed LSTM model.

Above studies have demonstrated promising results applying machine learning for challenging weather prediction problems. Even though these modern technologies dominate the old weather prediction methods, only handful of research are conducted regarding using machine learning for predicting weather in tropical countries like Sri Lanka. Time series prediction of temperature and precipitation based on past weather observational data using a multivariate LSTM model is mostly undetermined. The main authority of providing meteorological services in Sri Lanka is Department of Meteorology. One of the main methods they use for predicting weather is creating three-dimensional picture of the atmosphere by analyzing surface and upper-air information, and then, an experienced meteorologist predicts expected weather using climatology and persistence of weather systems. This method totally relies on the expertise and the experiences of the meteorologist. Other method is the use of numerical weather products output by global numerical weather prediction models.

But these NWP methods are less accurate when giving predictions in tropical countries like Sri Lanka. High turbulence in the atmosphere in tropical countries is difficult to be solved by equations of NWP and huge computational power is needed to solve all the equations that describe chaotic nature of the tropical atmosphere. Hence, the forecasts become less accurate and more expensive. Therefore, it is evident that tropical countries like Sri Lanka require a state of art, reliable solution for weather prediction problem. The current study aims to present a machine learning based weather prediction model for Sri Lanka to address existing gaps of traditional weather predicting methods

The paper is structured as follows: Section 2 describes the design methodology and Section 3 presents the evaluation details. Section 4 concludes the paper.

Table 01: Summary of the literature review

Research title	Applied Method	Accuracy/Error	Limitations
Mohammed et al (2020)	Forecast rainfall using three regression models namely multiple linear regression, support vector regression and lasso regression	Mean absolute error was used to measure error and support vector regression showed the minimum error of 4.35 of MAE and linear, lasso regression models have shown 11.71 and 10.95 of error	The study has used only rainfall data as input and has used linear models to predict nonlinear weather.
Abdul Keder et al (2020)	Forecast rainfall by combining Particle Swarm Optimization (PSO) and Multi-Layer Perceptron (MLP) and compare RMSE with various Back Propagation (BP) algorithms such as Levenberg-Marquardt (LM)	RMSE for MLP based PSO is 0.14 and RMSE for MLP based LM is 0.18	Used only RMSE error as performance metric. There is no in-depth evaluation of error between actual and predicted rainfall.
Zaytar & Amrani (2016)	Implemented two LSTM models per city for nine cities in Morocco to make 24 and 72 forecasts for temperature, humidity, and wind speed variables.	MSE error was calculated separately for each model The lowest MSE for 24-hour prediction is 0.00516 and the lowest MSE for 72-hour prediction is 0.00675	There is no indication of a method to determine the error individually for different parameters predicted.
Aswin, Geetha & Vinayakumar (2018)	LSTM and a ConvNet for making predictions for the global monthly average rainfall data	achieved 2.55 of RMSE and 1.6897 of MAPE value for predictions made by LSTM and 2.44 of	Only compared RMSE and MAPE errors. Cannot identify the actual error

		RMSE and 1.7281 of MAPE value for the predictions made by ConvNet	between actual and predicted values.
Salman et al, (2018)	Single layer and multi-layer LSTM models with an intermediate variable for weather forecasting. The output variable is visibility. temperature, pressure, humidity and dew point are used as intermediate variables	The model has gained the validation accuracy 0.8060 and RMSE 0.0775 using the pressure variable	Few input parameters are used and cannot identify the actual error between actual and predicted values.
Singh et al, (2019)	Support Vector Machine (SVM), Artificial Neural Network (ANN) and a Time Series based Recurrent Neural Network (RNN) for predicting temperature	The RMSE is 6.67 for SVM, 3.1 for ANN and 1.41 for RNN	There are considerable deviations between the actual and predictions but the authors have not discussed this aspect.
Thilakarathne & Premachandra (2017)	Time series based ARIMA model for predicting future weather attributes and binary classifier ANN for predicting flood occurrence probability	The RMSE values for forecasted rainfall, minimum and maximum temperatures by ARIMA model are 115.58, 0.42 and 0.56 The ANN model has shown 91.7% of accuracy of predicting flood probability	There are considerable deviations between the actual and predictions of rainfall data. But the authors have not discussed about the actual error.

Narvekar, & Fargose (2015)	Neural network with back propagation technique for daily minimum and maximum temperature, relative humidity and rain fall prediction	MSE error is used to measure performance	Authors have not discussed the value of the MSE of the model.

2 RESEARCH METHODOLOGY

The proposed weather forecasting model is based on implementing two multivariate LSTM models for predicting two main weather attributes namely temperature and precipitation. The methodology consists of several sub tasks such as collecting past weather observation data, data pre-processing, designing, and implementing two separate LSTM models for predicting temperature and precipitation, train LSTM models on training set, test trained LSTM models by making predictions on keep aside test set and finally evaluate prediction performances of two models' using standard evaluation metrics. Figure 01, shows the flowchart diagram of the proposed methodology.

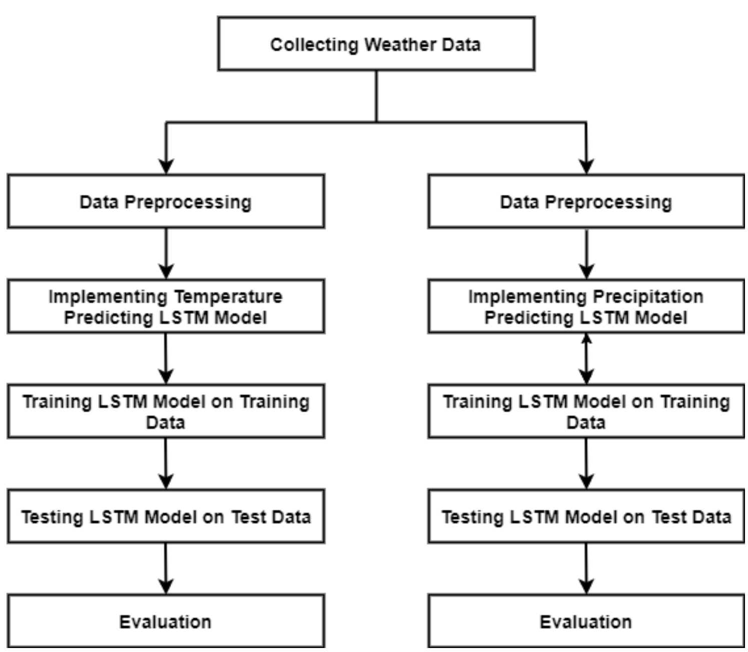

Figure 01: Methodology Workflow Diagram

2.1 Collecting Weather Observation Data

Machine Learning algorithms require input data to learn insights and build an appropriate model which can then be used for making predictions. Input data has a major influence on the prediction accuracy of a model. Historical weather observation dataset is collected from NOAA (National Oceanic and Atmospheric Administration), NCDC (National Climate Data Centre) for Colombo weather station in Sri Lanka (Latitude 6.9, Longitude 79.867). The collected dataset is a time series, and it contains 10 years of periodically measured data starting from 2010/01/01 to 2019/11/01. Moreover, the dataset consists of daily ground observational weather parameters including mean temperature for day, mean dew point for the day, mean sea level pressure for day, mean station level pressure for day, mean visibility for day, mean wind speed for day, maximum wind speed for day, maximum temperature for day, minimum temperature for day, total precipitation for day etc.

2.2 Data Pre-processing

The collected dataset is a raw dataset which needs pre-processing post applying on machine learning algorithm. Data pre-processing needs to be applied to the raw real-world dataset to transform it into the appropriate state that can be easily interpreted by machine learning algorithm. The main steps carried out in data pre-processing are data cleaning, feature selection, inspecting the dataset and visualization, data normalization, dividing the data set into training and test sets. Even though the implementation of the data pre-processing step is the same for both LSTM models, subtasks such as inspecting the dataset and visualization, data normalization, dividing the data into train and test sets are carried out separately for each model. In the data cleaning step, missing data and null values are imputed using next observation carried backward technique. This method is used because of the persistence nature of the weather variables. Input features are selected manually since there were some weather parameters such as “snow depth” which are irrelevant for predicting weather in selected station. For machine learning models the correlations between the variables in the dataset are quite important. Therefore, in feature selection step, the correlation between complex and nonlinear weather parameters is analyzed by preparing a correlation matrix. Then the dataset is visualized using visualization methods such as histograms, and scatterplots in order to observe hidden patterns of the data.

Figure 02, indicates the fluctuations of daily temperature throughout the ten years of time starting from 2010 to 2019 for Colombo station. Figure 03 is the year wise plot of temperature distribution for 4 years starting from 2010 to 2013. Figure 04, indicates the distribution of daily measured precipitation starting from 2010 to 2019 for Colombo station while figure 05, presents yearly distribution of daily precipitation for 4 years of time starting from 2010 until 2013

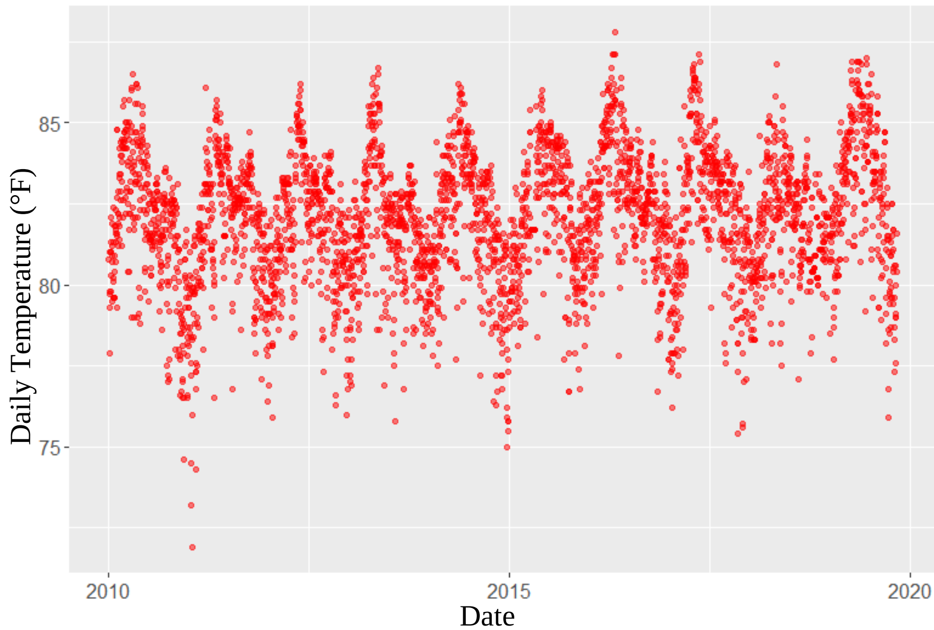

Figure 02: Daily temperature from 2010 to 2019 for Colombo station

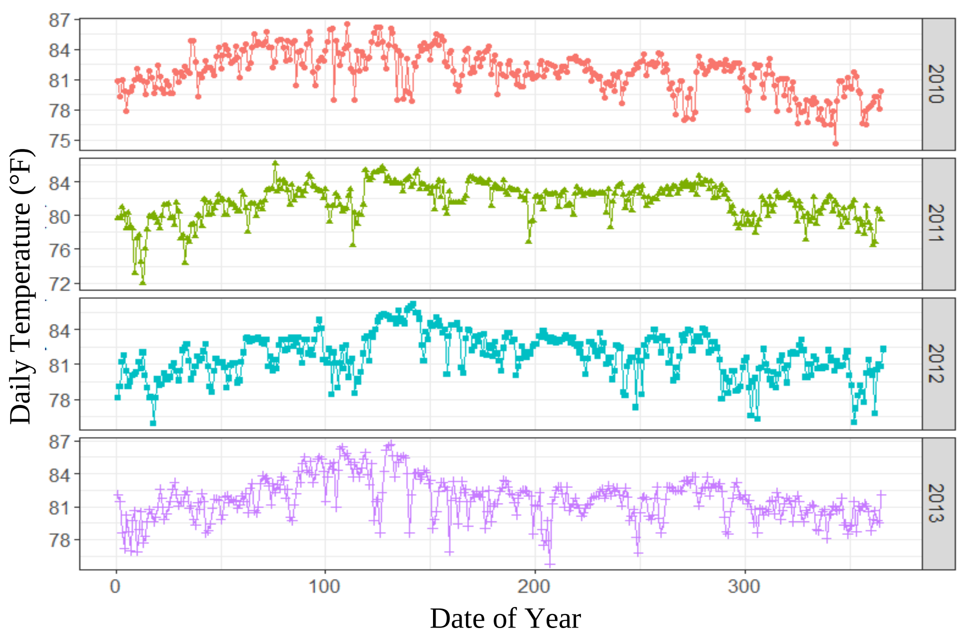

Figure 03: Comparison of daily temperature of four years starting from 2010 to 2013 for Colombo station

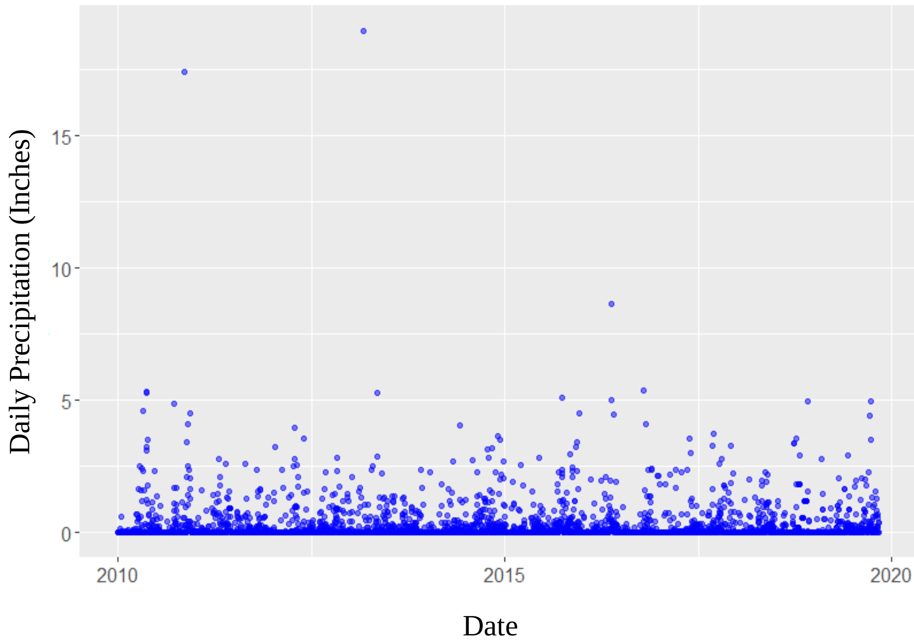

Figure 04: Daily precipitation from 2010 to 2019 for Colombo station

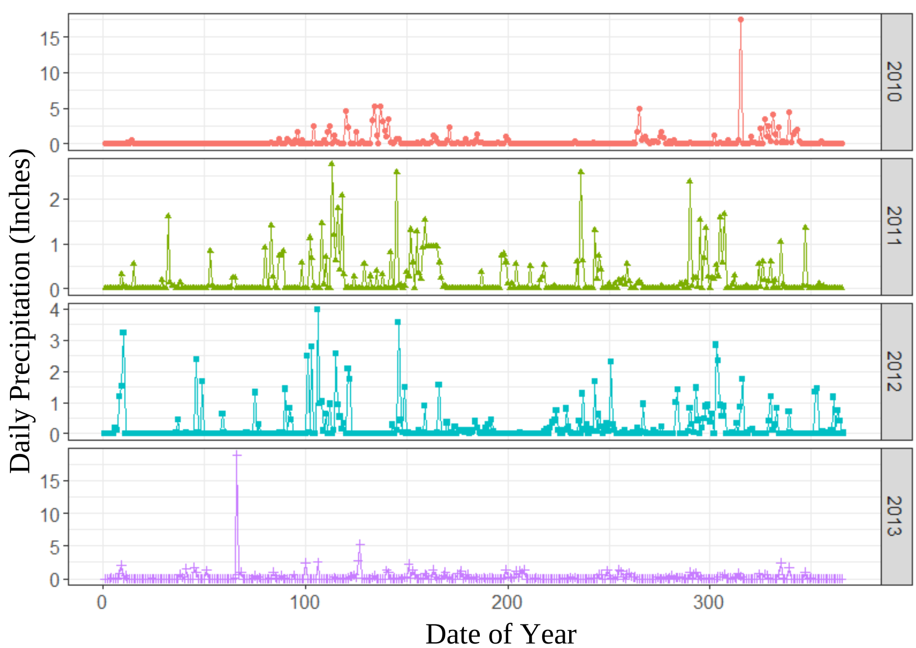

Figure 05: Comparison of daily precipitation of four years starting from 2010 to 2013 for Colombo station

The dataset used in current research is a time series. Therefore, it can be further decomposed and plotted to inspect level, trend, seasonal and random noise components individually in order to get a better understanding of the dataset. Figure 06, shows the decomposition plot of the temperature time series for ten years of time starting from 2010 to 2019. The decomposition plot clearly shows a nonlinear curved trend of temperature data which tends to increase and decrease over the time and moreover a definitive seasonality of the temperature time series can be observed.

Figure 07, indicates the decomposition plot of the precipitation time series for ten years of time from 2010 to 2019. A fluctuating trend component, and clear seasonal behavior of precipitation time series can be observed. The trend component is then removed using differencing technique. Differencing of the time series is done by getting the difference between two consecutive values in the series.

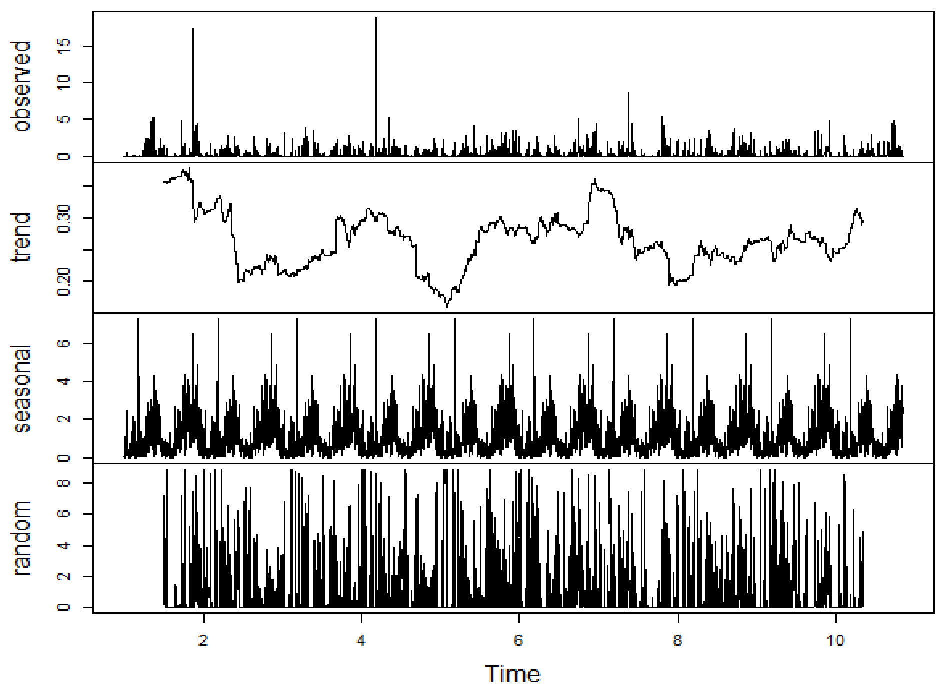

Figure 06: Decomposition plot of temperature time series

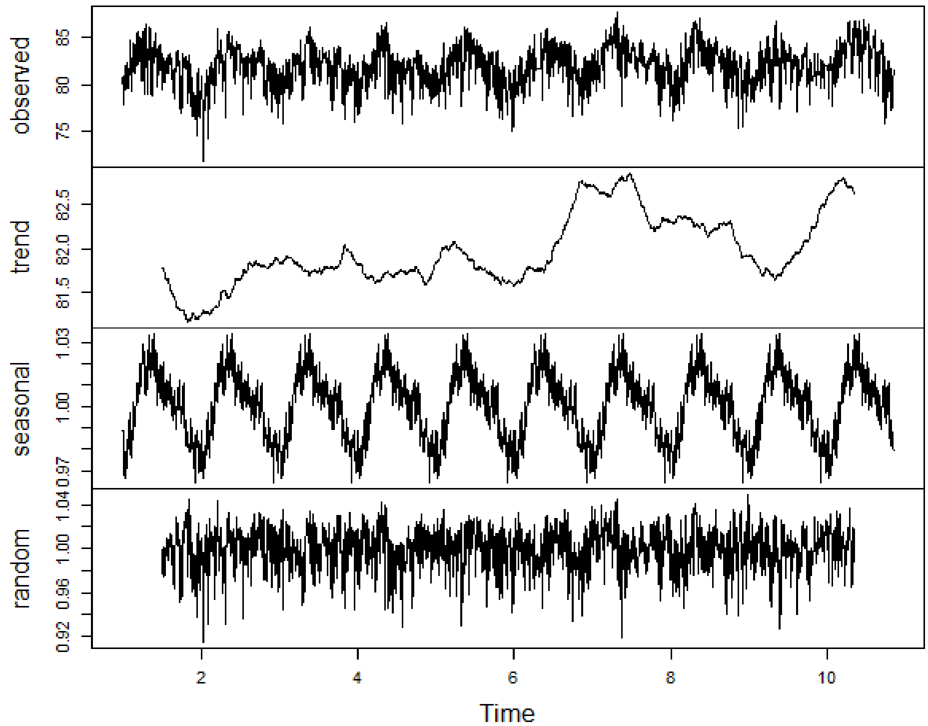

Figure 07: Decomposition plot of precipitation time series

The two LSTM models implemented in current research are based on supervised learning. Supervised learning algorithms expects the training data as a combination of input variables X and corresponding target variable Y. Therefore, the time series dataset should be reframed by lagging the dataset by 1-time step. Therefore, the value of the input variable X is equal to the value at (t-1) time step and the value of the output variable Y is the value at the time t.

The data should be normalized by scaling the dataset into a common scale between 0 and 1. Then the dataset is split in to training set and testing set. 80% of the pre-processed data is divided into the training set and 20% of dataset is reserved for testing set. Each Training set is made of 2 sets namely, x_train, y_train and each test set are composed of 2 sets namely, x_test and y_test. x_train includes predictor parameters in the training set and y_train includes the target variable for the training set. x_test consists of predictor variable for the test set and y_test includes the target variable for the test set. y_train and y_test sets are different for two LSTM models since temperature prediction LSTM, the target variable is temperature and for precipitation prediction LSTM, the target variable is precipitation. The dataset is divided in to training and testing sets by preserving the sequence of the time series since the order is important when making predictions for a time series.

Proposed LSTM networks are implemented using Keras package in R programming language with the help of

TensorFlow backend. Keras LSTM expects the input tensors in the specific format of a three-dimensional array. The three dimensions are namely batch size, timesteps and features. Therefore, the training data and testing data are reshaped into the above stated format before feeding them as inputs to both models

2.3 Design LSTM Algorithm

Long Short-Term Memory Network (LSTM) are a variation of Recurrent Neural Networks (RNN). They were introduced in 1997 by Hochreiter and Schmidhuber. LSTMs have the capability of learning long-term dependencies of sequential data for a long period of time and use that information for current processing without forgetting important information from earlier time steps. This ability makes LSTMs different from RNNs and makes powerful in recognizing patterns and make predictions in complex time series problems. LSTMs consists of memory blocks which enable the long-term memory true. The architecture of a LSTM memory block is shown in figure 08,

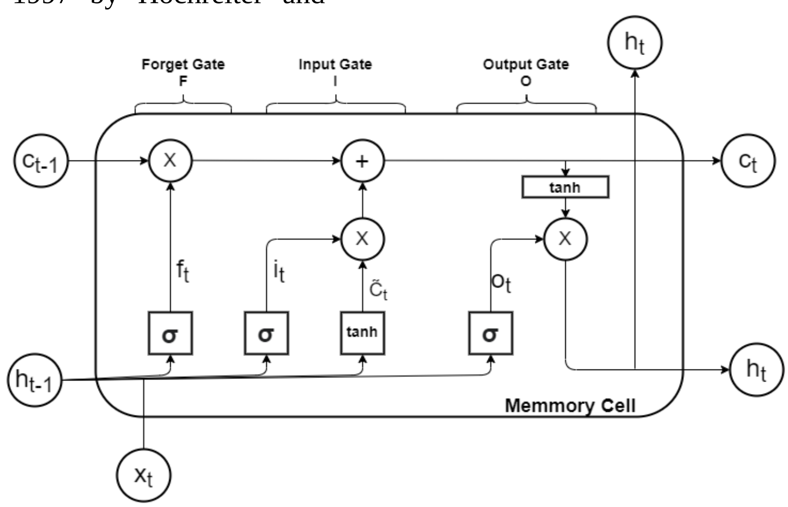

Figure 08: LSTM memory block

LSTM memory block consists of a memory cell and three gated units called forget gate, input gate, output gate. These gates control the data flow into the memory cell and out of it.

LSTM memory block receives three input signals. The input signal at the current time step (X_t), hidden state of previous unit (h_{t-1}) and memory of the previous unit (C_{t-1}) and LSTM memory block outputs two output signals namely, hidden state of the current unit (h_t) and memory of the current unit (C_t).

At the forget gate, $h_{t-1}$and X_t are taken in to account and outputs a value between 0 and 1 by the sigmoid activation function. 0 represents no information should pass and 1 represents all information should be passed. Therefore, forget gate determines what information should be passed through the gate.

$$f_t = \sigma\left(W_f[h_{t-1}, X_t] + b_f\right). \qquad (1)$$

At the input gate, takes the hidden state of previous unit (h_{t-1}) and the input signal at the current time step (X_t) together and update the cell state using a sigmoid layer. This gate returns a value between 0 and 1. The output of the input gate is then multiplied with the output of the candidate layer in order to update the cell state.

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

Then at the candidate layer a hyperbolic tangent layer is applied to the mix of the input signal at the current time step and hidden state of previous unit and outputs a vector of all possible candidate values that can be added to the cell state. The hyperbolic tangent function outputs values between -1 to +1. Then this candidate vector is added to the internal state.

$$\tilde{C}_t = tanh(W_i[h_{t-1}, X_t] + b_i). \qquad (3)$$

The previous memory state is multiplied by the output of the forget gate, and then added to the fraction of the new candidate value generated by the output gate to update the memory state.

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t. \qquad (4)$$

The output gate decides the fraction of the memory state which can be passed through to the output. First sigmoid layer is applied to the memory state and filters the memory state.

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

Then the memory state is put through a hyperbolic tangent function and pushed to scale the values between -1 to +1. And finally, the scaled memory state is multiplied by the output of the sigmoid layer in order to create the hidden state of the current unit.

$$h_t = O_t * tanh\, C_t. \qquad (6)$$

Implementation of two LSTM models for making temperature forecasts and precipitation forecasts is done using R programming language with the help of Keras and TensorFlow libraries. LSTM models developed in the current study are multivariate models by reason of multiple time dependent variables are used as inputs to the models in order to get optimal predictions for temperature and precipitation. There is no rule of thumb for selecting the number of layers or number of nodes for LSTMs. Therefore, starting with very simple LSTM networks, various experiments are carried out and compared RMSE values for finding minimal error LSTM models for the temperature and precipitation prediction. Since current research addresses a regression problem, one output node is used in output layers. Different architectures of LSTMs are tested by adjusting hyperparameters such as number of units in LSTM layer, number of hidden layers, and number of neurons in hidden layers. Learning rate was specified as 0.0001 and was not changed during tests. Moreover, the number of epochs also changed and observed the results.

Keras LSTMs expects its inputs to be in three-dimensional array namely, batch size, time steps and features. Batch size is set to 1 and look back is set to 14. For the temperature forecasting, a sequential model with multiple layers is finalized. The architecture of the temperature prediction LSTM model includes 1 LSTM layer consisting of 128 neurons which is followed by a dropout layer which is important for preventing over fitting the model, three fully connected dense layers and one output dense layer with one neuron.

The architecture of the precipitation prediction LSTM is also a sequential model with 1 LSTM layer consisting of 64 neurons which is followed by a dropout layer. Moreover, the model consists of two fully connected dense layers. The first dense layer includes 32 neurons, and the second dense layer includes 16 neurons. The precipitation predicting LSTM model also has one output dense layer with one neuron since it makes predictions for one parameter.

For both LSTM models, Adaptive Moment Estimation (ADAM) was used as the optimization algorithm and the learning rate was specified as 0.0001 and the learning rate decay over each update by 1×10^{-6}. Mean absolute error was used as the loss function of both LSTM networks.

2.4 Training LSTM Models

Training of two LSTM models is done using the training sets which consist of 80% of previously preprocessed data. Out of 3592 observations of the whole dataset 2872 observations are divided for the training set in order to train one LSTM model. In the training process 0.2 of observations from training data are divided for validation purpose. The weather attributes used as input features are mean temperature for the day, mean dew point for the day, mean sea level pressure for the day, mean station level for the day, mean visibility for the day, mean wind speed for the day, maximum sustained wind speed reported for the day, maximum temperature for the day, minimum temperature for the day, total precipitation reported during the day.

Both LSTM models are trained on training data in 100 epochs with early stopping function. The training of 2 models was observed using training history plots. Training history plot of a model indicates how MAE and Loss values change with the number of epochs. The training of the temperature prediction model was stopped after 88 epochs. Figure 09, shows the training history plot of the temperature prediction LSTM model. Training of the precipitation prediction LSTM model was stopped after 100 epochs. Figure 10, shows the training history plot of the precipitation prediction LSTM model.

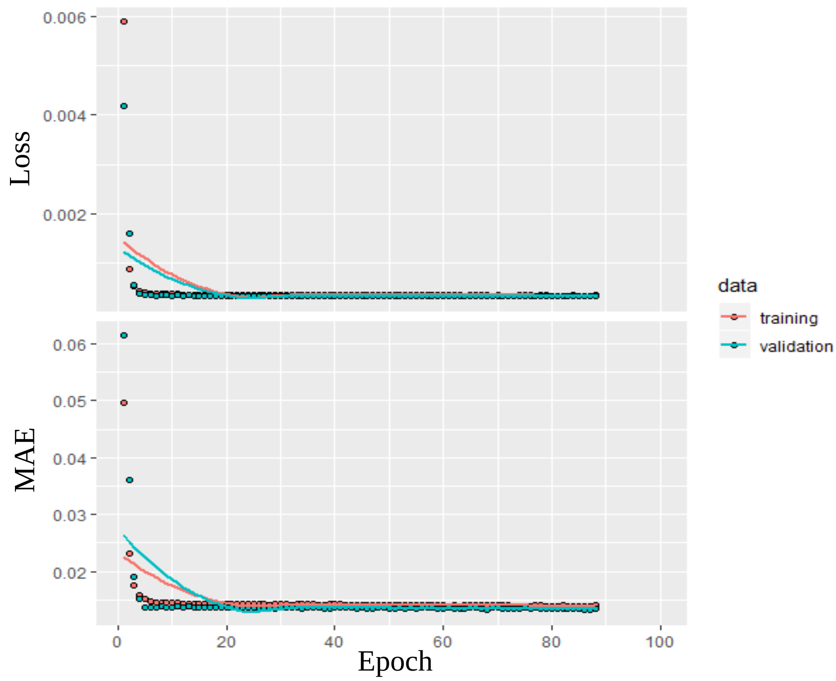

Figure 09: Training history plot of temperature prediction LSTM model

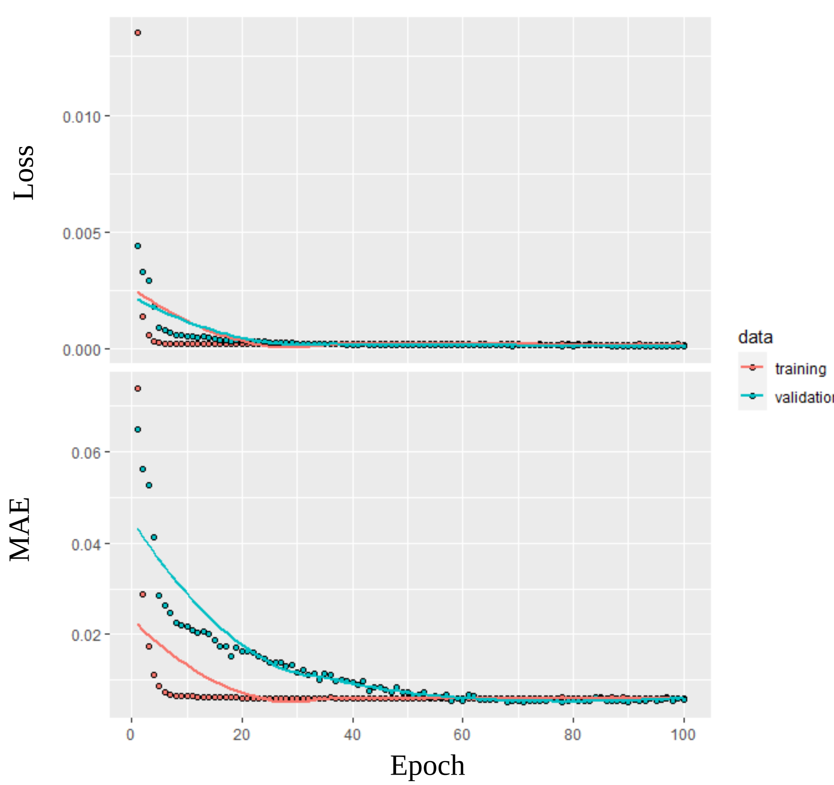

Figure 10: Training history plot of precipitation prediction LSTM model

Figure 09 shows the loss and MAE for the training set and the validation set are gradually decrease with the number epochs increases in the training of the temperature prediction LSTM model. The same behavior can be observed in the figure 10 which reflects the training history plot of the precipitation prediction LSTM model.

2.5 Testing LSTM Models

Making predictions on the test sets was done after the training of the two LSTM models is completed. Testing is done using separate test sets.

3 RESULTS & DISCUSSION

The performance of the developed LSTM models on predicting temperature and precipitation was evaluated using RMSE and MAE metrics. RMSE is the measure of standard deviation of the prediction errors which is calculated using following equation

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}\,. \qquad (7)$$

MAE is the measure of errors between predicted versus observed observations which is calculated using below equation.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\,y_i - \hat{y}_i|\,. \qquad (8)$$

Above evaluation metrics measures the difference between actual observations in the test set and the predicted results by the 2 models. Therefore, the accuracy of the model increases when the RMSE and MAE values decrease. These values are calculated on the model predictions made on the test set. Calculated RMSE, MAE values for temperature prediction results and the precipitation prediction results of two LSTM models are as follows.

Table 02: Performance Evaluation of two LSTM models

Model	RMSE	MAE
Temperature Prediction LSTM	1.37482° F	0.9898567° F
Precipitation Prediction LSTM	0.6658987 (Inches)	0.3365786 (Inches)

According to the table 02, temperature prediction LSTM model has shown 1.37482°F of RMSE value, 0.98985°F of MAE value for temperature predictions on test set for Colombo weather station. Precipitation prediction LSTM model has shown 0.6658987 (Inches) of RMSE value, 0.3365786 (Inches) of MAE value for precipitation predictions on test set for Colombo weather station.

Figure 11, presents the predicted temperature values by the temperature prediction LSTM model with the actual temperature observations of the data set while figure 12, shows the predictions made by the temperature prediction LSTM model with the actual temperature observations of the test set. Figure. 13, shows the scatterplot of the precipitation prediction LSTM model's predictions on observed actual precipitation data in the entire data set. Figure. 14, presents the precipitation predictions made by the precipitation predicting LSTM model with the actual precipitation observations of the test set.

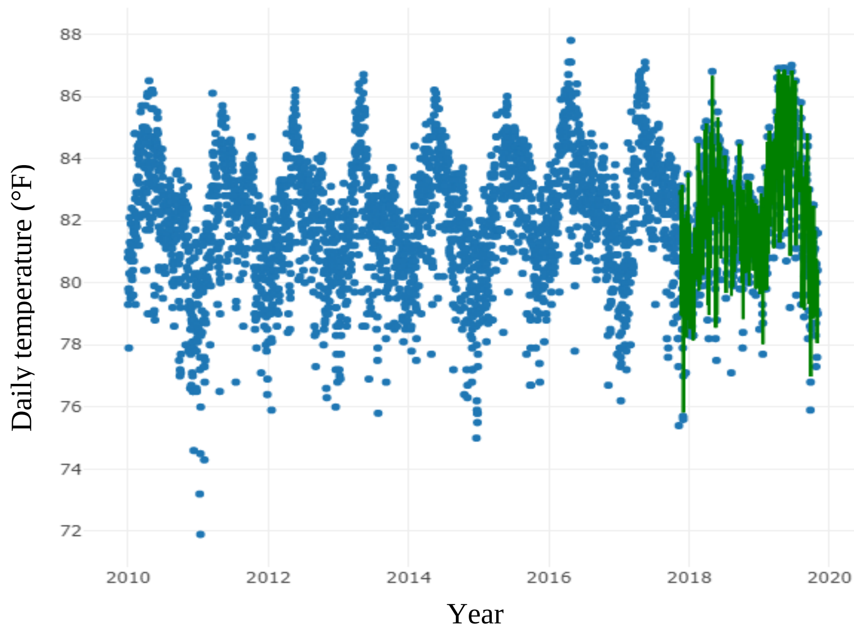

Figure 11: Model prediction on whole temperature observation dataset

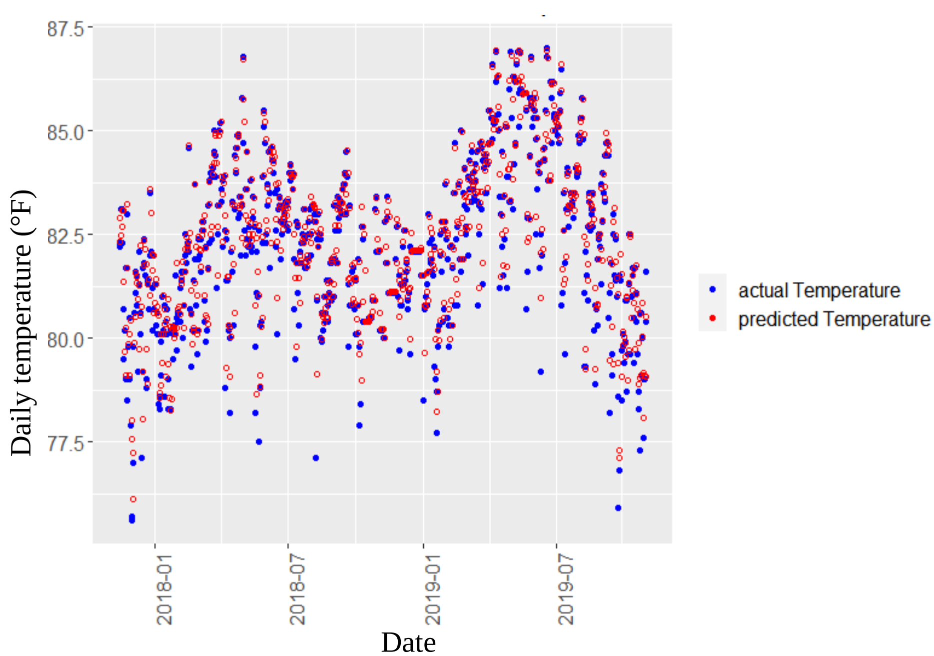

Figure 12: predicted temperature with the actual observations of the test set

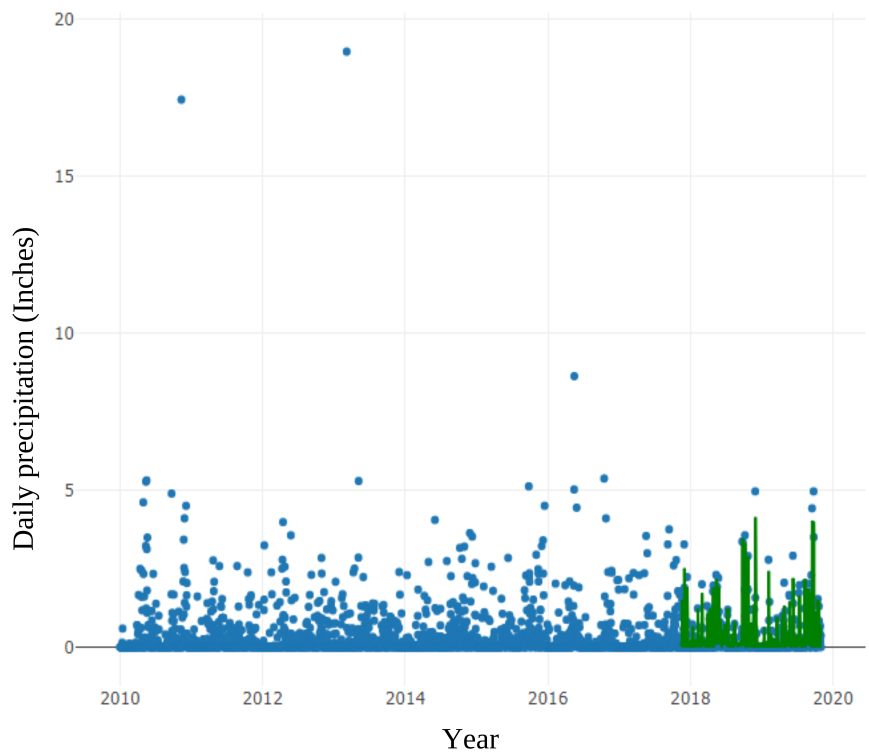

Figure 13: Model prediction on whole precipitation dataset

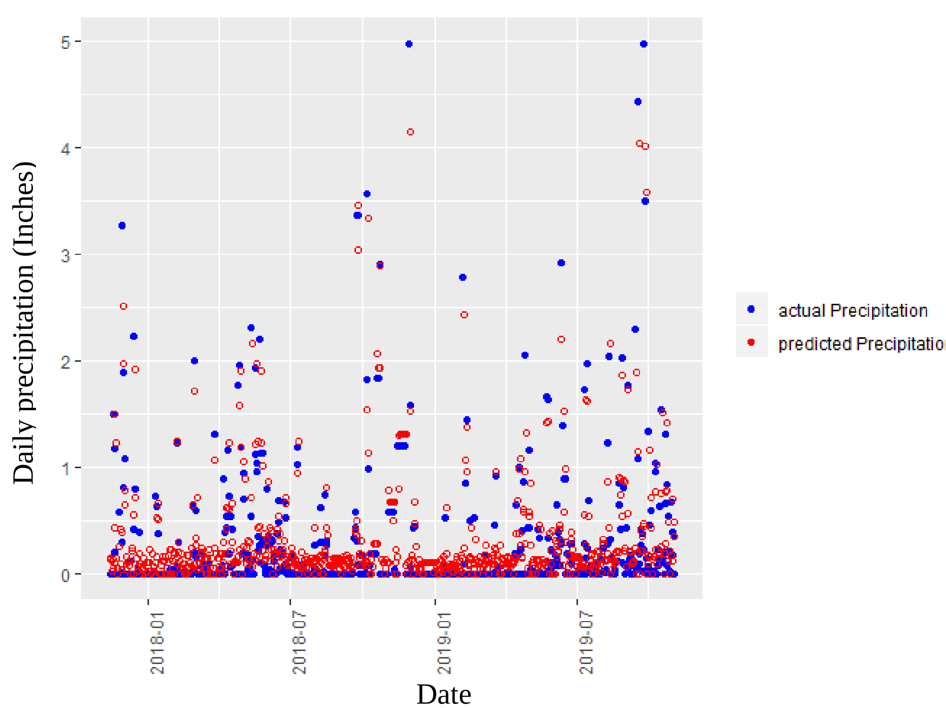

Figure 14: predicted precipitation values with the actual observations of the test set

Though the RMSE and MAE values for temperature and precipitation predictions are low, figure 12 and figure 14 show that there are considerable deviations between the actual and predicted values for some of the test data points. Table 03 and table 04 provide a better insight of the deviations of the predicted and actual temperature and precipitation values of the test set. These tables provide the percentage of test data points that belongs to each % of the actual error.

Table 03: Percentage of temperature predictions and the percentage of actual error

Percentage of actual error %	Test Data Points	Percentage of predictions %
0-0.02	573	79.8050
0.02-0.04	121	16.8523
0.04-0.06	21	2.9247
0.06-0.08	3	0.4178

Table 04: Percentage of precipitation predictions and the percentage of actual error

Percentage of actual error %	Test Data Points	Percentage of predictions %
0-10	587	81.7548
10-20	59	8.2172
20-30	34	4.7353
30-40	21	2.9247
40-50	6	0.8356
50-60	4	0.5571
60-70	4	0.5571
70-80	0	0.0000
80-90	1	0.1392
90-100	2	0.2785

According to the table 03, temperature prediction LSTM model has shown higher accuracy predicting test data points. The actual error between predicted and actual test data is extremely low. The maximum error shown by the Temperature prediction LSTM is 0.0727% which is very low and the majority of test data points are belongs to the minimum percentage error between 0-0.02%.

According to the table 04, precipitation prediction LSTM model has predicted the majority of test data points (81.7548% of data points) with minimum error (0-10%). The maximum error for the predicted precipitation data is 99.7078% which is encountered for a data point with actual precipitation of 4.96 Inches which can be one of the outliers of the precipitation dataset. Few outliers can be detected in the precipitation dataset which have affected the accuracy of the precipitation predictions in negative way.

Above results have demonstrated how well two LSTM models have uncovered underlying weather patterns in the historical training datasets and tried to adhere their predictions to those patterns. The results of the current research show that the temperature prediction LSTM has given extremely high accuracy predictions with minimum error between actual and predicted points of the test data set while precipitation prediction LSTM model has predicted majority of test data points with minimum error. Even though the most of the researches have discussed the performances of their models using RMSE or MAE, the current research has provided in depth insight for the prediction errors and performances of two LSTM models by analyzing actual percentage errors between predictions and each data points in the test set.

4 CONCLUSION & RECOMMENDATIONS

In current study, we experimented predictions of temperature and precipitation using several weather attributes as inputs for Colombo weather station in Sri Lanka using two separate LSTM models. The evaluation of the performance of the two LSTM models has been done using standard and popular evaluation metrics. The results have shown low RMSE and MAE values for temperature predictions and precipitation predictions made by two models on the set aside test set.

Even though predicting weather is challenging in a tropical country like Sri Lanka, the results of the study has proven that LSTM networks’ potential of identifying highly nonlinear and complex patterns of historical weather data and therefore can be used to make predictions for weather attributes such as temperature, precipitation with good accuracy. Also the results of the current research confirm that LSTMs have the ability of recognizing underlying patterns and long-term dependencies of time series and making reliable predictions based on past observations. The errors of the ground observation data may impact the prediction accuracy of models. The RMSE and MAE error values can be further reduced by using a larger dataset than the dataset used in the current study since LSTM networks perform well in larger datasets and by increasing the complexity of the LSTM models.

The two LSTM models implemented in the current study can be used to make predictions on any other weather attribute by making simple changes to the model architecture and the training process. The current study suggests that there should be more attention in the future on developing machine learning based weather prediction models in order to overcome the current limitations of weather prediction in tropical countries by replacing traditional numerical weather prediction models.

ACKNOWLEDGMENTS

NOAA (National Oceanic and Atmospheric Administration), NCDC (National Climate Data Centre) is acknowledged for collecting the weather dataset used in this study.

REFERENCES

Holmstrom, M, Liu, D & Christopher, V 2016, ‘Machine Learning Applied to Weather Forecasting’ Stanford University.

Coiffier, J 2011, *Fundamentals of Numerical Weather Prediction*, Cambridge University Press, New York, p. 1,11,12,13,14,15,16,17,18,19,20,21.

Larraondo, P,R, Inza, I & Lozano, J,A 2017, ‘Automating Weather Forecasts based on Convolutional Networks’, ICML 17 Workshop on Deep Structured Prediction, Sydney, Australia, PMLR 70

Warner, T,T 2011, *Numerical Weather and Climate Prediction,* Cambridge University Press, New York p. 6,7,10,11,12.

Krasnopolsky, V , Fox-Rabinovitz, M,S & Belochitski, A 2008, 'Using Neural Network Emulations of Model Physics in Numerical Model Ensembles', IEEE World Congress on Computational Intelligence, Neural Networks, IJCNN.

Elhoseiny, M, Huang, S & Elgammal, A 2015, 'Weather Classification with Deep Convolutional Neural Networks', IEEE international conference on Image Processing (ICIP), Quebec City, QC, Canada.

Mohammed, M, Kolapalli, R, Golla, N & Maturi S, S 2020, 'Prediction of Rainfall Using Machine Learning Techniques', International Journal of Scientific & Technology Research, vol.9, No.1.

Abdul-Kader, H, Salam, M, A & Mohamed, M 2020, 'Hybrid Machine Learning Model for Rainfall Forecasting', Journal of Intelligent and Internet of Things, vol.1, No.1.

Zaytar, M,A & Amrani, C,E 2016, 'Sequence to Sequence Weather Forecasting with Long Short-Term Memory Recurrent Neural Networks', International Journal of Computer Applications, vol. 143, No.11.

Aswin, S, Geetha, P & Vinayakumar, R 2018, 'Deep Learning Models for the Prediction of Rainfall', International Conference on Communication and Signal Processing (ICCSP).

Salman, A,G, Heryadi, Y, Abdurahman, E & Suparta, W 2018, 'Single Layer & Multi-layer Long Short-Term Memory (LSTM) Model with Intermediate Variables for Weather Forecasting', 3rd International Conference on Computer Science and Computational Intelligence.

Singh, S, Kaushik, M, Gupta, A & Malviya, A,K 2019, 'Weather Forecasting using Machine Learning Techniques', 2nd International Conference on Advanced Computing and Software Engineering (ICACSE) .

Thilakarathne, H & Premachandra, K 2017, 'Predicting Floods in North Central Province of Sri Lanka using Machine Learning and Data Mining Methods', International Conference on Artificial Intelligence.

Narvekar, M & Fargose, P 2015, 'Daily Weather Forecasting using Artificial Neural Network', International Journal of Computer Applications (0975 – 8887), vol. 121, No.22.

Subashini, A, Thamarai, S,M & Meyyappan, T 2019, 'Advanced Weather Forecasting Prediction using Deep Learning', International Journal for Research in Applied Science and Engineering Technology (IJRASET), vol. 7, pp. 939-945.