International Journal of Applied Information Systems (IJAIS) – ISSN : 2249-0868
Foundation of Computer Science FCS, New York, USA
Volume 12– No.46, January 2025 – www.ijais.org

A Comparative Study of Long Short-Term Memory and Gated Recurrent Units for Forecasting Rainfall: A Case Study of Nigeria

Oguche Ajah
Department of Computer Science
Federal University of Technology, Akure, Nigeria

Abimbola H. Afolayan
Department of Information Systems
Federal University of Technology, Akure, Nigeria

Akintoba E. Akinwonmi
Department of Computer Science
Federal University of Technology, Akure, Nigeria

ABSTRACT
Heavy rainfall may trigger floods, destroying human life and the agricultural economy. Accurate rainfall forecasting is vital for mitigating the effects of severe flooding. This study compares deep supervised learning (DSL) algorithms for rainfall forecasting in Nigeria, using a large dataset from ECMWF archives that includes Kaduna, Bauchi, Kwara, Imo, Ondo, and Cross River states representing the six geopolitical zones in Nigeria. Critical attributes in the dataset include rainfall, relative humidity, temperature, dew point, surface pressure, and wind speed. Following meticulous preprocessing and outlier removal, feature engineering was done using Pearson correlation and Spearman rank to ensure the most important predictors were chosen. Two sophisticated models—Long Short-Term Memory (LSTM), and Gated Recurrent Unit (GRU) were developed, trained, and tested extensively. The models were assessed using a variety of metrics, including Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), R-squared, and Accuracy. The findings were impressive, with the GRU model surpassing the LSTM model with the average MSE, MAE, R-Square, RMSE, and Accuracy values of 3.0504, 1.1644, 0.7978, 1.7381, and 98.43%, respectively. These findings illustrate GRU's exceptional capacity to capture complicated patterns of rainfall data, making it an effective tool for meteorological forecasting.

Keywords
Rainfall forecasting in Nigeria, Deep Supervised Learning Algorithms, Long Short-Term Memory, Gated Recurrent Unit, Evaluation Metrics

1. INTRODUCTION
The world's weather constantly changes, and rainfall forecasts are typically fascinating and helpful information for all societal segments. It is crucial to remember that rainfall forecasts affect not just the daily routines of individuals but also agriculture and various industries. The decision-making procedures used by different organizations to prevent disasters can also benefit from these forecasts. More specifically, rainfall matters for water resource management, food production strategies, agriculture, and other natural systems [15].

However, because of the intricate structure of weather forecasting modeling, many researchers still find it challenging to create an efficient prediction system for rainfall forecasting. Precise forecasting necessitates applying cutting-edge computer modeling and simulation techniques because of the extreme complexity and nonlinearity of weather and rainfall [7].

Predicting the rainfall conditions in a region can capture its diverse climatic variations, enhance forecast precision, deepen comprehension of how climate change influences rainfall trends, aid in devising strategies to alleviate its effects, assist in formulating water resource management plans, and support measures to manage extreme weather occurrences [24]. Moreover, as climate change continues to amplify extreme weather events [25], ensuring dependable rainfall forecasts is essential for enhancing Nigeria's disaster readiness and response efforts.

Historically, Nigeria's rainfall forecasting was initially made using statistical techniques and mathematical weather forecasting models. Although these techniques have yielded valuable insights, they often fall short of capturing the intricate and non-linear patterns present in rainfall datasets (Bai, 2024). On the other hand, machine and deep learning techniques typically require less processing power and are less expensive to implement [20].

The emergence of machine learning in recent years has brought a promising method to forecasting challenges. The potential for machine learning algorithms to reveal hidden patterns and relationships within rainfall datasets is enhanced by the availability of large volumes of weather records and increases in processing capacity. A subfield of machine learning (ML) called "deep learning" (DL) is inspired by the way the human brain processes information. Unlike traditional ML techniques, DL may learn feature sets for different jobs. [12].

DL approaches fall into three categories: semi-supervised, supervised, and unsupervised. Deep supervised learning (DSL) operates with labeled data. Deep neural networks (DNNs), convolutional neural networks (CNNs), and recurrent neural networks (RNNs) are a few examples of DSL techniques. The types in the RNN category are gated recurrent units (GRUs) and long short-term memory (LSTM) approaches [1]. DSL's main advantage is its ability to collect data or generate data output from prior information.

Recent research ([8]; [14]; [6]; [19]; [5]; [20]; [9]; and [23]) showcases the application of deep learning models in forecasting rainfall across diverse countries, indicating a growing interest in leveraging advanced techniques for predictive analytics. However, the design of an effective prediction system for rainfall forecasting in Nigeria using deep learning techniques is still in its early stages within Nigeria. However, designing an effective prediction system for rainfall

International Journal of Applied Information Systems (IJAIS) – ISSN : 2249-0868
Foundation of Computer Science FCS, New York, USA
Volume 12– No.46, January 2025 – www.ijais.org

forecasting in Nigeria using deep learning techniques is still in its early stages.

Thus, this research aims to develop LSTM and GRU deep-supervised learning models for daily rainfall forecasting for six states across the country, namely Kaduna, Cross River, Imo, Ondo, Bauchi, and Kwara, representing Nigeria's six geopolitical zones. The performances of these models are evaluated using the RMSE, MAE, MSE, R^2, and Accuracy. This investigation will contribute to the advancement of rainfall forecasting in Nigeria, thereby supporting sustainable development, improving livelihoods, enhancing resilience to climate change, and promoting economic growth across diverse sectors of the economy.

2. RELATED WORK

Many researchers have taken significant steps to solve rainfall forecasting problems with the use of LSTM and GRU.

2.1 LSTM for Rainfall Prediction

[11], employed Long Short-Term Memory, Artificial Neural Networks, Decision Trees, Random Forest algorithms, and Support Vector Regression in forecasting rainfall patterns. Employing average weekly rainfall data from the Kuala Krai station, the study meticulously scrutinizes the performance of these models. The findings demonstrate that the Long Short-Term Memory (LSTM) model outperforms other models., showcasing the most promising outcomes with notably lower values for both root mean squared error (RMSE) and mean absolute error (MAE).

[21], introduced an innovative rainfall prediction approach utilizing Artificial Intelligence, specifically LSTM and Neural Network architectures. Their study addressed the limitations of traditional rainfall prediction models and demonstrated the advantages of advanced AI and deep learning techniques. By incorporating LSTM and RNN models, they aimed to improve prediction precision, particularly for agricultural applications. Their model achieved a 76% accuracy rate with data from six diverse regions, showcasing its effectiveness in predicting rainfall.

[16], introduced two hybrid forecasting models, wavelet-LSTM (WLSTM) and convolutional LSTM (CLSTM), to enhance the accuracy of monthly streamflow and rainfall predictions. By combining LSTM networks with wavelet transforms and convolutional neural network techniques, they compared these models with a benchmark model (MLP). Their results showed that LSTM models, especially with integrated wavelet and convolutional layers, significantly improve forecast accuracy, particularly for long-term predictions.

[20], evaluates the performance of Long Short-Term Memory (LSTM) networks for rainfall prediction, showing that LSTMs generally outperform traditional Artificial Neural Networks (ANNs) in terms of metrics like MSE, RMSE, and MAE. The use of Seasonal Decomposition for data analysis and preprocessing adds to the robustness of the LSTM model. However, the study highlights that LSTMs struggle with outliers, leading to higher error rates in those cases. This limitation suggests a need for further research to improve outlier handling in LSTM-based models.

[22], proposed a rainfall forecasting model for Sub-Saharan Africa, specifically in Ghana, utilizing the LSTM deep learning approach. The primary objective of their research was to evaluate the performance and efficiency of an LSTM-based model for predicting rainfall in Axim, Ghana. The study also addresses the challenges associated with rainfall forecasting and underscores the importance of accurate predictions. The results indicate that the proposed model outperforms alternative methods, as evidenced by lower MSE and RMSE values.

[17], developed a weather forecasting model using the LSTM algorithm and time series data. Their study details how the LSTM model, trained on weather parameters, delivers accurate predictions of future weather conditions. While the research focused on temperature, humidity, wind speed, and rainfall, the authors suggested that incorporating additional meteorological variables could further enhance the model's performance.

[19], present a cutting-edge deep learning-powered TensorFlow solution for revolutionizing rain prediction in meteorology and emergency management. CNN was merged with LSTM networks to accurately forecast precipitation by capturing complex spatial and temporal relationships in meteorological data. By training the model on vast amounts of historical weather data, the CNN-LSTM model outperforms traditional statistical methods, enabling precise rainfall predictions.

[2], investigated monthly rainfall prediction in hyper-arid environments, specifically focusing on the United Arab Emirates. Their study analyzed a 30-year dataset (1991-2020) using various machine learning models, including XGBoost, LSTM, Random Forest, Gradient Boosting, Support Vector Machine, Multilayer Perceptron, Linear Regression, and ensemble methods. The research utilized both univariate and multivariate analyses to enhance prediction accuracy. The findings revealed that machine learning models, particularly XGBoost and LSTM, significantly outperformed other models in capturing the complex temporal patterns of rainfall data.

2.2 GRU for Rainfall Prediction

[13], present a dual deep-learning approach for rainfall forecasting in Thailand, integrating Convolutional Neural Networks (CNN) to analyze sensor relationships and Gated Recurrent Units (GRU) to capture time-series data. This cascading model was trained and tested using a six-year hourly rainfall dataset (2013-2018) provided by the Thai government. The model demonstrated robust performance across all regions, achieving an RMSE of 4.53%.

[3], developed a Deep BLSTM-GRU model for monthly rainfall prediction in Simtokha, Bhutan, and compared its performance with existing machine and deep learning models. Their hybrid model, combining Bidirectional Long Short-Term Memory (BLSTM) and Gated Recurrent Unit (GRU), achieved superior results with a Mean Squared Error (MSE) of 0.0075.

[18], developed a rainfall prediction model using a Gated Recurrent Unit (GRU), incorporating the Dipole Mode Index (DMI), Niño3.4 Index, temperature, humidity, sunlight duration, and wind speed. Utilizing monthly time series data, the study compared the performance of 1D-CNN, RNN, LSTM, and GRU algorithms. The GRU model demonstrated

International Journal of Applied Information Systems (IJAIS) – ISSN : 2249-0868
Foundation of Computer Science FCS, New York, USA
Volume 12– No.46, January 2025 – www.ijais.org

superior performance, with 32 hidden neurons, a batch size of 32, and a learning rate of 0.001, achieving a Mean Arctangent Absolute Percentage Error (MAAPE) of 0.42 and an R-squared value of 0.79.

[27], proposed a novel precipitation forecasting approach using CEEMD-WTD-GRU. This method aims to enhance the accuracy of regional precipitation predictions by incorporating Wavelet Threshold Denoising (WTD) and Complete Ensemble Empirical Mode Decomposition (CEEMD). WTD targets high-frequency components (IMF1–IMF3) to minimize signal noise, while CEEMD addresses modal aliasing. The Gated Recurrent Unit (GRU) model is employed to manage long-term memory and reflection issues. The study demonstrated that the second decomposition of CEEMD-WTD-GRU effectively extracts complex time series data, addressing challenges of uncertainty and unpredictability in precipitation forecasting. However, the study noted the absence of a noise reduction procedure in the CEEMD model used.

[6], introduced a Gated Recurrent Unit for rainfall forecasting in Pakistan. This model was trained on 30 years of climatic data (1991-2020). The performance of the model was assessed by comparing it with RNN, LSTM, DNN, and CNN using Normalized Mean Absolute Error (NMAE) and Normalized Root Mean Squared Error (NRMSE). The GRU model demonstrated superior accuracy compared to other models.

3. RESEARCH METHOD

This study used two deep-supervised learning algorithms (LSTM, and GRU) to forecast rainfall in six states (Kaduna, Bauchi, Kwara, Imo, Ondo, and Cross-River) representing Nigeria's six geopolitical zones. The section consists of three major sub-sections. The first and second subsections discuss the data collection and preprocessing respectively, while the third sub-section discusses the experimental approach utilized in this study. The stages of the research architecture are logically ordered and described in Figure 1.

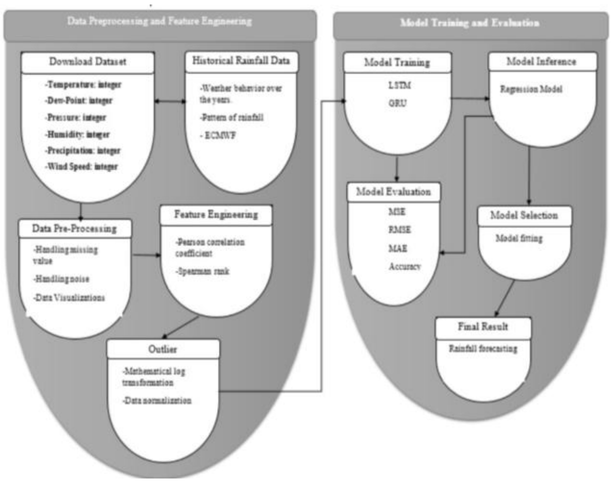

Figure 1. Conceptual Framework

3.1 Data Collection

The rainfall dataset used in this research was downloaded from the archive of the European Centre for Medium-Range Weather Forecast (ECMWF) ERA5. Its spatial resolution is 0.25 degrees by 0.25 degrees. The downloaded dataset was in .netcdf (NC) format and was extracted into .csv, which is suitable for machine learning using a Ferret. The dataset spans 31 years (1992-2022) for six states (Ondo, Cross River, Imo, Kwara, Bauchi, and Kaduna), each representing one of Nigeria's six geopolitical zones. Each state's dataset has 11,323 rows and encompasses a range of in-situ atmospheric surface parameters such as temperature, rainfall, wind speed, pressure, humidity, and dew point.

3.2 Data Preprocessing

To minimize the impact of outliers and variations in time-series data on the models and to accurately identify underlying patterns, preprocessing techniques were employed. This study employed Matplotlib and Seaborn libraries to visualize the datasets collected from ECMWF. The visualization's primary goal was to comprehend the rainfall dataset's essential features thoroughly. A histogram was used to represent the distribution of the critical parameters. Upon visual inspection, it was noted that the dataset contained missing values. A data cleaning subtask was undertaken to rectify this by removing the rows containing missing data in cases where the number of missing values was judged to be reasonably minimal. The remaining data preprocessing techniques are detailed in the following sub-sections.

3.2.1 Outlier Detection and Removal

The analysis indicated that outliers were primarily found in the rainfall parameters. To address these outliers, a mathematical log transformation was utilized to diminish their impact. This technique involves applying a logarithmic function to the data, which compresses the range of values and reduces the extremity of outliers. The log transformation is applied to each data point, as represented by Equation (1), and the graphical representation of the outlier removal is shown in Figure 2.

$$y = log(x) \qquad (1)$$

where y is the transformed data, and x is the data point.

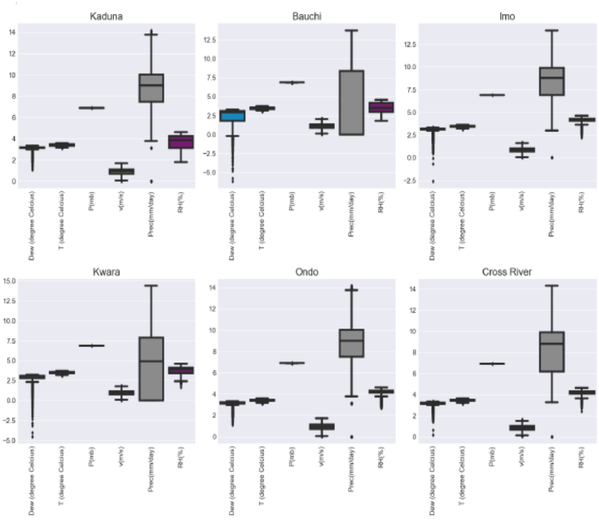

Figure 2: Graphical Representation of Outlier Removal

3.2.2 Feature Engineering

Pearson correlation coefficients represented by Equation (2) were employed to examine the dataset's properties and determine any associations between rainfall and other parameters.

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}} \quad (2)$$

where r is the Pearson correlation coefficient, (x_i) and (y_i) are the individual data points of variables x and y, and $(\bar{x})$ and $(\bar{y})$ are the means of variables x and y, respectively.

The correlation heatmap results for Kaduna, Bauchi, Kwara, Imo, Ondo, and Cross-River states are presented in Figures 3, 4, 5, 6, 7, and 8 respectively.

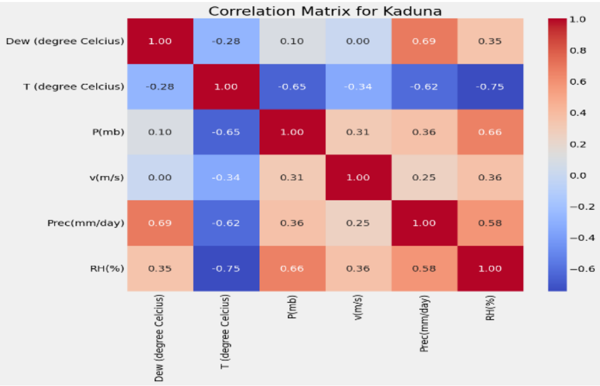

Figure 3. Correlation Heatmap for Kaduna

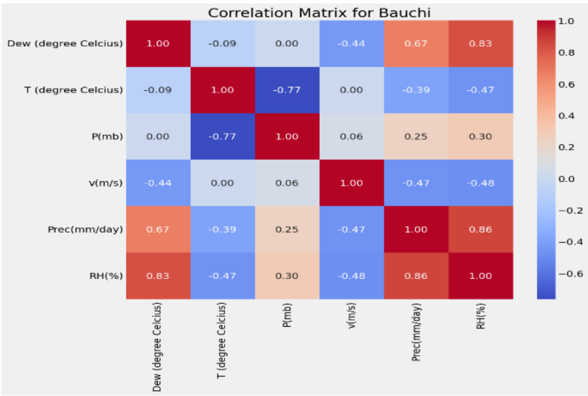

Figure 4. Correlation Heatmap for Bauchi

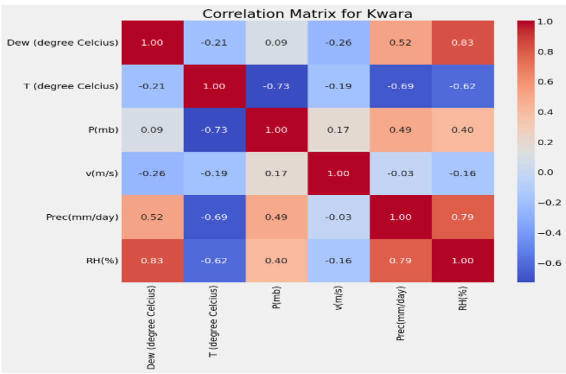

Figure 5. Correlation Heatmap for Kwara

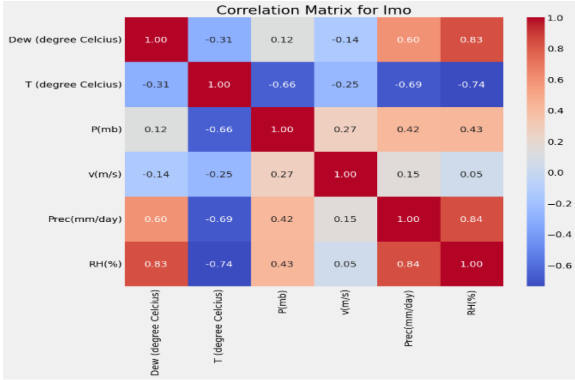

Figure 6. Correlation Heatmap for Imo

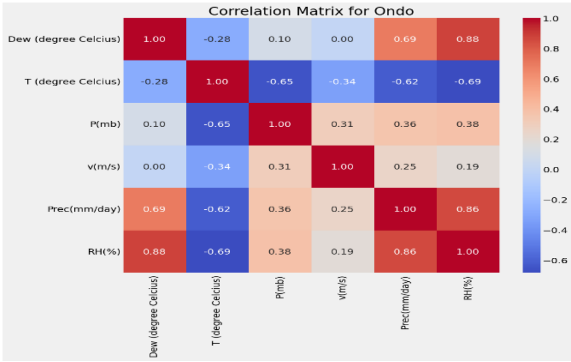

Figure 7. Correlation Heatmap for Ondo

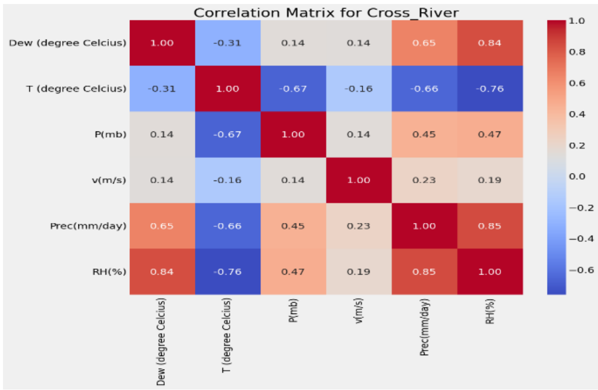

Figure 8. Correlation Heatmap for Cross-River

The Spearman rank correlation coefficient was used to identify monotonic correlations. When dealing with non-normally distributed or ordinal data, Spearman correlation is beneficial. The analysis's conclusions add to our understanding of the dataset's underlying structure on a more thorough level. The Spearman's correlation for non-linear relationships is mathematically expressed as shown in Equation (3) and the results are shown in Table 1.

$$r = \frac{\sum_{i=1}^{n} R(x_i)R(y_i) - n\left(\frac{n+1}{2}\right)^2}{\sqrt{\sum_{i=1}^{n} R(x_i)^2 - n\left(\frac{n+1}{2}\right)^2}\sqrt{\sum_{i=1}^{n} R(y_i)^2 - n\left(\frac{n+1}{2}\right)^2}} \quad (3)$$

where r is the resulting value, which will be between -1 and 1, $R(x)$ is the rank of variable x, and $R(y)$ is the rank of variable y.

Table 1. Spearman's Correlation Coefficient

S/No	State	Parameters					
		Prec and Prec	Prec and Temp	Prec and Pressure	Press and H	Prec and Dew-P	Prec and W-Speed
1	Kaduna	1	0.62	0.36	0.58	0.69	0.25
2	Bauchi	1	-0.39	0.25	0.86	0.67	-0.47
3	Kwara	1	-0.69	0.49	0.76	0.52	-0.03
	Imo	1	-0.69	0.42	0.84	0.6	0.15
5	Ondo	1	-0.62	0.36	0.86	0.69	0.25
6	C-River	1	-0.66	0.45	0.85	0.65	0.23
	Average	**1**	**-0.405**	**0.3883**	**0.7917**	**0.6367**	**0.0633**

The results indicate that, across all the six geo-political zones in Nigeria, Relative Humidity (RH%) and Dew Point are consistently strong predictors of precipitation due to their direct reflection of atmospheric moisture. In contrast, atmospheric pressure and wind speed have moderate to weak correlations, as their effects on precipitation are less direct and variable. Temperature has a strong negative correlation with precipitation, likely because high temperatures are often associated with dry conditions and lower moisture levels.

3.2.3 Data Normalization

In this study, the min-max normalization method defined in Equation (4) with an interval of 0 to +1 was employed to eliminate noise and incomplete information.

$$Y = \frac{(x - x_{min})}{(x_{max} - x_{min})} \quad (4)$$

Where x represents the data to be normalized x_min indicates the minimum value within a specific feature category, x_max denotes the maximum value within the same feature category, Y represents the normalized data.

3.2.4 Dataset Splitting

The total dataset has 11,323 rows, one for each state. The study uses an 80:20 split ratio to produce different training and testing sets. The dataset consists of seven features for each instance, each representing a different parameter. The training set consists of 9,058 samples, which uses 80% of the dataset. The model has many learning possibilities from this large amount of data, which helps it identify underlying patterns and relationships. The test set consists of 2,265 samples and makes up 20% of the data. This distinct collection is essential for assessing the model's effectiveness using fresh, untested data.

3.3 Experimental Approach

This sub-section discusses the experimental approach utilized in this study, including the LSTM and GRU deep-supervised learning models, the evaluation metrics used in assessing the models, and the experimental setup.

3.3.1 Long Short-Term Memory

The Long Short-Term Memory Network (LSTM), introduced by Hochreiter and Schmidhuber (1997), is a specialized type of recurrent neural network (RNN) designed to recognize long-term dependencies in sequential data. Unlike traditional RNNs, which struggle with issues such as vanishing gradients, LSTMs address these challenges by using a unique architecture with four interacting layers instead of one. The expanded configuration of a single-layer LSTM cell is illustrated in Figure 9.

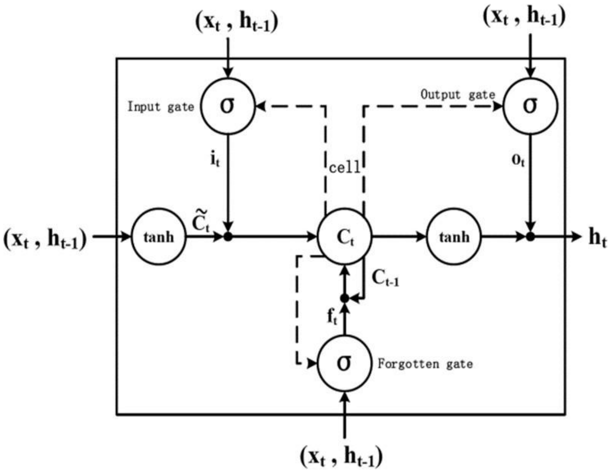

Figure 9. Schematic Representation of LSTM Memory Cell

This architecture includes a "memory unit" or cell that retains information over extended periods, resolving the long-range dependency problem. LSTM networks feature three key components: an input gate, an output gate, and a forget gate. These gates regulate the flow of information into and out of the memory cell, ensuring that only relevant information is retained. The forget gate uses a sigmoid function to control the information based on the previous output and current input, thereby maintaining effective long-term memory. The design of these gates helps LSTMs manage sequential data more efficiently than standard RNNs.

At time t, the memory cell module input is denoted as X_t, the output as H_t, and the unit status as C_t. Equations (5) to (10) outline the mathematical expressions governing the input gate, forget gate, output gate, input conversion, unit state update, and hidden layer output of the memory cell module.

$$\mathrm{i_t} = \sigma(\mathrm{w}_{xi}x_t + \mathrm{w_{hi}h_{t-1}} + \mathrm{b_i}) \quad (5)$$

$$\mathrm{f_t} = \sigma(\mathrm{w_{xf}}x_t + \mathrm{w_{hf}h_{t-1}} + \mathrm{b_f}) \quad (6)$$

$$o_t = \sigma(w_{xo}x_t + w_{ho}h_{t-1} + b_o) \quad (7)$$

$$\widetilde{C_t} = tanh(wx_t + w_{hC}h_{t-1} + b_C) \quad (8)$$

$$C_t = f_t \otimes C_{t-1} + i_t \otimes \widetilde{C_t} \quad (9)$$

$$h_t = o_t \otimes \tanh(C_t) \quad (10)$$

Where σ denotes the sigmoid function; tanh represents the hyperbolic tangent function; $\mathrm{i_t}$, $\mathrm{f_t}$, $o_{t,}$ and C_t signify the input gate, forget gate, output gate, and conversion unit of the input, respectively; w_{xi}, $\mathrm{w_{xf}}$, w_{xo}, $w_{xC,}$ and $\mathrm{w_{hi}}$, $\mathrm{w_{hf}}$ w_{ho} and W_{hC} correspond to the input weights associated with the input gate, forget gate, output gate and input transition for x_t and h_{t-1}

respectively. Additionally, b_i, b_f, b_o, and b_C denote the offset vectors of the input gate, forget gate, output gate, and input transformation, respectively.

3.3.2 Gated Recurrent Unit

Chung et al. (2014) addressed the vanishing gradient issue with conventional RNN by introducing Gated Recurrent Units (GRU). Because of its comparable designs and often equally exceptional results, GRU is similar to LSTM. GRU uses reset and update gates to solve the typical RNN's vanishing gradient issue. The reset and update gates are the two vectors deciding which data should be transferred to the output. What makes them unique is their ability to learn to remember information from the past without losing it over time or eliminating data that is not needed for the forecast. This simplification, using only two gates (update and reset), enhances computational efficiency during training, as illustrated in Figure 10.

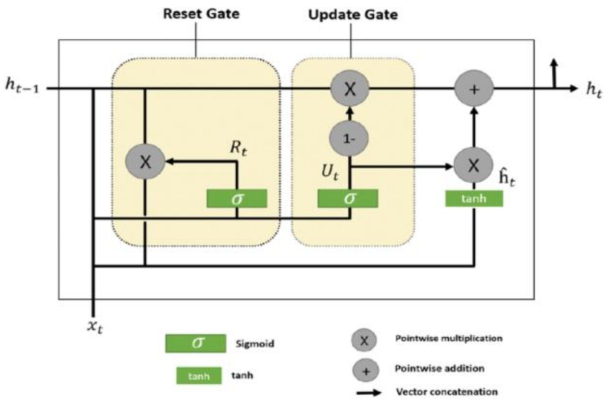

Figure 10. Schematic Representation of GRU Cell.

The reset gate controls the extent to which the previous hidden state is forgotten, as represented by Equation (11). In contrast, the update gate dictates the amount of update applied to the GRU unit and is mathematically expressed in Equation (12). The hidden state of the GRU is given by Equation (13). Additionally, the hyperbolic tangent function of the reset gate, known as the new memory gate, is detailed in Equation (14). A Sigmoid gate squashes numbers between 0 and 1, describing how much information should be passed through. Here, 1 means letting all the information go through the cell states. 0 means no information is allowed to pass the cell states. Tanh gates are used to put the value between −1 and 1.

$$R_t = \sigma(w_R * [h_{t-1,} x_t] + b_R) \quad (11)$$

$$U_t = \sigma(w_U * [h_{t-1,} x_t] + b_U) \quad (12)$$

$$\bar{h}_t = (1 - U_t) * h_{t-1} + U_t * h_t \quad (13)$$

$$h_t = \tanh(w_h * [R_t * h_{t-1}, x_t]) \quad (14)$$

Where R_t is the reset gate, U_t is the update gate, $\bar{h}_t$ is the hidden state, h_t is the hyperbolic tangent function of the reset gate, $\boldsymbol{\sigma}$ is the sigmoid activation function, h_{t-1} is the previous hidden state, x_t is the current input, w_R, w_U, and w_h terms represent weight matrices, b_R and b_U are the biases, respectively.

3.3.3 Evaluation Metrics

The evaluation metrics discussed in Equation (15) – (18) are used to compare the LSTM and GRU models:

(a) Mean Absolute Error (MAE): The Mean absolute error represents the average absolute difference between the actual and predicted values in the dataset.

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i| \quad (15)$$

where $\hat{y}_i$ represents the predicted values, y_i represents the observed values, n represents the total number of observations.

(b) Mean Squared Error (MSE): represents the average of the squared difference between the original and predicted values in the dataset.

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (16)$$

(c) Root Mean Squared Error (RMSE): is the square root of mean squared error. It measures the standard deviation of residuals.

$$\text{RMSE}=\sqrt{MSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (17)$$

(d) Accuracy is a metric that measures how often a machine learning model correctly predicts the outcome.

$$Accuracy = 100 - RMSE \quad (18)$$

3.3.4 Experimental Setup

The LSTM and GRU models were implemented using Python 3 and Google Colab's Jupyter Notebook on Google Colab (https://colab.research.google.com). The GPU backend in Google Colab was selected for training the models in this study to maximize the training procedure. Colab allocated 12.67 GB of RAM and 107.72GB of disk space, of which 0.76GB and 34.55GB were used for the training. This study examined various hyperparameters to train the models, Table 2. shows the final chosen hyperparameters.

Table 4. Model Training Parameters

S/N	Description	Value
1	Original data shape	(11323, 7)
2	Transformed data shape	(11323, 9)
3	Transformed train set shape	(9,058, 9)
4	Transformed test set shape	(2,265, 9)
6	Categorical features	None
7	Preprocess	TRUE
8	Imputation type	Simple
9	Numeric imputation	Mean
10	Categorical imputation	Mode
13	Normalize	TRUE
14	Optimizer	Adams
15	Epochs	30
16	Learning_rate	0.001
17	batch_size	32
18	Use CPU	TRUE

4. RESULTS AND DISCUSSION

This section presents the results of daily rainfall forecasts for six states (Kaduna, Cross River, Imo, Ondo, Bauchi, and Kwara) across six geopolitical zones in Nigeria, utilizing the

LSTM and GRU models. The evaluation metrics used to assess the performance of these models are also discussed.

4.1 Daily Rainfall Forecasting

The LSTM and GRU models successfully forecasted daily rainfall in the selected states in the six geo-political zones in Nigeria and the forecasted results of LSTM and GRU for Kaduna, Bauchi, Kwara, Imo, Ondo, and Cross-River are shown in Figures 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 and 22, respectively. These results offer insight into the effectiveness of LSTM and GRU in forecasting rainfall across different states.

Figure 11. Kaduna LSTM Daily Forecast

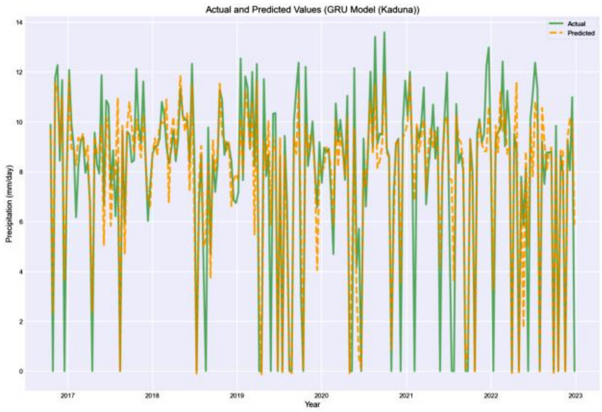

Figure 12. Kaduna GRU Daily Forecast

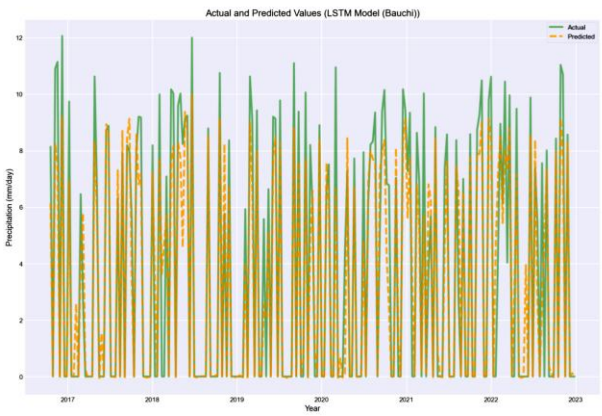

Figure 13. Bauchi LSTM Daily Forecast

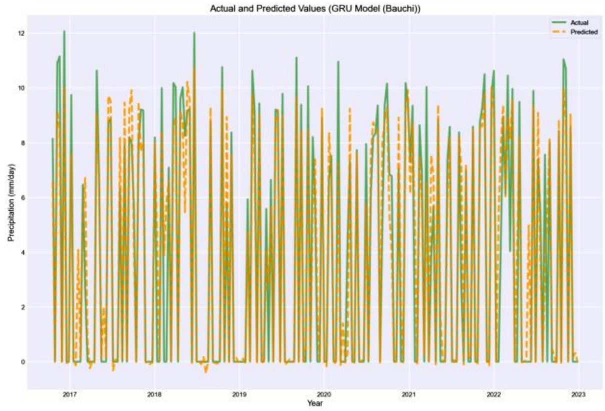

Figure 14. Bauchi GRU Daily Forecast

Figure 15. Kwara LSTM Daily Forecast

Figure 16. Kwara GRU Daily Forecast

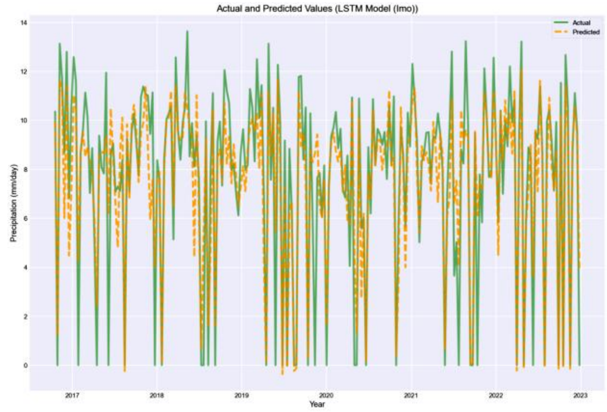

Figure 17. Imo LSTM Daily Forecast

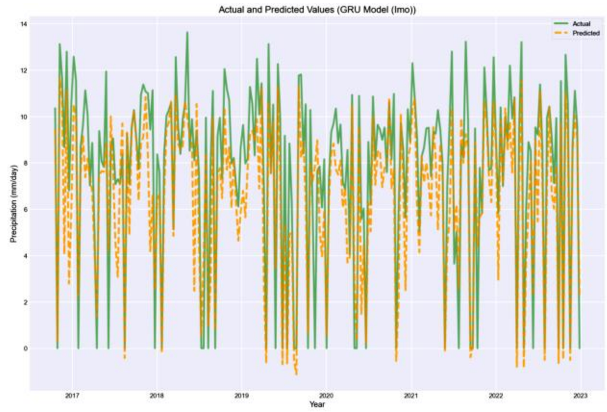

Figure 18. Imo GRU Daily Forecast

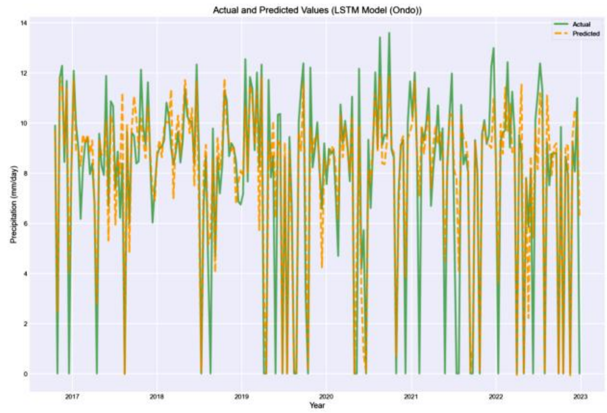

Figure 19. Ondo LSTM Daily Forecast

Figure 20. Ondo GRU Daily Forecast

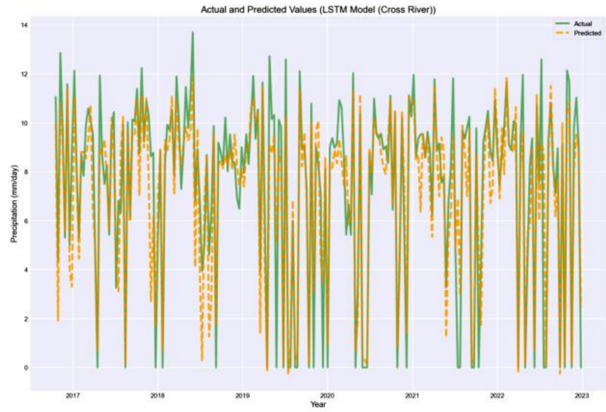

Figure 21. Cross River LSTM Daily Forecast

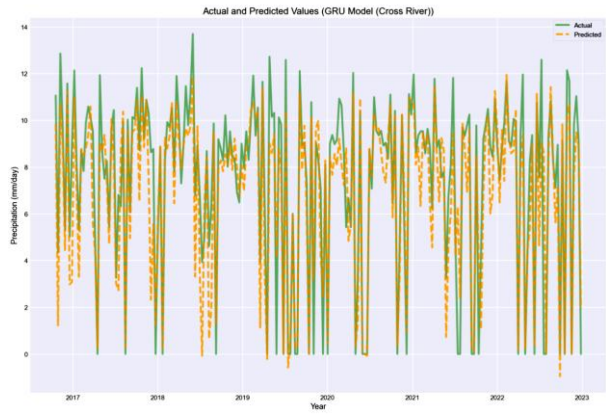

Figure 22. Cross River GRU Daily Forecast

4.2 Evaluation Metrics

The LSTM and GRU results were compared using standard evaluation metrics, and the results are shown in Tables 3 and 4, respectively. Table 3 indicates that LSTM has an average MSE, MAE, R2, RMSE, and Accuracy value of 3.3192, 1.2540, 0.7792, 1.8163, and 98.3705%, while Table 4 shows that GRU has an average MSE, MAE, R2, RMSE, and Accuracy values of 3.0504, 1.1644, 0.7978, 1.7381, and 98.4306%, respectively.

Table 3. Summary of LSTM Daily Evaluation Metrics

S/No	State	MSE	MAE	R^2	RMSE	Accuracy
1	Kaduna	2.5548	1.0634	0.8034	1.5984	98.4016
2	Bauchi	3.4773	1.0895	0.8193	1.8648	99.2564
3	Imo	2.9000	1.2115	0.7938	1.7029	98.2971
4	Kwara	4.2028	1.4495	0.7436	2.0501	97.9499
5	Ondo	3.2475	1.3470	0.7501	1.8021	98.1979
6	Cross-River	3.5329	1.3633	0.7654	1.8796	98.1204
	Average	**3.3192**	**1.2540**	**0.7792**	**1.8163**	**98.3705**

Table 4. Summary of GRU Evaluation Metrics

S/No	State	MSE	MAE	R^2	RMSE	Accuracy
1	Kaduna	2.5594	1.0777	0.8037	1.5998	98.4002
2	Bauchi	3.0457	0.9685	0.8417	1.7452	99.2668
3	Imo	2.8444	1.1878	0.7978	1.6865	98.3135
4	Kwara	4.3732	1.4815	0.7332	2.0912	97.9088
5	Ondo	2.4511	1.0741	0.8114	1.5656	98.4344
6	Cross-River	3.0287	1.197	0.7989	1.7403	98.2597
	Average	**3.0504**	**1.1644**	**0.7978**	**1.7381**	**98.4306**

[22], stated that models with lower MSE and RMSE values outperform others. Therefore, based on the results from Tables 5 and 6, it can be concluded that GRU outperformed the LSTM for forecasting rainfall in Nigeria. This may be because it offers a simpler, more efficient alternative to LSTM networks, requiring fewer resources and faster training while effectively capturing long-term dependencies. [26], stated that GRU's architecture is simpler and has one fewer gate than LSTM, which decreases matrix multiplication, thus GRU can save a lot of time without compromising performance Therefore, this study aligns with the assertion in these authors' work.

5. CONCLUSION

The study demonstrates the effectiveness of utilizing LSTM and GRU for daily rainfall forecasting in Nigeria by analyzing data from six states representing the six geopolitical zones in Nigeria over 31 years (1992-2022). Preprocessing, outlier detection and removal, and feature engineering were performed on the dataset. The performances of the LSTM and GRU models are evaluated in terms of MSE, RMSE, MAE, R2, and accuracy. The results showed that the GRU model outperformed LSTM models across all selected states in Nigeria's six geopolitical zones, achieving an average accuracy of 98.43%, MSE of 3.0504, RMSE of 1.7381, MAE of 1.1644, and R^2 of 0.7978 for forecasting daily rainfall, respectively. Although the study initially planned to utilize high-resolution data from the Nigerian Meteorological Agency (NiMET), bureaucratic challenges necessitated using ECMWF data instead. Despite this limitation, the findings remain valuable and applicable, providing a baseline for rainfall forecasting in Nigeria. This study sets the path for future studies that attempt to increase and broaden rainfall forecast accuracy in Nigeria. Future work will include utilizing datasets from multiple sources, including NiMET (if available), additional meteorological stations, and global climate models, to improve the model's geographical and temporal coverage. Furthermore, future research aims to employ deep ensemble learning models to improve prediction accuracy, stability, and dependability.

6. REFERENCES

[1] Alzubaidi, F., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M. A., Al-Amidie, M., & Farhan, L. (2021**).** Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions**.** *Journal of Big Data*, *8*(1), 53. https://doi.org/10.1186/s40537-021-00444-8

[2] Baig, F., Ali, L., Faiz, M. A., Chen, H., and Sherif, M. (2024). How accurate are the machine learning models in improving monthly rainfall prediction in hyper-arid environments," Journal of Hydrology, vol. 633, https://doi.org/10.1016/j.jhydrol.2024.13

[3] Chhetri, M., Kumar, S., Roy, P.P., & Kim, B. (2020). Deep BLSTM-GRU Model for Monthly Rainfall Prediction: A Case Study of Simtokha, Bhutan. Remote. Sens., 12, 3174.

[4] Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555*.

[5] Endalie, D., Haile, G., & Taye, W. (2022). The deep learning model for daily rainfall prediction: a case study of Jimma, Ethiopia. Water Supply, 22(3), 3448 doi: 10.2166/ws.2021.391

[6] Fahad, S., Su, F., Khan, S. U., Naeem, M. R., & Wei, K. (2023). Implementing a novel deep learning technique for rainfall forecasting via climatic variables: An approach via hierarchical clustering analysis. Science of The Total Environment, 854:1-9, 158760, http://dx.doi.org/10.1016/j.scitotenv.2022.158760

[7] Ganachari P. V, and Kumari, R., (2021), Machine Learning Based Rainfall Analysis. International Journal of Scientific Research in Science and Technology, 7(4), 208-212, doi: https://doi.org/10.32628/CSEIT217462

[8] Haq, D. Z., Novitasari, D. C. R., Hamid, A., Ulinnuha, N., Arnita, Farida, Y., Nugraheni, R. D., Nariswari, R., Ilham, Rohayani, H., Pramulya, R., Widjayanto, A. (2021). Long Short-Term Memory Algorithm for Rainfall Prediction Based on El-Nino and IOD Data. Procedia Computer Science, 179, 829–837.

[9] He, R., Zhang, L., & Chew, A. W. Z. (2024). Data-driven multi-step prediction and analysis of monthly rainfall using explainable deep learning. Expert Systems with Applications, 235, 121160. https://doi.org/10.1016/j.eswa.2023.121160

[10] Hochreiter, S., and Schmidhuber, J. (1997) "Long short-term memory," Neural Computation, vol. 9(8), pp. 1735–1780.

[11] Khairudin, N.B.M, Mustapha, N.B., Aris, T.N.B.M. and Zolkepli, M.B. (2020). Comparison of Machine Learning Models for Rainfall Forecasting. In 2020 International

Conference on Computer Science and Its Application in Agriculture (ICOSICA) (pp. 1-5), 2020. IEEE.

[12] Mahmood A., and Shrestha A. (2019), "Review of Deep Learning Algorithms and Architectures," in IEEE Access, vol. 7, pp. 53040-53065, 2019, doi: 10.1109/ACCESS.2019.2912200.

[13] Manokij, F., Sarinnapakorn, K., & Vateekul, P. (2019). Forecasting Thailand's Precipitation with Cascading Model of CNN and GRU. 2019 11th International Conference on Information Technology and Electrical Engineering (ICITEE), 1-6

[14] Mungale, N., & Shinde, J. (2024). Rainfall Forecasting: A Comparative Analysis of Deep Learning and Machine Learning Models with Application to Environmental Data. International Journal of Intelligent Systems and Applications in Engineering, 12(12s), 380–393.

[15] Narejo S, Jawaid MM, Talpur S, Baloch R, Pasero EGA (2021). Multi-step rainfall forecasting using a deep learning approach. PeerJ Comput. Sci. 7:e514 DOI 10.7717/peerj-cs.514

[16] Ni, L., Wang, D., Singh, V. P., Wu, J., Wang, Y., Tao, Y., and Zhang, J. (2020). Streamflow and rainfall forecasting by two long short-term memory-based models. Journal of Hydrology, 583, 124296.

[17] Nugraha, Y.E., Ariawan, I., & Arifin, W.A. (2023). Weather forecast from time series data using LSTM algorithm. Jurnal Teknologi Informasi dan Komunikasi, 14(1), 144-152 doi:10.51903/jtikp.v14i1.531

[18] Nurrohman, H.F., Novitasari, D.C., Setiawan, F., Rochimah, Taufiq, A., & Hamid, A. (2022). Rainfall Prediction Using Gated Recurrent Unit Based on DMI and Nino3.4 Index. 2022 IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology (IAICT), 191-196.

[19] Prakaash, A. S. K., Kalaivani, K. C., Rajheshwari, P., Dhanalakshmi, & Mohanaprakash, T. A. (2023). Revolutionizing Rain Prediction: Deep Learning-Powered TensorFlow Solution for Meteorology and Emergency Management. *International Journal of Intelligent Systems and Applications in Engineering*, *14*(4), 844-852.

[20] Samad, A., Jain, P., Bhagyanidhi, S., Gautam, V., and Sarkar, K., (2021). An Approach for Rainfall Prediction Using Long Short-Term Memory Neural Network, IEEE 5th International Conference on Computing Communication and Automation (ICCCA) Galgotias University, Greater Noida, UP, India, pp. 190-195, 2021.

[21] Salehin, I., Talha, I.M., Mehedi Hasan, M., Dip, S.T., Saifuzzaman, M., & Moon, N.N. (2020). An Artificial Intelligence-Based Rainfall Prediction Using LSTM and Neural Network. 2020 IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-ECE), 5-8.

[22] Shiwei C., Adjei, C.O., Tian, W., nzo, B.O., Adu, E., Kedjanyi, G., & Darteh, O.F. (2021). Rainfall Forecasting in Sub-Sahara Africa-Ghana using LSTM Deep Learning Approach, International Journal of Engineering Research & Technology (IJERT), 10(03), 464-470.

[23] Sun, D., Wu, J., Huang, H., Wang, R., Liang, F., & Xinhua, H. (2021). Prediction of Short-Time Rainfall Based on Deep Learning. Mathematical Problems in Engineering, Hindawi, 1-8. https://doi.org/10.1155/2021/6664413

[24] Tuysuzoglu G, Birant K. U., and Birant D.(2023). "Rainfall prediction using an ensemble machine learning model based on K-Stars," Sustainability, 15, 5889, 2023. https://doi.org/10.3390/su15075889

[25] Valavanidis A (2023), "Extreme Weather Events Exacerbated by the Global Impact of Climate Change: A Glimpse of the Future, if Climate Change Continues Unabated", 2023. Retrieved from https://www.researchgate.net/publication/368468689

[26] Yang, S., Yu, X., & Zhou, Y. (2020). LSTM and GRU Neural Network Performance Comparison study: Take the Yelp review dataset as an example. In 2020 International Workshop on Electronic Communication and Artificial Intelligence (IWECAI) (pp. 98-101). IEEE.

[27] Zhang, X., Duan, B., He, S., Wu, X., & Zhao, D. (2022). A new precipitation forecast method based on CEEMD-WTD-GRU. Water Supply (2022) 22 (4): 4120–4132. https://doi.org/10.2166/ws.2022.037