Intelligent Decision Technologies 17 (2023) 1031–1060
DOI 10.3233/IDT-220157
IOS Press

Adaptive membership enhanced fuzzy classifier with modified LSTM for automated rainfall prediction model

Nishant Nilkanth Pachpor[a,*], B. Suresh Kumar[b] and Prakash S. Prasad[c]
[a]*Computer Science Engineering, IIMS, Chinchwad, India*
[b]*Computer Science Engineering, Sanjay Ghodawat University, Atigre Taluka, Maharashtra, India*
[c]*IT Department, Priyadarshini College of Engineering, Nagpur, Maharashtra, India*

Abstract. Nowadays, various research works is explored to predict the rainfall in the different areas. The emerging research is assisted to make effective decision capacities that are involved in the field of agriculture broadly related to the irrigation process and cultivation. Here, the atmospheric and climatic factors such as wind speed, temperature, and humidity get varies from one place to another place. Thus, it makes the system more complex, and it attains higher error rate during computation for providing accurate rainfall prediction results. In this paper, the major intention is to design an advanced Artificial Intelligent (AI) model for rainfall prediction for different areas. The rainfall data from diverse areas are collected initially, and data cleaning is performed. Further, data normalization is done for ensuring the proper organization and related data in each record. Once these pre-processing phases are completed, rainfall recognition is the main step, in which Adaptive Membership Enhanced Fuzzy Classifier (AME-FC) is adopted for classifying the data into low, medium, and high rainfall. Then for each degree of low, medium, and high rainfall, the prediction process is performed individually by training the developed Tri-Long Short-Term Memory (TRI-LSTM). Additionally, the output achieved from the trained TRI-LSTM rainfall prediction in cm for each low, medium, and high rainfall. The meta-heuristic technique with Hybrid Moth-Flame Colliding Bodies Optimization (HMFCBO) enhances the recognition and prediction phases. The experimental outcome shows that the different rainfall prediction databases prove the developed model overwhelms the conventional models, and thus it would be helpful to predict more accurate rainfall.

Keywords: Rainfall prediction, rainfall recognition, adaptive membership enhanced fuzzy classifier, modified long short-term memory, hybrid moth-flame colliding bodies optimization

1. Introduction

Rainfall is defined as the quantity of rain falling within a particular area at a specific time. It is a decisive phenomenon in a climatic system because the turbulent nature was directly influenced the planning of water resources, crop production in agriculture, and the biological environment [1]. Agriculture mainly depends on monthly rainfall and water resources. Prediction of rainfall levels over a particular period is very important for increasing economic growth and financial security in all countries [2]. Rainfall forecast helps farmers in different ways, such as hydrology function, weather data mining, forecasting numerical data, and environmental machine learning conditions [3]. Increasing greenhouse gases and global warming may change the weather and climatic conditions, and also changes some patterns in rainfall. As a result, rainfall prediction is required to avoid floods, increasing crop yield, and reducing financial risk [4]. The

*Corresponding author: Nishant Nilkanth Pachpor, Computer Science Engineering, IIMS, Chinchwad, India. E-mail: nishantnilkanthp@gmail.com.

meteorological department provides information related to the substantial rainfall process and climatic conditions. Rainfall estimation is done by contemplating the volume of rainfall in a specific area, prediction error, and prediction accuracy [5]. Processes involved in the estimation of rainfall in a specific area are a collection of rainwater, analysis of some features, verification using common derivatives, modeling simulation, and manipulation research based on distinct meteorological features and information [6]. Rainfall estimation is carried out using the parameters like average monthly temperature, intensity level, relative humidity, and probability of annual rainfall in millimeters. This estimation helps humans and farmers to provide weather alerts to forbid the occurrence of disaster [7].

In the hydrological cycle, rainfall prediction and weather forecasting are obligatory for governing the agricultural yields, operations of reservoir resources, and scheduling the irrigation resources to improve the financial status [8]. Therefore, various approaches and methodologies like physical-based time series, stochastic and soft computing are developed for forecasting and modeling rainfall with different hydrological parameters in the atmosphere [9]. To provide accurate prediction results, some of the machine learning and deep learning-based methodologies like fuzzy, Artificial Neural Networks (ANN), and Convolutional Neural Networks (CNN) are proposed for managing water resource engineering and weather forecasting. These deep learning-based approaches consider parameters such as drainage, water quality, drought water, groundwater availability, and rainfall estimation [10]. On earth, these parameters vary from one place to another place and one time to another. Monthly analysis of climatic conditions includes the beginning level, intermediate level, and final level of the rainy season, which alleviates giving accurate information about rainfall in the distribution of intra-year while considering seasonal rainfall [11]. The agricultural area might be mostly perturbed using monthly rainfall, and so, the exact estimation of rainfall plays an important role to improve the quality of decisions among climatic conditions in particular places [12].

Various approaches are used for obtaining the required information based on irrigation and storage of water needed for farming which takes a huge process to provide support for farmers to increase economic growth [13]. Multilayer Perceptron (MLP) and Support Vector Machine (SVM) are the commonly used classifier for the estimation of rainfall [14]. Naïve-based methods like Artificial Neural Networks (ANN) and Arti?cial Intelligence (AI) are valuable methods for forecasting rainfall in particular areas [15]. Deep learning techniques such models are Hidden Markov Model (HMM), Deep CNN (DCNN), Restricted Boltzmann Machine (RBM), convolution and pooling, Long Short Term Memory (LSTM), Recurrent neural network (RNN), Conditional RBM and Autoencoder are used to provide the time-series information [16]. Non-linear data are handled very efficiently in rainfall prediction using these deep learning approaches, and it provides emerging and promising rainfall prediction results based on low-level rainfall and high-level rainfall and also forecasting rainfall with meteorological parameters [17]. Thus, this research work aims to develop a new deep learning-based rainfall prediction model to give accurate results during rainfall.

The main contributions related to the knowledge of this research work are summarized below:

- To develop an adaptive deep learning-based rainfall prediction system with improved fuzzy-based rainfall recognition by utilizing the optimization strategy for the precise prediction of seasonal and annual rainfall at different states at different times.
- To develop an enhanced fuzzy logic for the recognition of rainfall to give the accurate level of rainfall, whether it is low, medium, and high by optimizing the membership function using the developed HMFCBO.
- To introduce a TRI-LSTM-based rainfall prediction model with the training of low, medium, and high rainfall prediction separately and obtain the rainfall output in cm. Separate training of low, medium, and high rainfall prediction in TRI-LSTM is performed for reducing the computational complexity and error rate. To implement HMFCBO to optimize parameters in the proposed rainfall estimation model. In fuzzy, the membership function is optimized for improving the recognition performance, and in LSTM, epochs are optimized for increasing the prediction performance with a low mean square error rate.

– A comparison of the performance of the proposed HMFCBO-TRI-LSTM over different classifiers and various algorithms is performed to validate the effectiveness of the system.

The upcoming sections to elaborate on the developed rainfall prediction model are given as follows: Section 2 summarizes the previously used rainfall prediction model and also their limitations. Section 3 elucidates the developed algorithm description. Section 4 gives the preprocessing techniques, the rainfall recognition process using fuzzy, and the rainfall prediction using LSTM. Section 5 describes the results and performance analysis of the modified rainfall prediction model. Section 6 shows the conclusion of the proposed rainfall prediction model.

2. Literature survey

2.1. Related works

In 2020, Nam and wang [18] developed a new rainfall prediction model based on a deep learning approach, which combined a stacked autoencoder and sparse autoencoder to give information about landslide susceptibility. The data has been applied for the feature extraction process and, further, performed with classification in three different layers like input, hidden, and output layers. This model was performed with real landslide and non-landslide information and compared the performance that has been executed with the SVM classifier and random forest classifier. This model has given higher performance when considering optimal non-linear features. In 2021, Venkatesh et al. [19] promoted generative adversarial networks for estimating the annual and monthly rainfall data in India and also performed the prognosis of upcoming rainfall. The developed rainfall estimation approach has considered the generator and discriminator that were performed in LSTM and CNN. The time-series information related to rainfall was precisely estimated using LSTM. The effectiveness of the suggested generative adversarial networks has attained more efficient prediction results, which has assisted the farmers to cultivate their crops and also raised the economic growth of the country.

In 2020, Zhang et al. [20] implemented an altitude-combined rainfall forecasting model for estimating short-term monthly rainfall with promising accuracy. Firstly, the weighted k-means grouping model has been used for collecting meteorological information from the target center with various environments. The high altitude shear value of the specified target station was computed for all dimensions behind surface factors, and it was decreased by using Principal Component Analysis (PCA) method. The effectiveness of the model has been established by the terms Mean Square Error (MSE) and threat rating. In 2018, Xiang et al. [21] proposed an ensemble Empirical Mode Decomposition (EMD) rainfall detection model while considering long and short time series information from rainfall areas. The features from the rainfall were extracted information from the EMD model, and the efficiency was computed with some performance metrics like precision, sensitivity, time efficiency, and accuracy.

In 2020, S and P [22] have suggested a new rainfall prediction system by adopting the deep learning network called the convolutional LSTM approach for accurate estimation along with the spatial and temporal patterns in the rainfall. This model has adopted a stochastic gradient descent method and Salp Swarm Algorithm (SSA). To optimize the weights in LSTM, this has been implemented mainly for global optimal tuning of the weights. The big data has been efficiently handled in this model and provided better performance regarding RMSE and MSE rates. In 2018, Zhang et al. [23] developed a novel rainfall forecasting model based on non-monsoon and annual seasons employed with MLP and support vector regression for the estimation of annual rainfall in both seasons. Parameters should be considered for the calculations that were cloud cover, humidity, wind velocity, and average monthly temperature. Different metrics were used for comparing the efficiency of the developed rainfall estimation model mean average error, coefficient of efficiency, correlation and MSE.

Table 1
Benefits and drawbacks of previously developed rainfall estimation models

Author [citation]	Technique	Benefits	Drawbacks
Nam and wang [18]	Stacked autoencoder (StAE) and a sparse autoencoder (SpAE)	– It extracts the features efficiently. – It also improves classification performance.	– The prediction performance can be affected by fewer hidden neurons in the spare encoder.
Venkatesh et al. [19]	LSTM	– It improves prediction accuracy and performance.	– This model suffers from computational complexity.
Zhang et al. [20]	Weighted k-means clustering	– It attains a high accuracy rate. – It also offers more efficient training.	– However, this model is only applicable for limited time intervals, and it does not have an optimal structure.
Xiang et al. [21]	SVR-ANN	– It offers more reliability and performance. – It also reduces the error rate in terms of MAE and RSME.	– It gets the lower discrete degree. – It is a time-consuming process.
S and P [22]	convLSTM	– It attains higher prediction performance and attains high accuracy.	– The ability of rainfall prediction has to be enhanced, and this model is not suitable for other countries.
Zhang et al. [23]	SVR-MLP	– It attains superior generalization capability and attains higher accuracy.	– This model does not suitable for LSTM and multi-variate time series prediction.
Guhathakurta et al. [24]	NN	– This model attains higher performance even for smaller spatial scales. – It also reduces the RMSE.	– Though, less performance is attained while using a more homogeneous rainfall time series.
Thai et al. [25]	PSOANFIS	– It increases the prediction performance for daily rainfall. – This model attains lower MAE and higher R-mean values.	– This model does not suitable for 24 h scale models.

In 2020, Guhathakurta et al. [24] employed the NN technique to forecast monsoon rainfall data. This model has been employed at different meteorological stations from monthly rainfall time series data. The developed system has attained suitable input and output non-linear relationships and also accurately predicted the seasonal rainfall at a certain duration. The developed area-weighted rainfall forecasting of entire sub-divisions was used for generating the rainfall forecasting with all India monsoon details. In 2020, Thai et al. [25] employed AI based Fuzzy clustering system optimized with Particle Swarm Optimization (PSO) for the estimation of monthly rainfall and daily rainfall. This model has collected the meteorological parameters from the environment such as solar radiation and maximum temperature. Various measures were validated to show the higher performance of the developed model.

2.2. Problem statement

Rainfall prediction is a complex task, and thus, the outcomes must be accurate. However, short-term rainfall prediction can be explored exactly, whereas long-term prediction is a challenging task. Moreover, different hardware devices are used for rainfall prediction in real time that, includes weather factors such as pressure, humidity, and temperature. On the other hand, these existing approaches cannot perform efficiently, and thus, there is a need to adopt deep learning approaches for rainfall prediction. However, the size of historical data affects the prediction performance and also requires more storage, less accuracy rate, and computation time. Numerous approaches have been proposed in recent years, which have several benefits and limitations, as given in Table 1. StAE and SpAE [18] extract the features efficiently and also improve the classification performance. Conversely, the prediction performance can be affected by fewer hidden neurons in the spare encoder. LSTM [19] improves the prediction accuracy and performance, but this model suffers from computational complexity. Weighted k-means clustering [20] attains a high accuracy rate and also offers more efficient training. However, this model is only applicable for limited time intervals, and it does not have optimal structure. SVR-ANN [21] offers more reliability and performance and also reduces the error rate in terms of MAE and RSME. Conversely, it gets the lower discrete degree. ConvLSTM [22] attains higher prediction performance and attains high accuracy. The ability of rainfall prediction has to be enhanced, and this model is not suitable for other countries. SVR-MLP [23] attains superior generalization capability, and higher accuracy, and this model does not suitable for LSTM and multi-variate time series prediction. NN [24] attains more efficient performance even for smaller spatial scales and also reduces the RMSE. Though, less performance is attained while using a more homogeneous rainfall time series. PSOANFIS [25] increases the prediction performance for daily rainfall and attains lower MAE and higher R-mean values. Though, this model does not suitable for 24 h scale models.

2.2.1. Research gaps and challenges

Rainfall is a critical phenomenon in the weather and climatic system. It is mainly influenced by water resources, ecosystems, management of irrigation, and agriculture. All people in the world broadly depend on water for their entire life. Rainfall prediction in a specified area is helpful for farmers to cultivate their crops without any water deficiency resources. Being aware of annual rainfall estimation and analysis is useful for human beings to find solutions during surplus rainfall and deficit rainfall. Estimation of rainfall grabbed the attention of governments, industries, scientific communities, and all the risk management entities. Rainfall occurs randomly in all areas hence, it is very difficult to predict [35–37] the accurate rainfall level based on the weather and climatic conditions. Hence, the accurate prediction of rainfall becomes a crucial issue for gaining information from atmospheric conditions. Therefore, traditional forecasting methodologies are implemented for the prediction of rainfall through the analysis of weather and climatic conditions. To improve the effectiveness of the prediction, numerous methodologies like data mining, AI, statistics, deep learning, and machine learning have been developed. Real-time estimation of rainfall is a challenging task because it is a stochastic and non-linear process. Some dynamic models are introduced by physically generating equations based on the atmospheric conditions for forecasting rainfall. Yet, it does not provide effective results over rainfall estimation using mathematical expressions. Hence, empirical approaches are suggested for estimating rainfall by using regression methods, and deep learning approaches [34]. Neural networks with feed-forward neural networks and backpropagation networks are developed to predict rainfall. But, it does not handle the time series-based information very efficiently. Fuzzy logic is used for the prediction of rainfall in many years and performs the prediction not only in numerous scales, but also labeled the variables in stochastic environmental conditions. The analysis of the fuzzy-based rainfall prediction model is highly effective, but it relies on more real-time problems among several factors. MLP is the most widespread technique in the estimation of rainfall based on atmospheric factors like wind speed, temperature, relative humidity, and wind pressure. The drawbacks of the MLP method are insufficient memory capacity and the occurrence of gradient problems. The gradient problems are vanished by using the LSTM approach for the accurate prediction of the rainfall along with optimization algorithms. But, the results acquired a high error rate using this model. Hence, this proposed method focuses mainly on eliminating the problems that occur in the previous models, and some predictive analyses are performed for validating the performance of the newly suggested rainfall prediction model.

2.2.2. Innovations

Various existing issues are considered for implementing a new rainfall prediction model. The new rainfall prediction model is developed using the HMFCBO algorithm and TRI-LSTM deep-structured architectures. These types of weather forecasting are fascinating concepts, and also it is utilized for accurate prediction. It is utilized for the investigation of seasonal variability, and to forecast yearly/monthly rainfall over some geographical areas. The developments of this implemented model are useful for agricultural applications. The impacts of this rainfall prediction help to avoid the destruction of crops and farms and damage to property. This early warning is rescued for the human risks to life and property and helps to manage agricultural farms. The experimental results for this offered method proved its effectiveness, and it is utilized to predict the faults accurately.

3. Architectural illustration of rainfall prediction system with heuristic strategy

3.1. Proposed rainfall prediction model

The schematic representation of the offered rainfall prediction model is shown in Fig. 1.

A newly enhanced deep learning framework is offered for the prediction of rainfall using atmospheric factors like temperature, wind speed, wind pressure, and relative humidity. This technique is mainly used for the accurate prediction of rainfall and reduces the RMSE and MSE rates. This is broadly helpful for the farmers using improving agricultural yields and also helpful for people following the prevention of drought and natural disasters. The dataset contains the weather details for all states, but we collect the data regarding different states like Arunachal Pradesh, Assam, Manipur, Tripura, Mizoram, West Bengal, Sikkim, and the Himalayas. The data collected from the above-mentioned states are initially given to the pre-processing technique. This pre-processing method used in rainfall prediction is mainly for eliminating unwanted data and reducing noise present in the original information. It consists of two stages likely, data cleaning and data normalization. In data cleaning, the missing values in the original information are reduced to get high-quality data without loss of information. In normalization, the noises and distortions present in the data are removed. After pre-processing, the incorrect data are eliminated, and useful information regarding rainfall prediction is obtained. The preprocessed output is given to the rainfall recognition process, where AME-FC is used to improve the recognition performance in terms of accuracy. Parameter optimization is preferred in this developed AME-FC, where membership functions are optimized using developed HMFCBO and predicted the rainfall to be low, medium, and high levels. The recognized low, medium, and high rainfall outcomes are given to the rainfall prediction process, where the developed TRI-LSTM is used for the accurate prediction of rainfall. Here, the training of low, medium, and high rainfall is performed. This gives results about the rainfall in cm, and the individual prediction process decreases the error rate and computational complexity. By using developed HMFCBO in LSTM, the epochs in the classifier are tuned to improve the prediction performance regarding accuracy and precision. Finally, MSE and RMSE rates are calculated to validate the performance by comparing the proposed model with the existing classifiers and heuristic algorithms.

3.2. Proposed heuristic algorithm

The developed hybrid deep learning-based rainfall prediction model uses the HMFCBO optimization algorithm in the rainfall recognition phase and the rainfall prediction phase. In the recognition phase, the membership function is optimized in fuzzy, and in the rainfall prediction phase, the epochs are tuned in the LSTM classifier. The MFO algorithm [26] is generally used in the rainfall prediction model because of solving real-world problems in constrained unknown search space and provides solutions for far distances. However, it suffers from problems like less population diversity, low premature convergence, local optimization of entrapment, and poor efficiency in exploration and exploitation. These challenges in the conventional MFO motivate us to integrate MFO with CBO, and it is termed as HMFCBO. Because of these drawbacks, the CBO heuristic algorithm is preferred for the accurate prediction of rainfall.

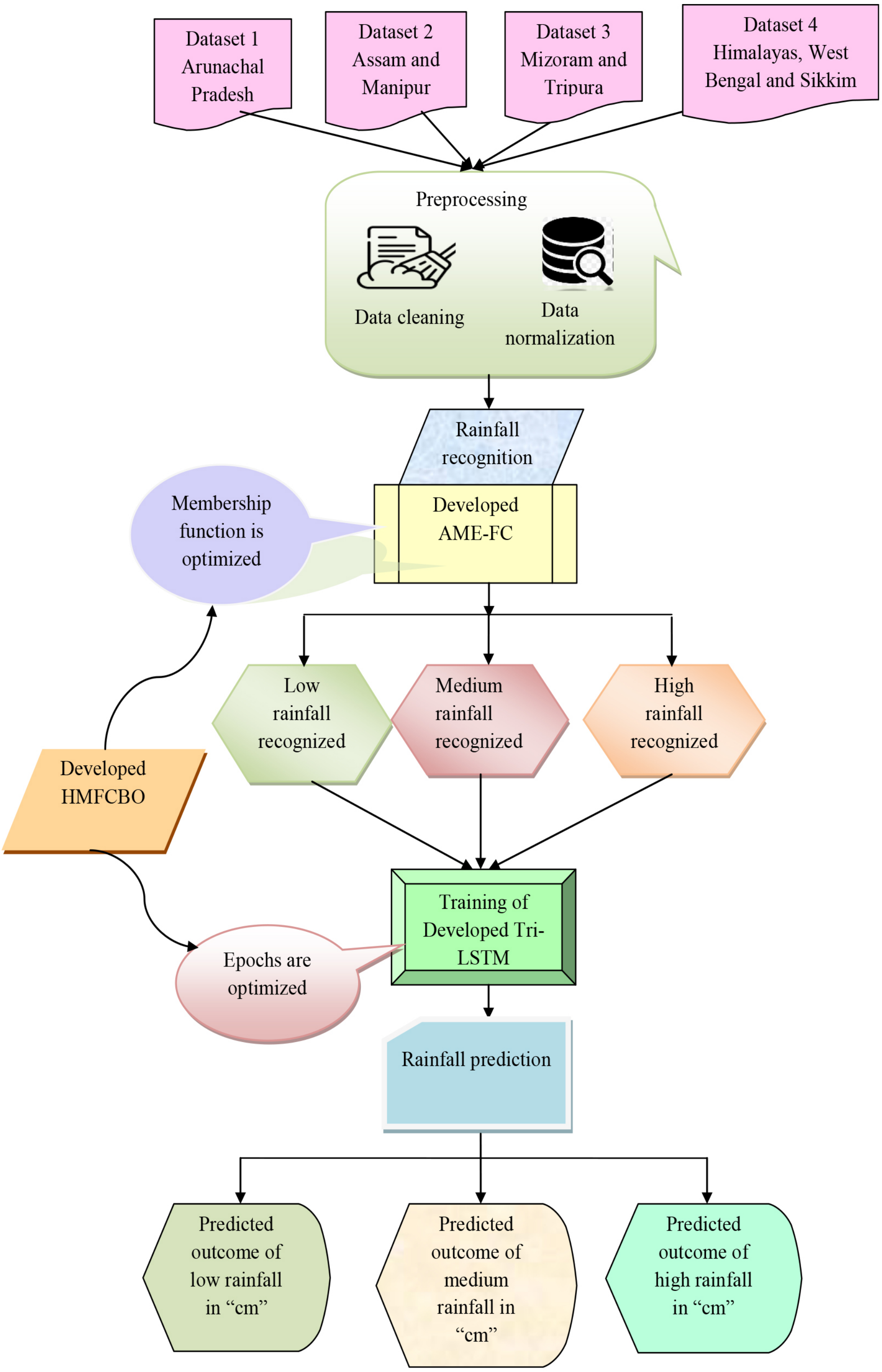


Fig. 1. Schematic representation of the rainfall estimation model using a modified deep learning technique.

CBO algorithm [27] requires less memory space for colliding between two bodies, and it generates the right balance between local and global search space. In this proposed HMFCBO, the parameter t is introduced to update the solution based on the variable a, and it is computed by below Eqs (1), (2) respectively.

$$a = -1 + i * \left(\frac{(-1)}{\max - it}\right) \tag{1}$$

$$t = abs\left(\frac{(a-1)}{it}\right) \tag{2}$$

The variable it is defined as the number of iterations. If $(t < 0.5)$ the condition is satisfied, then upgrade the solution using the CBO algorithm, otherwise, it is updated by the MFO algorithm.

MFO: It is nature-inspired algorithm [36] broadly depends on the moth's capacity to migrate in the dark that is predicted by using transverse propagation of the moth. It can fly straight lines over long distances. In the MFO algorithm, all moths are considered in the position of n-dimensional vector space.

In the MFO algorithm, all moths are the search agent, and all flames are in the best position of all moths. During the finding process, flames are generally relinquished by the moth, and it is updated based on the position. The updated position of all moths is calculated using Eq. (3).

$$c_{m,n} = D\left(MH_m, FL_n\right) \tag{3}$$

Here, the term c_m represents m^{th} the moth, FL_n defines the n^{th} flame, and D is the spiral function of the MFO algorithm. The logarithmic representation of the spiral function is calculated using Eq. (4).

$$D\left(c_m, FL_n\right) = S_m e^{xy} \cos\left(2\pi y\right) + FL_n \tag{4}$$

From Eq. (2), the term x is the constant value, and it describes the shape of the logarithmic spiral function, y is laid in the interval of $[-1, 1]$, and S_m is the distance calculated between m^{th} moth and n^{th} flame that is evaluated by Eq. (5).

$$S_m = |FL_n - c_m| \tag{5}$$

To improve the exploitation process in the MFO algorithm, the number of flames used in the entire process is decreased. It is mathematically defined in Eq. (6).

$$flame_{number} = rand\left(T - R * \frac{T-1}{Q}\right) \tag{6}$$

Moreover, the variable T denotes the number of current iterations, the term Q represents a total number of maximum iterations and R the total number of flames. The moths in the algorithm never lost the best flames during all the iterations. It chooses the best solution among all the flames.

CBO: In CBO, colliding bodies are split mainly into two divisions that are the moving group and the stable group. In the moving group, the object is propagated along with the stable group. The objects from these groups are colliding with each other and immigrate into a new place after the collision. The steps involved in the process of colliding are summarized below:

Step 1: Stimulate the position of all colliding bodies.
The position c is first initialized, selected $c_{\min}$ and $c_{\max}$ values randomly. The parameter c is initialized based on Eq. (7).

$$c_l^0 = c_{\min} + rand\left(c_{\max} - c_{\min}\right), l = 1, 2, \ldots, k \tag{7}$$

Step 2: Mass value M_x of all colliding bodies is computed by Eq. (8).

$$M_x = \frac{\frac{1}{obfn(x)}}{\sum_{l=}^{k} \frac{1}{obfn(l)}}, \quad x = 1, 2, \ldots, k \tag{8}$$

Here, the term $obfn(k)$ is the objective function of the k^{th} colliding body and k define the total number of populations.

Step 3: The colliding body values are arranged in ascending order based on the smaller objective function and larger objective function. The stationery group contains the smaller objective function colliding bodies, and the moving group consists of the larger objective function colliding bodies. Initially, the speed of the stationery group colliding bodies is set to 0 before the collision process. The moving group speed is calculated by Eq. (9) and Eq. (10), respectively.

$$u_l = 0,\ l = 1, \ldots, \frac{k}{2} \tag{9}$$

$$u_l = c_{l-\frac{k}{2}} - c_l,\ l = \frac{k}{2} + 1, \ldots, k \tag{10}$$

Here, the term c_l is the speed vector of the colliding body and u_l denotes the vector position of the colliding body.

Step 4: After performing a collision between two bodies, the speed of the stationary group and the moving group can be changed. The moving group speed is evaluated using Eq. (11).

$$u_l' = \frac{\left(M_l - \beta M_{l-\frac{k}{2}}\right) u_i}{M_l + M_{l-\frac{k}{2}}},\ l = \frac{k}{2} + 1, \ldots, k \tag{11}$$

Stationary group colliding body speed is calculated by using Eq. (12).

$$u_l' = \frac{\left(M_l - \beta M_{l-\frac{k}{2}}\right) u_i}{M_l + M_{l-\frac{k}{2}}},\ l = \frac{k}{2} + 1, \ldots, k \tag{12}$$

Here, the variable β is denoted as the coefficient of restitution, and it is calculated using Eq. (13).

$$\beta = 1 - \frac{it}{it_{\max}} \tag{13}$$

Here, the term *it* is the current iteration number and $it_{\max}$ denotes the maximum number of iterations involved in the process.

Step 5: Based on the speed of the moving object, the position of each group is calculated using Eq. (14).

$$c_l^{new} = c_{l-\frac{k}{2}} + rand \times u_i',\ l = \frac{k}{2} + 1, \ldots, k \tag{14}$$

Algorithm 1: Developed HMFCBO

Initiate the population in the search space
Update the parameter
  For all solutions
    Evaluate the value of a Eq. (1)
    Find the value of the parameter t by Eq. (2)
    If $(t < 0.5)$
      Upgrade the position of the solution by Eq. (15)
    Else
      Update the position of the solution by Eq. (3)
    End if
  End for
  Validate the parameters
Obtain the optimal solution

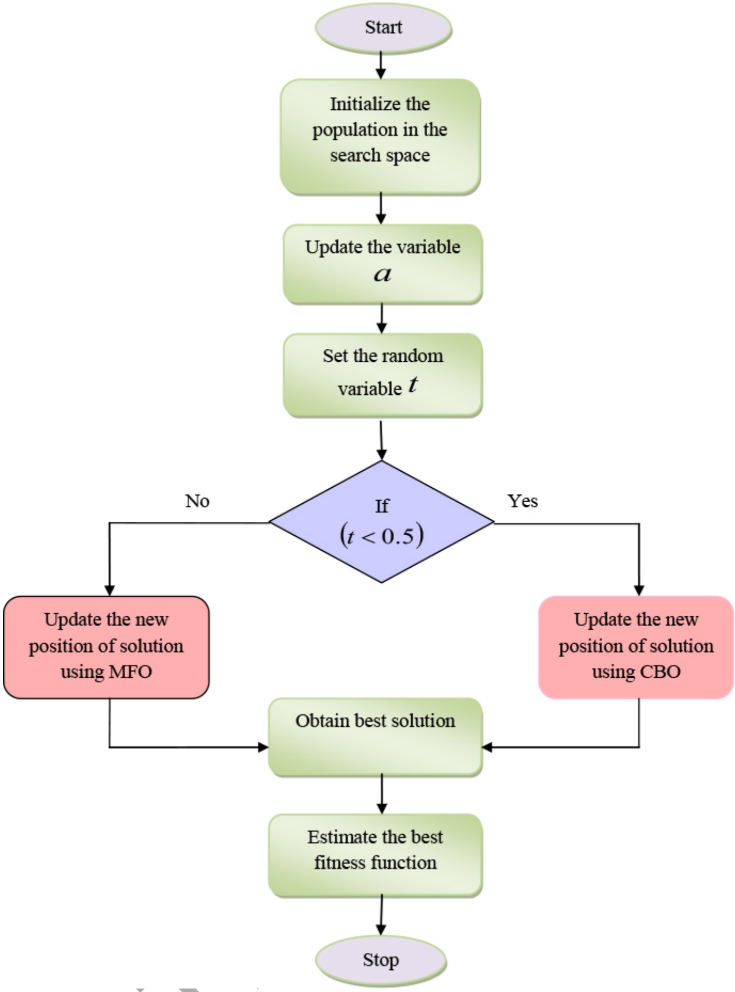


Fig. 2. Flowchart of a proposed HMFCBO.

However, the new position of the static group is expressed in Eq. (15).

$$c_l^{new} = c_{l-\frac{k}{2}} + rand \times u_i'', \; l = \frac{k}{2} + 1, \ldots, k \tag{15}$$

The random variable is distributed in the interval $[-1, 1]$ after the collision process.

Step 6: Whether the number of the current iteration is the same as the number of maximum iteration then the colliding process is terminated.

The flowchart of the developed HMFCBO is showcased in Fig. 2.

4. Processing steps for automated rainfall prediction system

4.1. Dataset description

The Subdivision wise Rainfall and its departure from 1901 to 2015 dataset are taken from the link.[1] The daily rainfall is measured by mm per day. This dataset includes the details like annual rainfall and seasonal

[1] https://data.gov.in/catalog/rainfall-india?page=1: access date is 2022-07-08.

rainfall in a particular area. The parameters used to monitor the climatic condition are wind pressure, temperature, relative humidity, wind direction, dew point, and wind pressure. The data has been gathered during the monsoon season. As per the requirement, the information has been separated, and the needed parameters are changed into categorical data.

The data dataset contains the weather information regarding all states, but, in the proposed model, these data are split into four datasets. However, the below mentioned datasets contain the time-series dataset.

Dataset 1: In dataset 1, the weather details are taken from Arunachal Pradesh state. The details are collected from the state of Arunachal Pradesh in India from year 1916 to 2017 with the months of January-December. The variable count is taken as 13×116. The datasets hold float values for providing rainfall prediction.

Dataset 2: In dataset 2, includes the weather details taken from Assam and Manipur states. From 1916 to 2017, the rainfall prediction is considered concerning the months of January-December. The 13×116 count variables are taken. The dataset includes float values for providing rainfall prediction.

Dataset 3: In dataset 3, it contains information regarding the states of Mizoram and Tripura states. In the years 1916–2017, the rainfall prediction is taken from the year 1916 to 2017 with the months of January-December. The count of variables is considered as 13×116. It takes float values for visualizing the rainfall prediction.

Dataset 4: In dataset 4, the information regarding rainfall conditions is taken from sub-Himalayas, West Bengal, and Sikkim states. Here, the prediction of rainfall is from the year 1916 to 2017, with the months of January-December. The similar count variable is considered in this dataset. It includes float values for revealing the rainfall prediction.

Here, the main thing is the given dataset splits into two phases. One is the training phase, and another one is the testing phase. 75% of the data are utilized for the training phase, and also 25% of the data are utilized for the testing phase. In this research work, the validation of this prediction is divided into 5 sets. For example, the author is considered for 100 data for processing. Moreover, the 1$^{\text{st}}$ set includes 1–20 data. The 2$^{\text{nd}}$ set holds 21–40 data. Accordingly, the 3$^{\text{rd}}$ set includes 41–60 data. In the same manner, the 4$^{\text{th}}$ set holds 61–80 data. Finally, the 5$^{\text{th}}$ set includes 81–100 data. At first, we have taken the 1$^{\text{st}}$ set, where testing is done on 1 set, and then the other sets of data have been used for the training phase. Based on this process, the implementation has been done. This process is repeated until accurate solutions have been obtained. The input data to be gathered from the dataset are denoted by Da_x^{in}, where x is distributed in the interval $x = 1, 2, ..., M$. Here, the term M denotes the total number of rainfall-based data.

4.2. Data pre-processing

The input Da_x^{in} is applied to the pre-processing technique to eliminate unwanted information from the gathered dataset. After pre-processing, high-quality data is obtained using the data cleaning and normalization methods.

4.2.1. Data cleaning

The input data Da_x^{in} is given for the data cleaning process. The cleaning process mainly deals with missing information by using the averaging method. This performs the cleaning process by taking the sum of all the instances of the specified parameter divided by the total number of samples taken in the process. The main sources of all the missing information were broken, then the information has never been stored, and there will be a break in the data transmission. The missing values in the dataset are removed by using the cleaning process. The output to be obtained from the data cleaning process is $Da_x^{c\ln d}$, and this output is applied to the data normalization process.

4.2.2. Data normalization

The normalization process takes the input data as $Da_x^{c\ln d}$. This process gives the parameter values in the specified range. Noises present in the information are removed by using the data normalization method within the particular interval. The input parameters are effectively mapped into the specified time interval for facilitating the prediction process. It is mainly a scaling technique to maintain the huge variations in the pre-processing stage. The output gets from the data normalization technique Da_x^{norm}, and this output is given for the further processing of the rainfall recognition phase.

4.3. Developed objective of rainfall recognition

Rainfall recognition is carried out by developed AME-FC to raise the outcome of the recognition process in terms of accuracy. The output Da_x^{norm} obtained from the normalization process is given to the developed AME-FC for the recognition of rainfall.

4.3.1. Fuzzy

Fuzzy system [28] is broadly used in the rainfall prediction approach because it provides higher performance results in decision-making and classification, where the solution is achieved by combining the input parameters with the output parameters. The steps involved in the adaptive fuzzy system are fuzzification, defuzzification, and rule estimation. In fuzzification, numeric values are changed into linguistic values. The membership function is adapted for solving complex situations and produces a high potential outcome. Some rules are available for all the combinations of input and output variable membership functions. The rules used for the prediction of rainfall are given below.

Rule 1: If s_1 is D_1^y, s_2 is D_2^y, s_i is D_i^y, then h is K_t.

In rule1, y is taken as $y = 1, 2, \ldots, Y$ that gives the number of rules, $t = 1, 2, \ldots, T$ describes the class number, i gives the feature number, and the feature variable is represented by D_n^y, where i is considered in the interval of $i = 1, 2, \ldots, j$. The membership functions are used to partition the linguistic ranges between low, medium, and high. The membership function is defined by $\chi_{D_n^y}$ and is mathematically represented in Eq. (16).

$$\chi_{D_n^y}(s_i) = \begin{cases} e^{\frac{1}{2}\left(\frac{s_i - t_i^y}{\rho_{i,n}^A}\right)^2}, s_i < t_i^y \\ e^{\frac{1}{2}\left(\frac{s_i - t_i^y}{\rho_{i,n}^B}\right)^2}, s_i > t_i^y \end{cases} \tag{16}$$

The membership function center is represented by the term t_i^y , and the right spread and left spread membership functions are expressed by the terms $\rho_{i,n,}^A$, and $\rho_{i,n,}^A$ respectively. The parameters involved in the data recognition are described in vector space that is denoted by $\hat{s}_n = [s_1, s_2, \ldots, s_l]$.

Rule 2: $\chi^y(\hat{s}_n) = \chi_{D_1^y}(s_1)\,\chi_{D_2^y}(s_2)\ldots\chi_{D_n^y}(s_n)$

The firing strength of the feature vector is calculated for recognizing rainfall. Firing strength is calculated using rule 2. The error rate of the fuzzy system is calculated for analyzing the accuracy of the prediction model. The error rate is given in Eq. (17).

$$\varepsilon_{rate} = \begin{cases} 0\,\textit{for correct classification of}\, s_n \\ 1\,\textit{for incorrect classification of}\, s_n \end{cases} \tag{17}$$

The output achieved from the fuzzy system is labeled as a low, medium, and high membership values for the accurate recognition of the rainfall. The membership function in developed AME-FC is optimized using developed HMFCBO to generate accurate results based on the low rainfall, medium rainfall and high

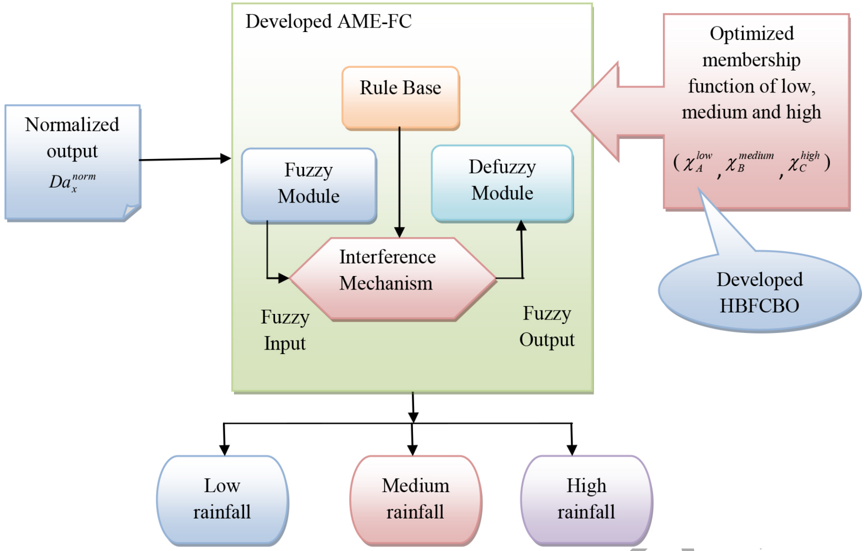


Fig. 3. Rainfall recognition model using developed AME-FC.

rainfall ranges at particular areas in specified time. The objective function of the developed AMEFC is defined F_1, and it is calculated using Eq. (18).

$$F_1 = \underset{\{\chi_A^{low}, \chi_B^{medium}, \chi_C^{high}\}}{\arg\min} \left(\frac{1}{accuracy}\right) \tag{18}$$

Here, the term χ_A^{low} denotes the optimized membership function at a low range and it is laid in the interval between [0, 3] and the term χ_B^{medium} represents the optimized membership function at a medium range and it is considered in the interval between [0, 3], and the term χ_C^{high} is the optimized membership function at high range in the interval between [0, 3]. The accuracy is explained as the true values of the predicted model to all observations. Here, P_T is the true positive, N_T is the true negative, P_F is the false positive, and N_F is the false negative value of the observations. Accuracy is examined using Eq. (19).

$$Accuracy = \frac{(P_T + N_T)}{(P_T + N_T + P_F + N_F)} \tag{19}$$

This recognized output is given for the rainfall prediction process. The rainfall recognition process with developed AME-FC is illustrated in Fig. 3.

5. Deep learning-based automated rainfall prediction system with objective function

5.1. Basic LSTM

LSTM [29] is used to predict rainfall in cm for the specified areas. Accurate estimation of the memory block is important for rainfall calculation. LSTM mainly consists of three layers the input layer, the output layer, and the forget layer. Layers in the LSTM are activated by the values 0 or 1. In this, 0 indicates block everything, and 1 indicates allow everything to pass through the system. Therefore, the gradient problem being backpropagated vanishes but remains constant for solving this gradient problem [38]. This network is employed to estimate background radiation from time-based atmospheric information. The input to be stored in the memory state is given in Eq. (20).

$$A_s = \alpha\left(Q_A \cdot \lfloor m_{y-1}, n_y \rfloor + w_A\right) \tag{20}$$

Here, the term m_{y-1} is the hidden state memory cell, and it is the formation of the input layer, which gives information related to the newly added information in the input layer of the LSTM cell. Forget layer input is represented by the term B_s, it is the working function of the forgotten layer, and it gives the details about the removed information from the LSTM cell. The hidden state information is kept based on below Eq. (21).

$$B_s = \alpha\left(Q_B \cdot \lfloor m_{y-1}, n_y \rfloor + w_B\right) \tag{21}$$

The input, forget and hidden state information are activated by below Eq. (22).

$$\tilde{C}_s = \tanh\left(Q_C \cdot [m_{y-1}, n_y] + w_C\right) \tag{22}$$

The current memory cell state is denoted by C_s, and the $\tanh$ function or sigmoid function is used to combine the data as in Eq. (23).

$$C_s = B_s * C_{s-1} + A_s * \tilde{C}_s \tag{23}$$

The information to be stored in the output state is given in Eq. (24), and it is denoted by the term D_s finally, the $\tanh$ function is used to combine the values in the output state between $[-1, 1]$ those that are represented in Eq. (25).

$$D_s = \alpha\left(Q_D \cdot \lfloor m_{y-1}, n_y \rfloor + w_D\right) \tag{24}$$

$$E_s = D_s * \tanh C_s \tag{25}$$

LSTM preserves the information in the forward direction and removes the information in the backward direction. Sliding windows are used for moving the information in the backward direction and the forward direction.

5.2. Modified LSTM

LSTM networks are broadly used rainfall prediction models with high performance. It is mainly used for solving the gradient problems that occur in the prediction model, and also it solves the long-term dependencies based on atmospheric and climatic factors. The complex autocorrelation sequence is highly performed using this LSTM architecture. Yet, it needs more training for real-world applications, and it requires more resources for the accurate prediction of rainfall. To obtain rainfall prediction results very quickly, and overcome these drawbacks in the LSTM model, the newly developed TRI-LSTM is demonstrated for the prediction of rainfall. Training of low, high and medium rainfall is processed individually in the TRI-LSTM. The predicted rainfall is measured in cm. This separate training of the prediction process improves the performance while reducing the computational complexity and error rate. Here, parameter optimization, like epochs in the LSTM classifier, is tuned using the proposed HMFCBO. The objective function of TRI-LSTM is represented by the term F_2, and it is calculated by Eq. (26).

$$F_2 = \underset{\{Ep_a^D\}}{\arg\min}\left(MASE + MAE\right) \tag{26}$$

Here, the term Ep_a^D denotes the optimized epochs in the LSTM classifier. The tuned epochs are lies in the range of $[10, 50]$. MAE is defined as the ratio of the difference among the sum of the magnitude of corresponding model output and observation to the total number of observations, and it is given in Eq. (27).

$$MAE = \frac{100\%}{l} \sum_{k=1}^{l} \frac{(|AVG - FC|)}{AVG} \tag{27}$$

The variable *FC* describes the forecasted value and *AVG* represents the average value.

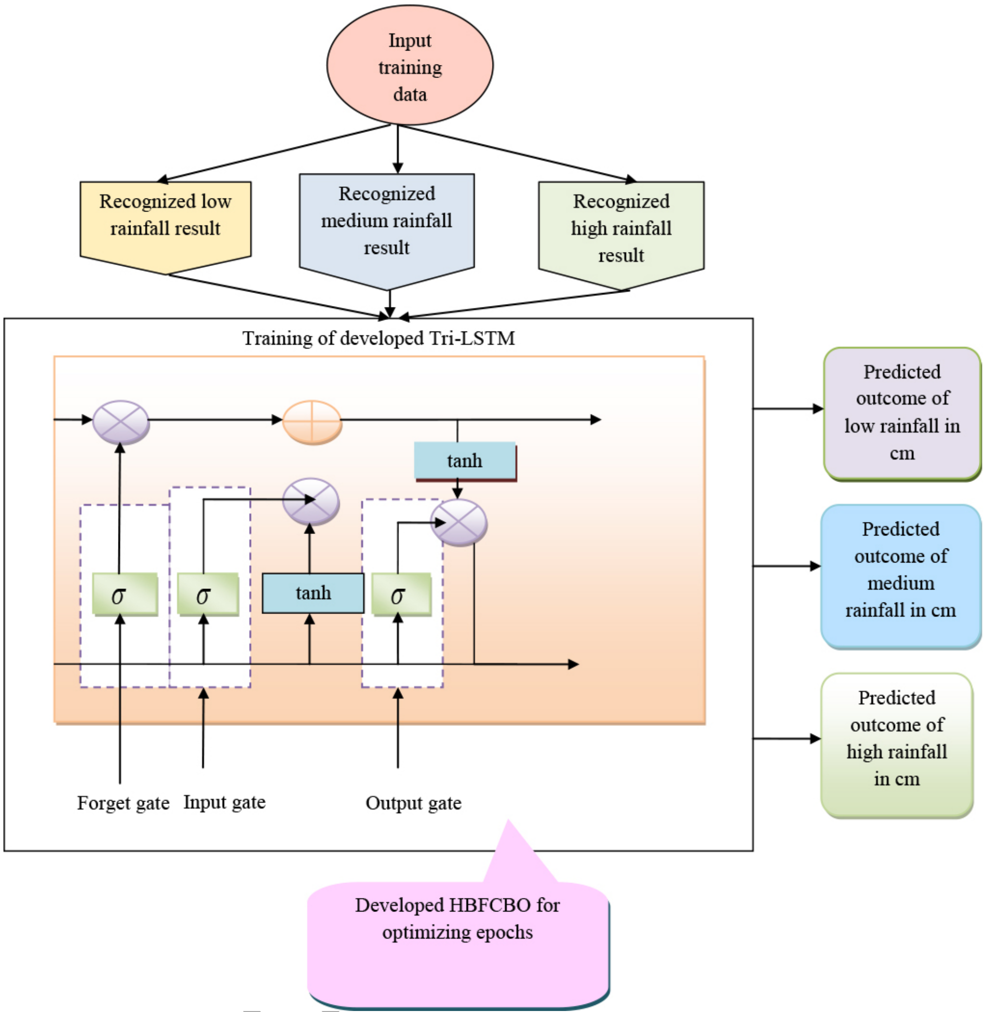


Fig. 4. Developed TRI-LSTM-based rainfall prediction model.

The Mean Absolute Scaled Error (MASE) is determined as the prediction error by the average of the prediction error, and it is computed by Eq. (28).

$$MASE = \frac{|FC|}{\frac{1}{l-1}\sum_{k=1}^{l}|AVG_k - AVG_{k-1}|} \tag{28}$$

Finally, the rainfall-predicted output is achieved with a low error rate using the proposed TRI-LSTM in the developed rainfall prediction model. The schematic representation of the suggested TRI-LSTM-based rainfall prediction model is depicted in Fig. 4.

6. Results

6.1. Simulation setting

The offered HMFCBO-TRI-LSTM rainfall prediction models were implemented in MATLAB 2020a, and the performance of the proposed model was compared with the previously used prediction techniques and optimization algorithms. The population size taken in the optimization process was 10, and the maximum number of iterations taken for the process was 10. The performance analysis used L2-Norm, L1-Norm, LINF- Norm, MAE, MASE, RMSE, SMAPE, and MEP. The outcome of the model was compared with the heuristic algorithms like "Dragonfly Algorithm (DA) [30], Electric Fish Optimization (EFO) [31], Moth Flame Optimization (MFO) [27], and Colliding Bodies Optimization (CBO) [26] and performance analysis over different prediction techniques like fuzzy [28], RNN [32], NN [33] and LSTM [29]".

6.2. Evaluation measures

The evaluation measures needed for the computation of the rainfall prediction model are summarized below:

(a) RMSE: It defines the sum of the difference among the output and observations and it is examined by Eq. (28).

$$RMSE = \sqrt{\frac{\sum_{k=1}^{l} (|AVG - FC|)^2}{l}} \tag{29}$$

(b) SMAPE: It is calculated based on the value of actual and forecast that is given in Eq. (30).

$$SMAPE = \frac{100\%}{l} \sum_{k=1}^{l} \frac{(|FC - AVG|)}{\left(\frac{|AVG| + |FC|}{2}\right)} \tag{30}$$

(c) MEP: It is computed by the average percentage error that differs by the actual value is expressed by Eq. (31).

$$MEP = \frac{100\%}{l} \sum_{k=1}^{l} \frac{AVG - FC}{AVG} \tag{31}$$

(d) L1-Norm: The magnitude sum of all the vectors in the search space is denoted in Eq. (31).

$$L1 - Norm = \sum_{k} |M_k| \tag{32}$$

Here, the term M_k is the matrix, where $k = 1, 2, \ldots, m$ and m denotes the matrix size.

(e) L2-Norm: It is defined as the shortest distance between two points, and it is evaluated using Eq. (33).

$$L2 - Norm = \left(\sum_{k=1}^{l} M_k^2\right)^{\frac{1}{2}} \tag{33}$$

(f) L-Infinity-Norm: The length is evaluated by using the maximum norm value and which is given in Eq. (34).

$$L - INF - Norm = \max |M_k| \tag{34}$$

6.3. Performance evaluation on dataset 1

The extensive experiment is conducted for evaluating the performance of the proposed HMFCBO-TRI-LSTM rainfall prediction model over different optimization algorithms and different classifiers for dataset 1 that are graphically presented in Figs 5 and 6, respectively. Performance analysis over heuristic algorithms shows existing DA has high MAE, MASE, SMAPE, and RMSE, but the proposed rainfall prediction model secures low RMSE, SMAPE, MEP, and MAE. The proposed HMFCBO-TRI-LSTM-based rainfall prediction model achieves 23.72%, 26.22%, 37.06%, and 46.74% improved MAE performance than the other heuristic algorithms. In graph analysis, the DA-TriLSTM algorithm attains a second higher

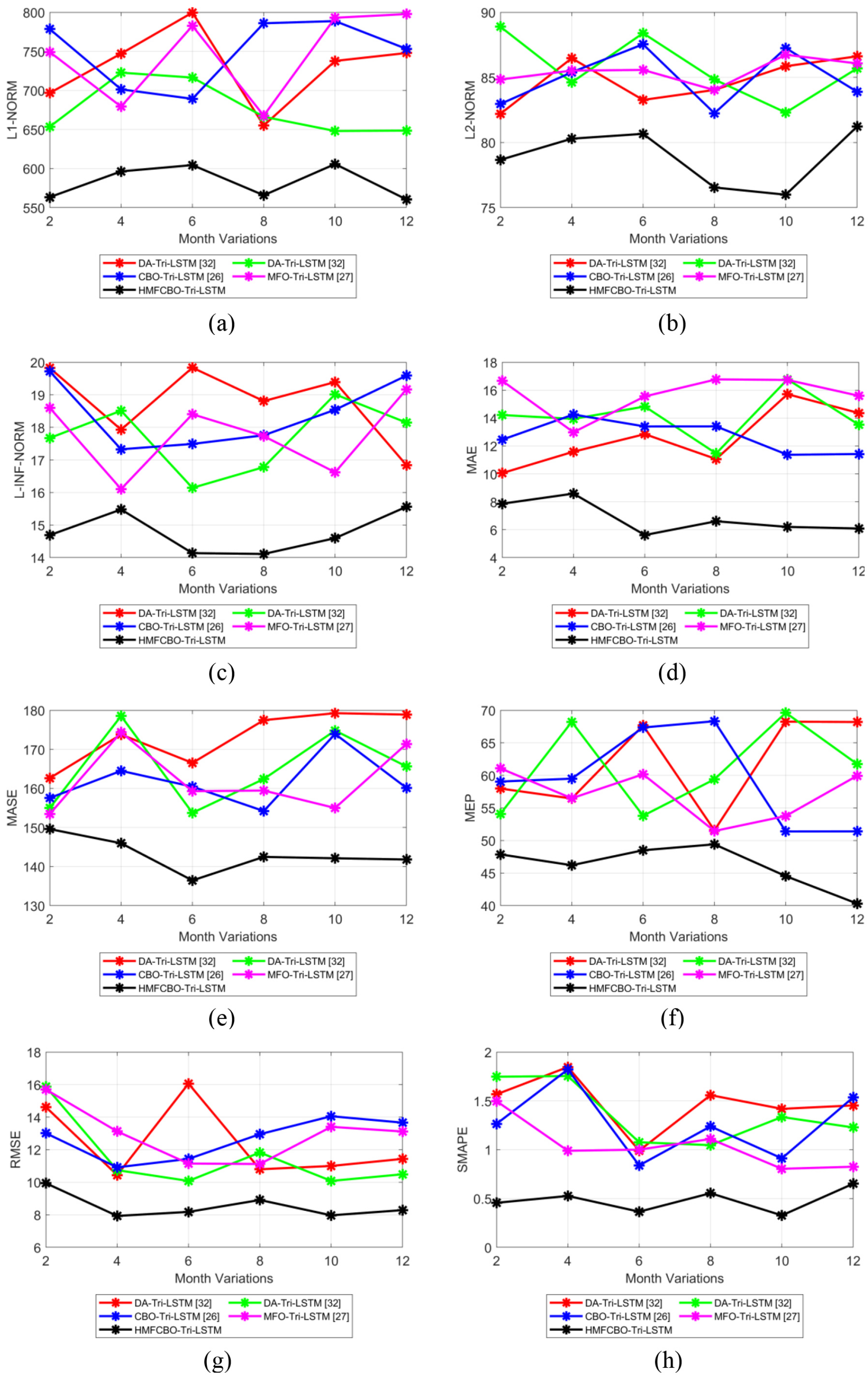


Fig. 5. Performance evaluation on HMFCBO-TRI-LSTM using dataset 1 over different heuristic algorithms with regards to "(a) L1-Norm, (b) L2-Norm, (c) L-INF-Norm, (d) MAE, (e) MASE, (f) MEP, (g) RMSE, (h) SMAPE".

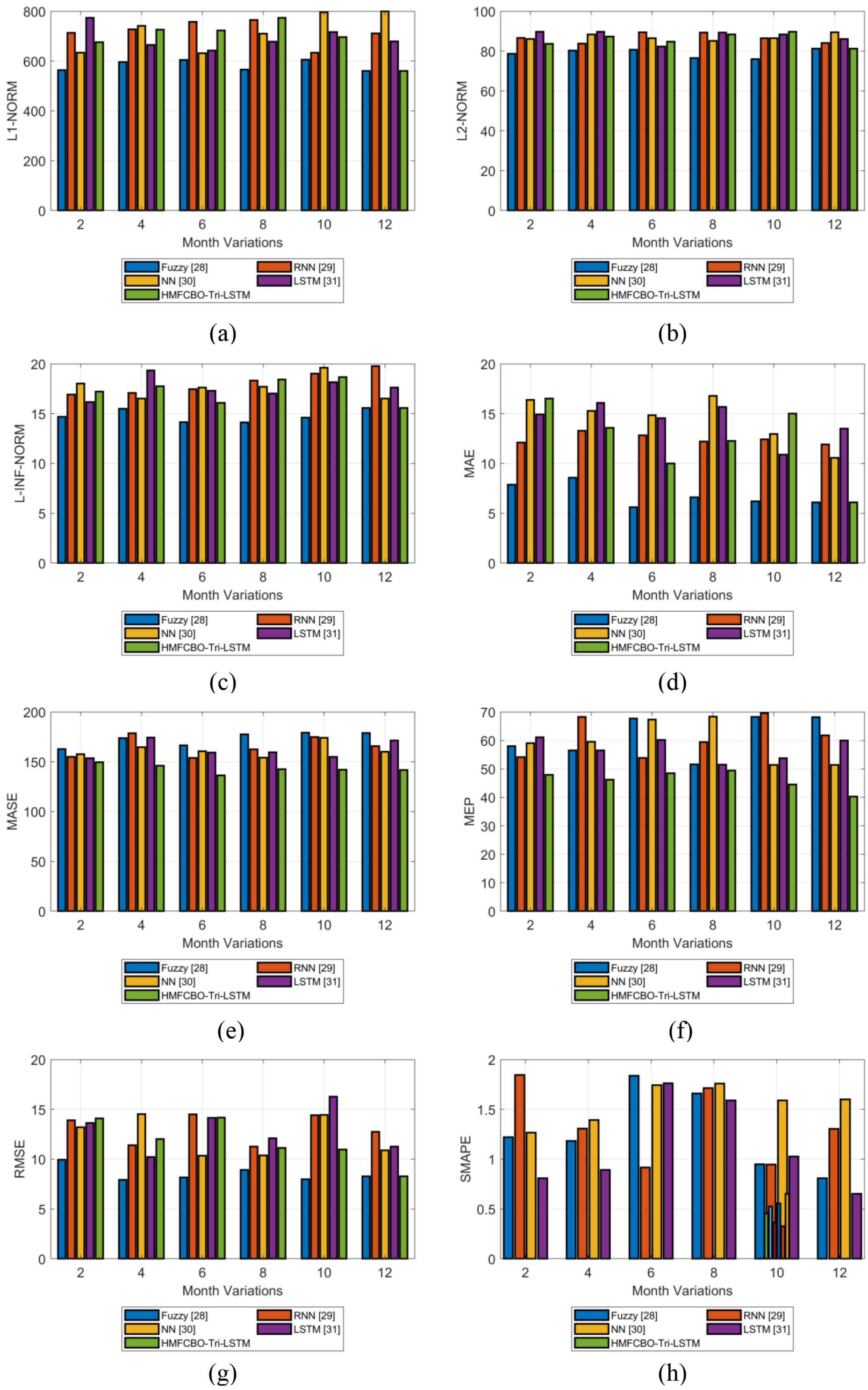


Fig. 6. Performance evaluation on HMFCBO-TRI-LSTM using dataset 1 over different prediction techniques with regards to "(a) L1-Norm, (b) L2-Norm, (c) L-INF-Norm, (d) MAE, (e) MASE, (f) MEP, (g) RMSE, (h) SMAPE".

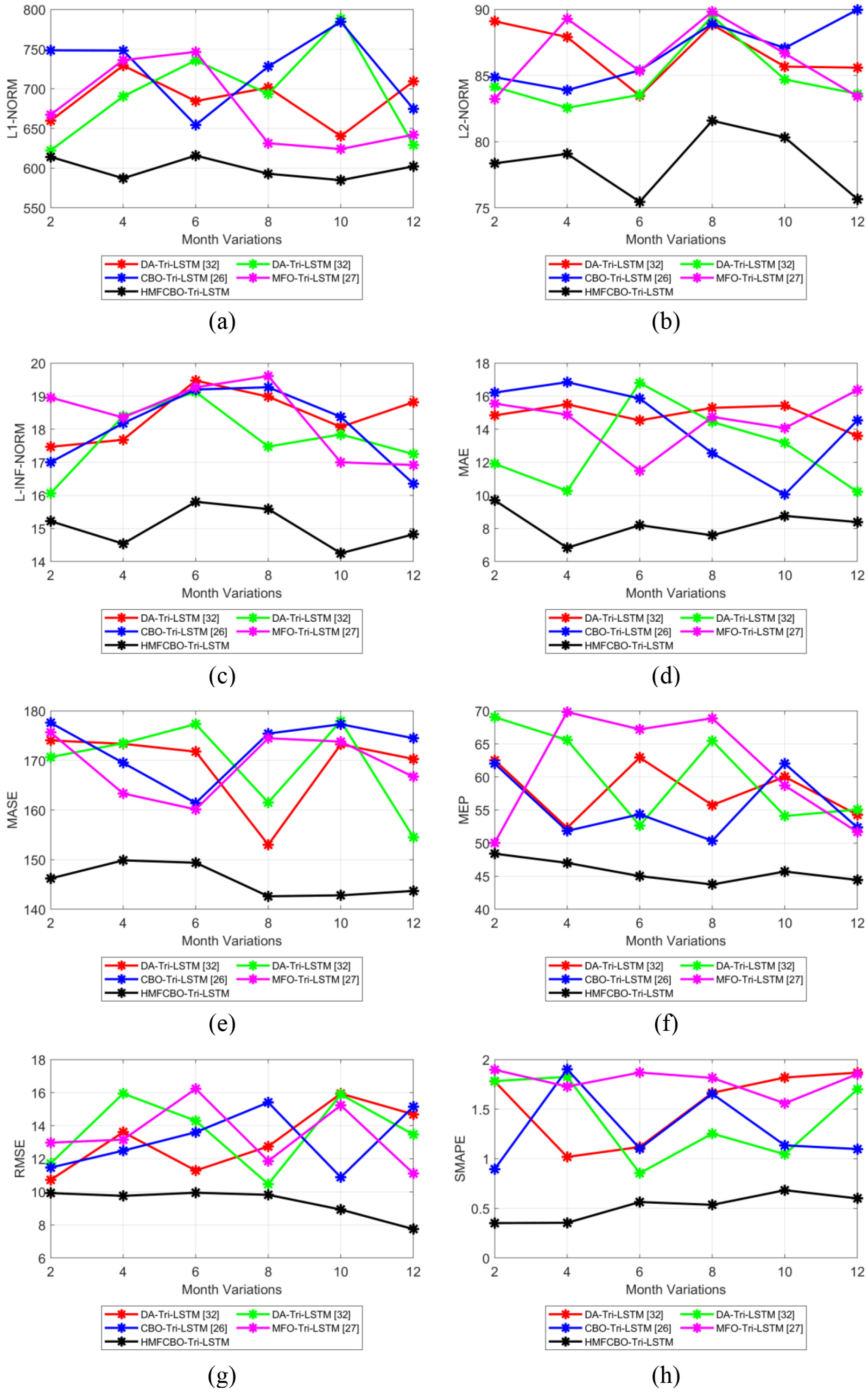


Fig. 7. Performance evaluation on HMFCBO-TRI-LSTM using dataset 2 over different heuristic algorithms with regards to "(a) L1-Norm, (b) L2-Norm, (c) L-INF-Norm, (d) MAE, (e) MASE, (f) MEP, (g) RMSE, (h) SMAPE".

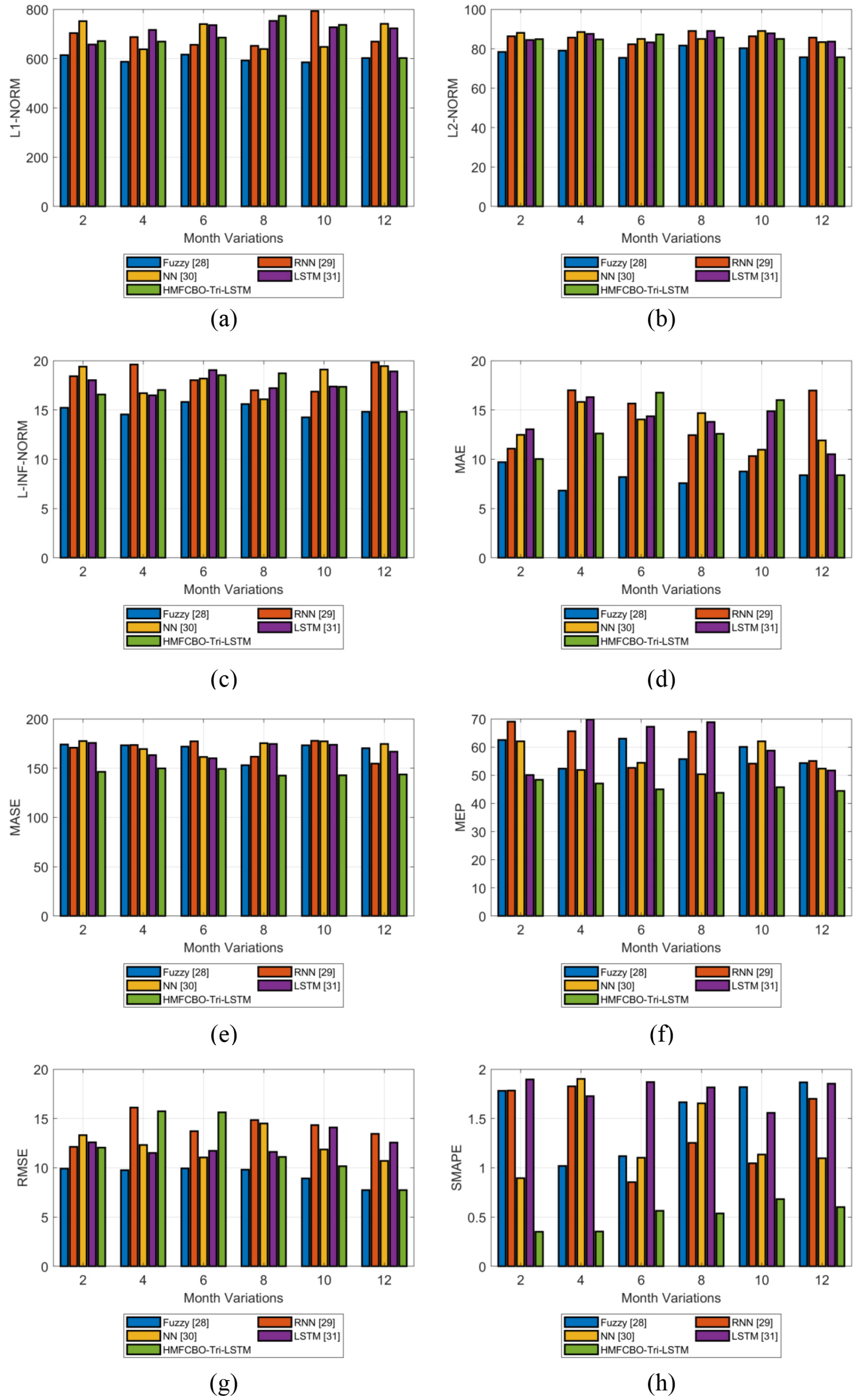


Fig. 8. Performance evaluation on HMFCBO-TRI-LSTM using dataset 2 over different prediction techniques with regards to “(a) L1-Norm, (b) L2-Norm, (c) L-INF-Norm, (d) MAE, (e) MASE, (f) MEP, (g) RMSE, (h) SMAPE”.

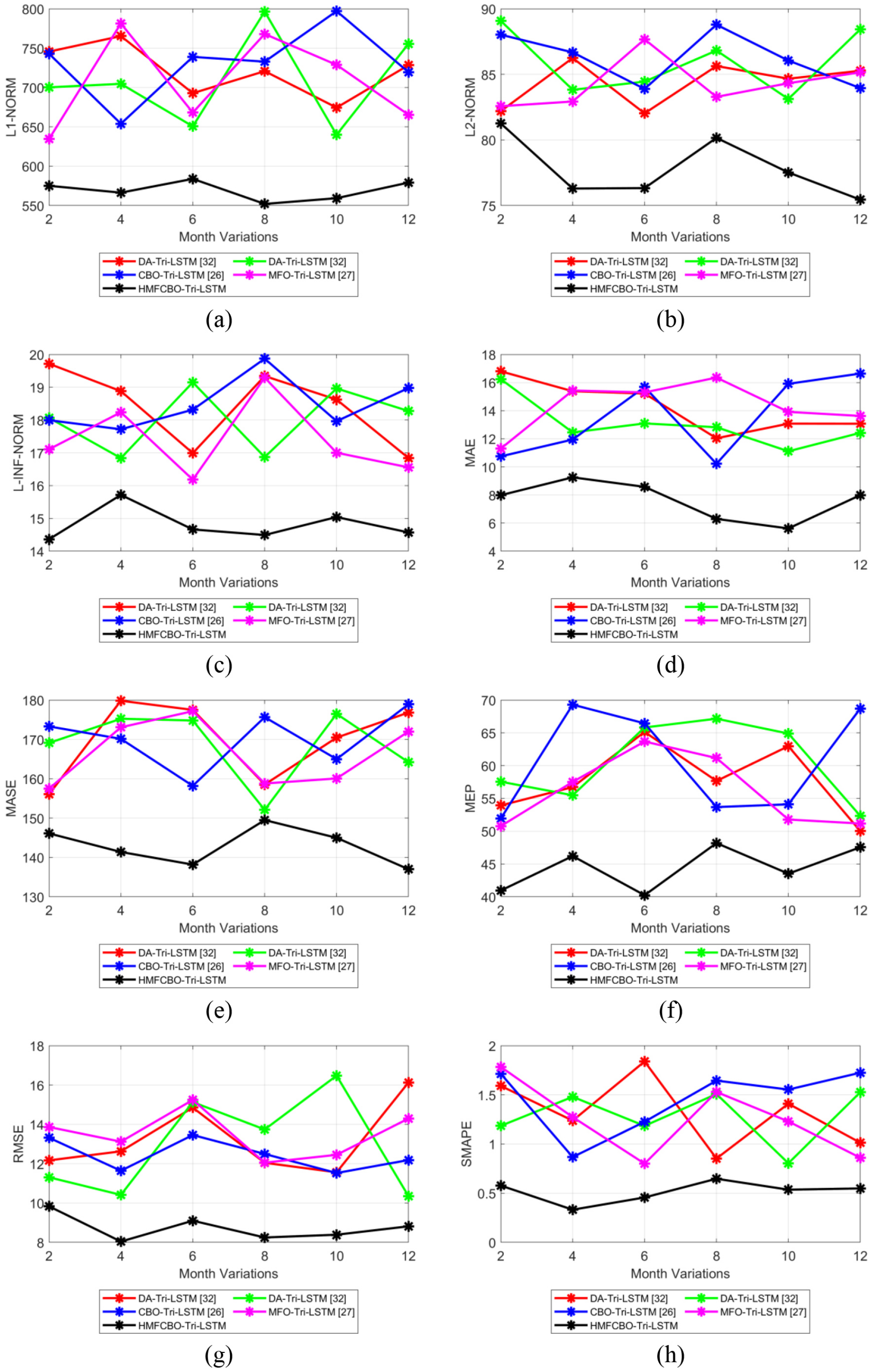


Fig. 9. Performance evaluation on HMFCBO-TRI-LSTM using dataset 3 over different heuristic algorithms with regards to "(a) L1-Norm, (b) L2-Norm, (c) L-INF-Norm, (d) MAE, (e) MASE, (f) MEP, (g) RMSE, (h) SMAPE".

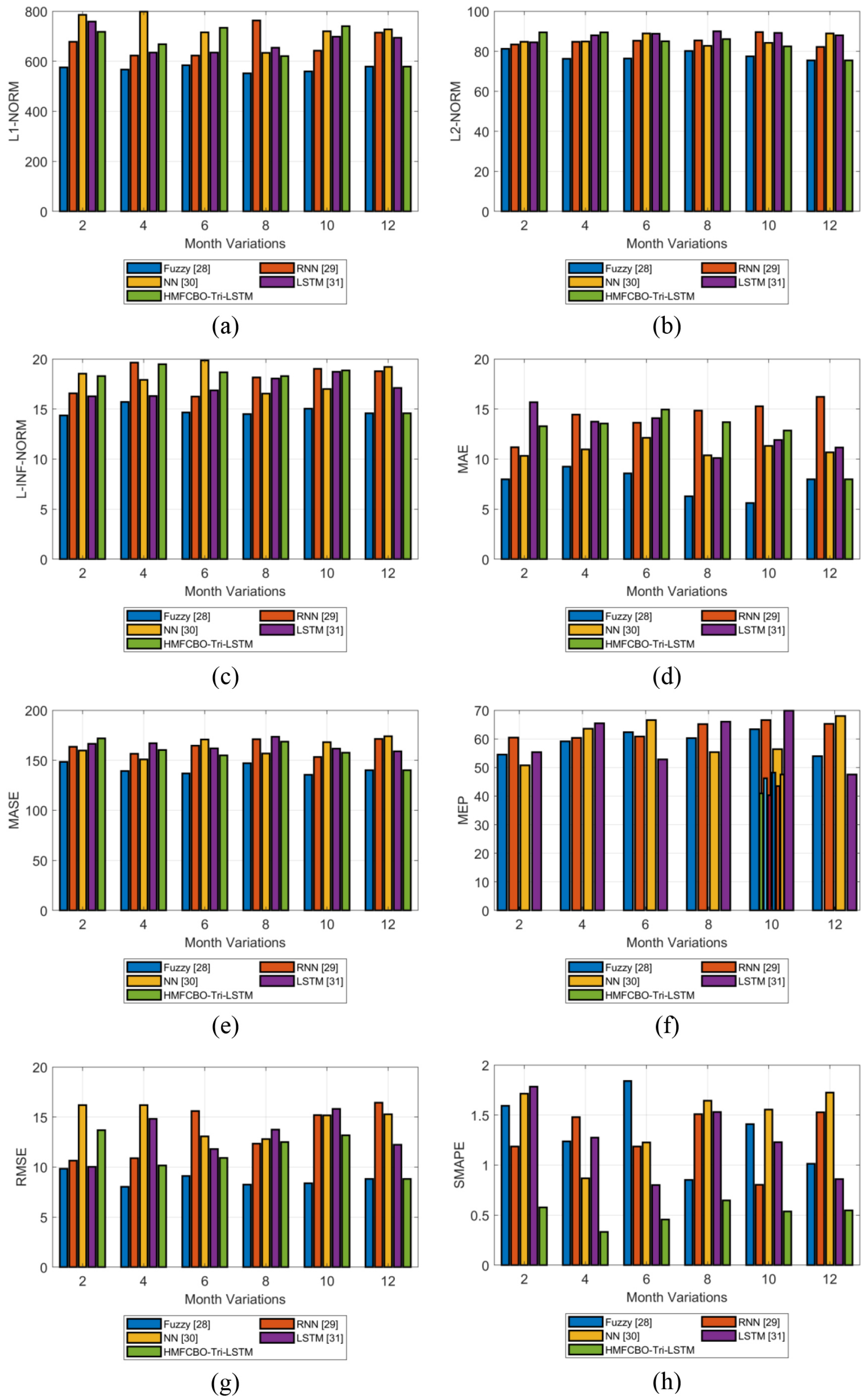
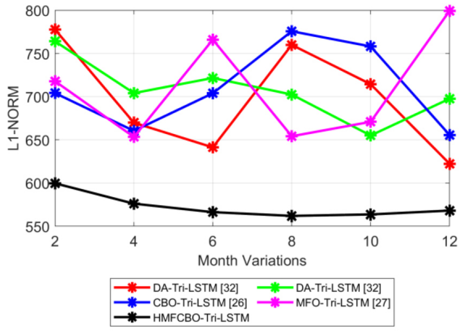
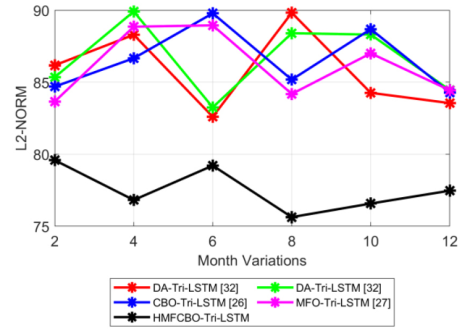


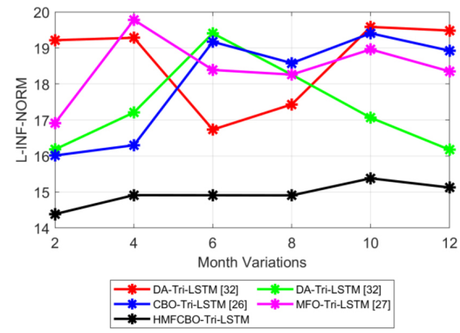
Fig. 10. Performance evaluation on HMFCBO-TRI-LSTM using dataset 3 over different prediction techniques with regards to "(a) L1-Norm, (b) L2-Norm, (c) L-INF-Norm, (d) MAE, (e) MASE, (f) MEP, (g) RMSE, (h) SMAPE".



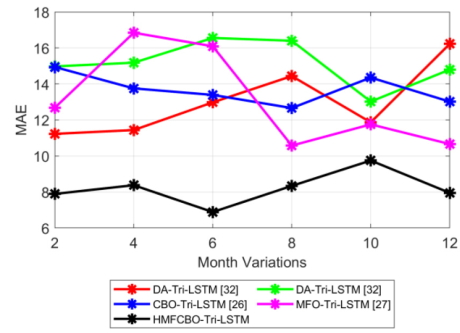
(a)



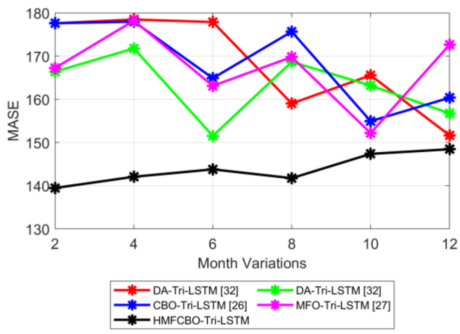
(b)



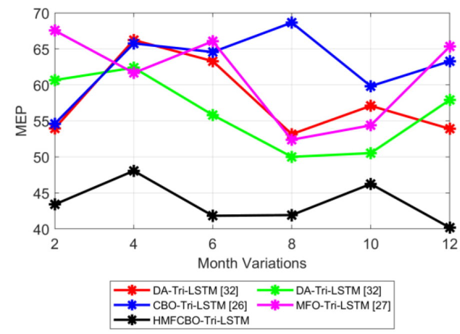
(c)



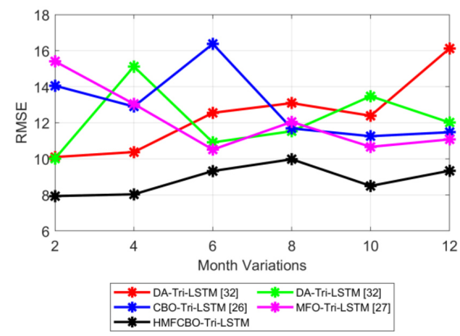
(d)



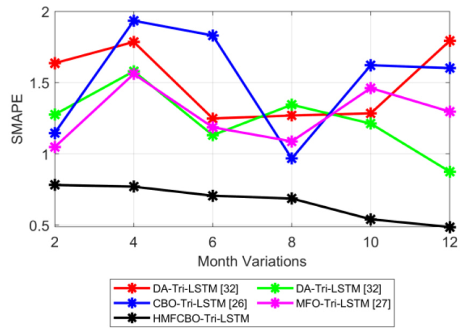
(e)



(f)



(g)



(h)

Fig. 11. Performance evaluation on HMFCBO-TRI-LSTM using dataset 4 over different heuristic algorithms with regards to "(a) L1-Norm, (b) L2-Norm, (c) L-INF-Norm, (d) MAE, (e) MASE, (f) MEP, (g) RMSE, (h) SMAPE".

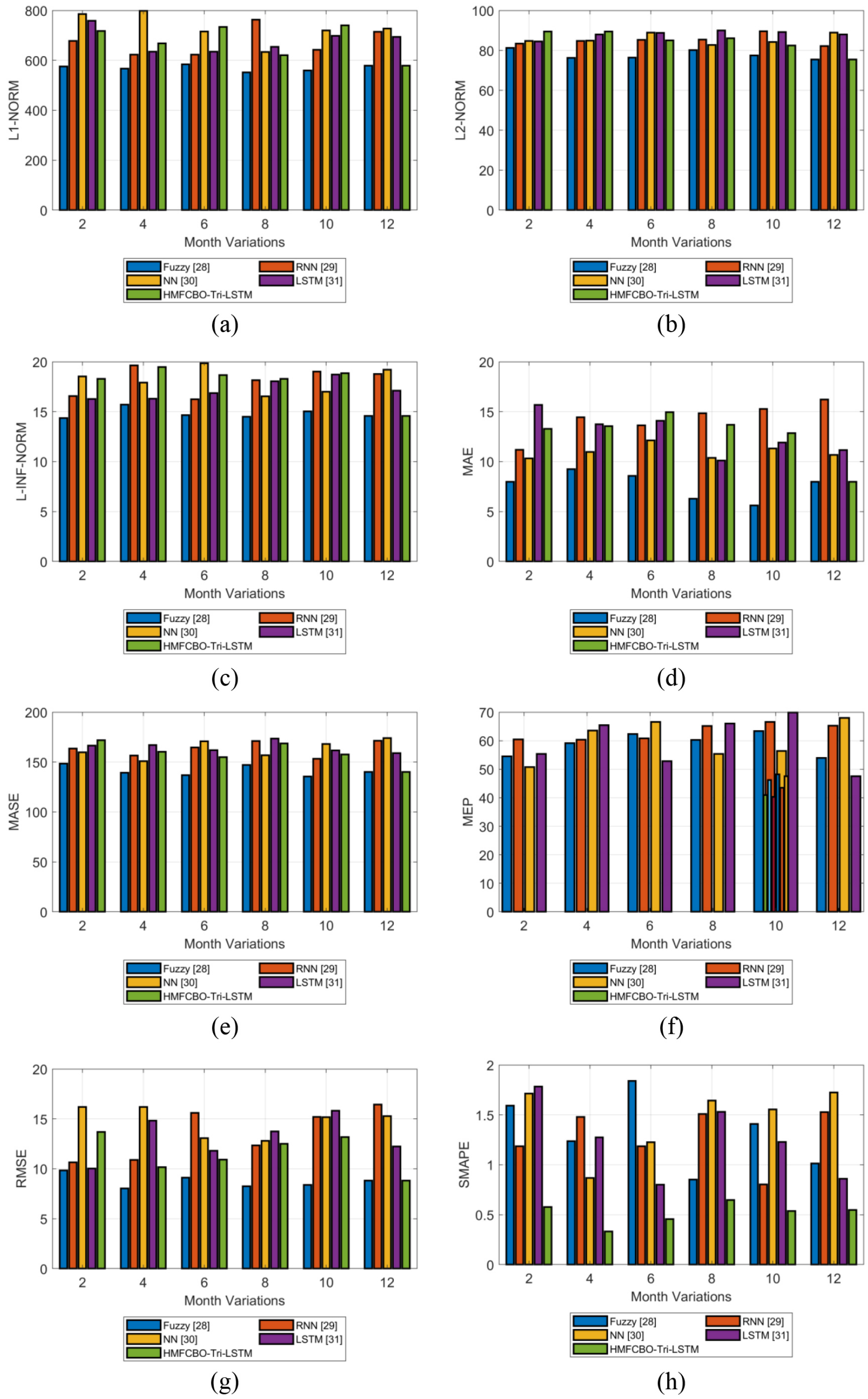


Fig. 12. Performance evaluation on HMFCBO-TRI-LSTM using dataset 4 over different prediction techniques with regards to “(a) L1-Norm, (b) L2-Norm, (c) L-INF-Norm, (d) MAE, (e) MASE, (f) MEP, (g) RMSE, (h) SMAPE”.

Table 2
Performance evaluation of the proposed rainfall prediction model with different optimization algorithms

Metrics	DA-MSLTM [30]	EFO-MSLTM [31]	CBO-MSLTM [26]	MFO-MSLTM [27]	HMFCBO-TRI-LSTM
Algorithmic comparison on dataset 1					
MEP	68.251	69.613	51.39	53.766	44.533
SMAPE	1.4181	1.3346	0.91128	0.8046	0.32747
MASE	179.25	174.81	173.93	154.99	142.12
MAE	12.869	13.481	10.184	11.94	9.8513
RMSE	12.749	14.878	14.787	10.125	9.3296
L1-NORM	773.14	722	646.53	685.51	608.94
L2-NORM	87.573	88.686	87.474	84.859	80.344
L-INF-NORM	17.612	16.782	17.867	17.149	15.928
Algorithmic comparison on dataset 2					
MEP	60.029	54.096	62.01	58.699	45.708
SMAPE	1.8185	1.0458	1.135	1.5584	0.68312
MASE	173.22	177.92	177.28	173.79	142.81
MAE	15.506	10.73	13.512	14.914	6.3365
RMSE	14.385	12.139	15.383	11.016	9.4072
L1-NORM	737.38	731.66	772.46	636.97	552.17
L2-NORM	89.644	85.775	85.673	87.66	78.243
L-INF-NORM	17.207	18.777	19.487	18.984	15.689
Algorithmic comparison on dataset 3					
MEP	62.967	64.892	54.099	51.78	43.524
SMAPE	1.4096	0.80447	1.5538	1.2299	0.53583
MASE	170.45	176.49	164.99	160.07	144.99
MAE	11.114	13.107	12.261	11.879	7.55
RMSE	12.161	15.936	12.985	14.041	7.9603
L1-NORM	632.32	640.47	717.58	760.72	607.58
L2-NORM	86.273	84.993	89.671	83.11	77.551
L-INF-NORM	18.283	16.622	17.372	18.985	15.112
Algorithmic comparison on dataset 4					
MEP	57.088	50.526	59.822	54.376	46.242
SMAPE	1.2839	1.2117	1.6214	1.4612	0.54022
MASE	174.11	176.92	169.09	167.37	148.79
MAE	10.335	10.752	13.236	12.013	5.5342
RMSE	15.696	15.406	14.296	14.153	7.8077
L1-NORM	649.89	726.35	799.44	760.42	617.19
L2-NORM	89.571	85.22	83.908	86.305	81.447
L-INF-NORM	17.856	17.739	18.577	17.943	15.014

performance when compared with other existing approaches. Owing to these, an error rate can be reduced while predicting rainfall. Performance analysis over classifiers gives that the proposed model has low MAE, SMAPE, and RMSE, and the overall effectiveness of the model is improved with a low error rate. Throughout the analysis, the developed model provides enhanced performance while validating with existing traditional algorithms.

6.4. Performance evaluation on dataset 2

Evaluation of performance using dataset 2 regarding diverse prediction techniques and various heuristic algorithms are graphically represented in Figs 8 and 9, respectively. While considering the 8^{th}-month variation, the proposed HMFCBO-TRI-LSTM-based rainfall prediction model has assured 11.76%, 6.25%, 13.79%, and 12.89% improved L2-norm performance among Fuzzy, RNN, NN, and LSTM prediction techniques on dataset 2. The algorithmic evaluation shows that the existing CBO algorithm has high SMAPE at 8^{th} to 10^{th}-month variations. The graph analysis shows the existing MFO-TriLSTM algorithm attains a second, more efficient performance to predict the accurate rainfall in L1-Norm. However, the CBO-TriLSTM algorithm attains a high error rate at 4^{th} variations in MAE. Thus, this algorithm does not

Table 3
Performance evaluation of the proposed rainfall prediction model with different prediction techniques

Metrics	FUZZY [28]	RNN [29]	NN [33]	LSTM [29]	HMFCBO-TRI-LSTM
Classifier comparison on dataset 1					
MEP	55.278	55.234	69.591	56.268	44.533
SMAPE	0.94865	0.94623	1.5895	1.0263	0.32747
MASE	157.84	159.57	172.92	156.1	142.12
MAE	13.867	15.946	12.606	16.111	9.8513
RMSE	14.789	11.636	15.796	13.658	9.3296
L1-NORM	724.24	645.73	671.62	709.69	608.94
L2-NORM	82.864	86.353	83.386	88.155	80.344
L-INF-NORM	18.538	19.891	19.299	19.7	15.928
Classifier comparison on dataset 2					
MEP	55.784	54.061	66.57	59.774	45.708
SMAPE	1.552	0.80675	1.3871	0.8458	0.68312
MASE	163.79	161.81	166.93	162.26	142.81
MAE	10.218	14.86	14.682	15.423	6.3365
RMSE	11.905	14.954	10.687	16.278	9.4072
L1-NORM	733.8	734.46	698.5	657.39	552.17
L2-NORM	88.651	89.065	85.691	82.248	78.243
L-INF-NORM	18	19.175	17.069	16.063	15.689
Classifier comparison on dataset 3					
MEP	55.784	54.061	66.57	59.774	45.708
SMAPE	1.552	0.80675	1.3871	0.8458	0.68312
MASE	163.79	161.81	166.93	162.26	142.81
MAE	10.218	14.86	14.682	15.423	6.3365
RMSE	11.905	14.954	10.687	16.278	9.4072
L1-NORM	733.8	734.46	698.5	657.39	552.17
L2-NORM	88.651	89.065	85.691	82.248	78.243
L-INF-NORM	18	19.175	17.069	16.063	15.689
Classifier comparison on dataset 4					
MEP	63.968	58.342	64.364	69.912	46.242
SMAPE	1.1972	1.4537	1.4564	1.6124	0.54022
MASE	156.33	173.28	173.38	162.95	148.79
MAE	14.958	11.193	12.705	10.778	5.5342
RMSE	10.357	13.301	15.431	16.459	7.8077
L1-NORM	674.36	683.2	716.58	771.91	617.19
L2-NORM	87.498	84.563	83.345	87.087	81.447
L-INF-NORM	17.325	19.1	17.912	16.401	15.014

provide accurate performance for rainfall prediction. Yet, our proposed model gets low SMAPE at the same month variation value. The efficacy of the proposed rainfall prediction model is high when compared to other prediction techniques. Throughout the analysis, the outcome shows enhanced performance when compared with existing approaches.

6.5. *Performance evaluation on dataset 3*

The performance calculation of the proposed HMFCBO-TRI-LSTM-based rainfall estimation model on dataset 3 among diverse heuristic algorithms are depicted in Fig. 9, and performance evaluation among different prediction techniques are diagrammatically represented in Fig. 10 respectively. The proposed model gets 8.47%, 18.79%, 33.74%, and 36.47% improved RMSE performance when compared to the DA-TRI-LSTM, CFO-TRI-LSTM, EFO-TRI-LSTM, MFO-TRI-LSTM, and CBO-TRI-LSTM heuristic algorithms in regards to 8th-month variation. This performance analysis shows LSTM prediction technique has low performance regarding MAE, MEP, and SMAPE when the 6th-month variation. In graph analysis, the variation takes place regarding the month. Here, the existing CBO-TriLSTM algorithm attains second higher performance. While validating with L1-Norm, the CBO-TriLSTM algorithm shows a second,

Table 4
Overall performance analysis of the developed model using algorithms

TERMS	DA-MSLTM [30]	EFO-MSLTM [31]	CBO-MSLTM [26]	MFO-MSLTM [27]	PROPOSED
MEP	48.513	59.725	46.429	51.147	38.649
SMAPE	0.91939	1.29	0.95461	1.1789	0.60271
MASE	167.63	149.6	146.34	147.54	141.42
MAE	13.052	12.583	11.65	14.972	8.4175
RMSE	12.045	10.95	9.0863	11.159	6.7247
L1-NORM	773.31	625.48	770.47	656.5	535.66
L2-NORM	80.728	81.488	81.377	80.434	77.772
L-INF-NORM	17.925	17.911	17.991	17.97	14.81

Table 5
Overall performance analysis of the recommended model using several classifiers

TERMS	FUZZY [28]	RNN [29]	NN [33]	LSTM [29]	PROPOSED
MEP	55.675	58.358	56.133	47.558	38.649
SMAPE	0.81461	1.3336	1.1163	1.277	0.60271
MASE	165.65	158.15	146.54	157.13	141.42
MAE	11.029	13.378	13.681	14.343	8.4175
RMSE	10.911	13.597	11.194	14.193	6.7247
L1-NORM	741.8	748.46	715.49	739.52	535.66
L2-NORM	80.072	82.016	85.189	80.359	77.772
L-INF-NORM	17.995	17.931	17.99	17.993	14.81

more efficient performance at the 10^{th} month variation. Thus, the simulation outcome of the developed HMFCBO-TRI-LSTM model shows greater performance while validating with other-state-of-art methods.

6.6. *Performance evaluation on dataset 4*

The performance analysis of the proposed HMFCBO-TRI-LSTM-based rainfall detection model on dataset 4 among different optimization algorithms are illustrated in Fig. 11, and performance calculations over different estimation models are diagrammatically represented in Fig. 10 respectively. The proposed model gets 9.32%, 22.36%, 35.89%, and 37.52% improved RMSE performance when compared to the DA-TRI-LSTM, CFO-TRI-LSTM, EFO-TRI-LSTM, MFO-TRI-LSTM, and CBO-TRI-LSTM heuristic algorithms in regards to 10^{th}-month variation. From the analysis, the proposed prediction technique has higher performance in terms of MAE, MEP, and SMAPE for the 8^{th}-month variation. Various performance measures are validated to show the effective performance of the developed model. In the graph analysis, the MFO-TriLSTM algorithm shows higher error rate while validating with L1-Norm measures. Throughout the entire analysis, the developed model has attained enriched performance when compared with other traditional approaches.

6.7. *Error analysis on the proposed HMFCBO-TRI-LSTM over different algorithms*

The error measures such as MEP, SMAPE, MAE, RMSE, L1-Norm, L12-Norm, and L-INF-Norm of the proposed HMFCBO-TRI-LSTM-based rainfall prediction model are analyzed through different optimization algorithms are depicted in Table 2. This comparison shows that the previously used heuristic algorithms have high MAE. MASE, SMAPE, and MEP but, our proposed model has low MAE, MEP, MASE, and SMAPE. The developed rainfall prediction model secures 14.70%, 7.16%, 12.61%, and 14.77% improved MASE performance on dataset 1 when compared with the existing techniques. The overall effectiveness of the proposed model is high when compared to other DA-TRI-LSTM, CFO-TRI-LSTM, EFO-TRI-LSTM, MFO-TRI-LSTM, and CBO-TRI-LSTM heuristic algorithms.

6.8. Error measures of the proposed HMFCBO-TRI-LSTM over different prediction techniques

The evaluation performed on the proposed rainfall prediction model over different prediction techniques is illustrated in Table 3. The proposed HMFCBO-TRI-LSTM-based rainfall prediction model achieves 46.81%, 57.38%, 39.02%, and 1.31% high SMAPE performance than the Fuzzy, RNN, NN, and LSTM prediction techniques. The overall performance of the proposed model is high in terms of MAE, SMAPE, MEP, and RMSE over different classifiers.

6.9. Overall performance analysis of the developed model for averaging all the datasets

The overall performance analysis of the offered model is evaluated to average the results for all the datasets using diverse algorithms and classifiers, and it is tabulated in Tables 4 and 5. Here, the performance of the developed model attains 30.7%, 14.3%, 30.4% and 18.4% enhanced performance than DA, EFO, CBO, and MFO. Throughout the analysis, the offered model shows superior performance when compared with existing algorithms and classifiers.

7. Conclusion

This paper has designed a hybrid deep learning-based rainfall prediction model among three states to predict low, medium, and high rainfall. The information has been taken from the standard website, and the collected data was given to the preprocessing technique. Data cleaning and normalization techniques were utilized in the pre-processing approach to remove unwanted data. Then, the normalized information has been given to the rainfall recognition phase. In this phase, the AME-FC model was introduced to accurately recognize the rainfall level. Here, the membership function was optimized using the proposed HMFCBO algorithm for enhancing the recognition performance and given the results about low, medium, and high rainfall. In the rainfall prediction phase, the TRI-LSTM model was implemented for the precise prediction of rainfall in three different states. The recognized low, medium, and high rainfall output was applied for the prediction phase, where the training of prediction is carried out separately for all rainfall regions and provides the rainfall prediction result in cm. Hence, individual training in the rainfall prediction has decreased the error rate and computational complexity. The proposed HMFCBO algorithm was utilized for improving the prediction performance in the proposed TRI-LSTM-HMFCBO-based rainfall prediction model. Here, the epochs in the TRI-LSTM classifier are tuned to enhance the prediction performance in terms of MAE, MSE, MEP, RMSE, and SMAPE. To validate the performance of the proposed rainfall prediction model was performed by comparing the error measures with the previously used heuristic algorithms and prediction techniques. Hence, this proposed rainfall prediction model secured 23.52%, 9.21%, 31.43%, and 17.67% more efficient RMSE performance than the fuzzy, RNN, NN, and LSTM prediction techniques. The overall performance of the network was improved by using the proposed rainfall prediction model. Overall, the result analysis of the designed method has shown the MAE has attained 6.3%, which is more effective. Additionally, the MASE of the offered HMFCBO-TRI-LSTM model has attained 142.81%. Thus, the recommended model has achieved enhanced performance while validating with other existing approaches. Hence, the designed HMFCBO-TRI-LSTM model shows more effective performance. Thus, it helps to provide accurate rainfall prediction based on the developed model. Thus, it has the capability to minimize the error rate to increase the system's performance. In future work, various novel deep learning models will be implemented to predict accurate rainfall in an effective manner.

References

[1] Vrieling A, Nelson A, Zambrano F, Meroni M, Tadess T. Prediction of drought-induced reduction of agricultural productivity in Chile from MODIS, rainfall estimates, and climate oscillation indices. Remote Sensing of Environment. 2018 December;

219: 15-30.
[2] Ghorbanzadeh O, Meena SR, van Westen CJ. Rapid mapping of landslides in the Western Ghats (India) triggered by 2018 extreme monsoon rainfall using a deep learning approach. Landslides. 2021.
[3] Chowdhur I, Pal SC, Chakrabortty R, Da B, Roy P, Pal SC, Malik S. Torrential rainfall-induced landslide susceptibility assessment using machine learning and statistical methods of eastern Himalaya. Nat Hazards, 2021.
[4] Manandhar S, Dev S, Lee YH, Meng YS, Winkler S. A Data-Driven Approach for Accurate Rainfall Prediction. IEEE Transactions on Geoscience and Remote Sensing. 2019 Nov; 57(11): 9323-9331.
[5] Cowan T, Wheeler MC, Griffith M, Stone R. Cowan T. Improving the seasonal prediction of Northern Australian rainfall onset to help with grazing management decisions. Climate Services. 2020 August; 19.
[6] Fernando DN et al., A process-based statistical seasonal prediction of May–July rainfall anomalies over Texas and the Southern Great Plains of the United States. Climate Services. 2019 December; 16.
[7] Manandhar S, Dev S, Lee YH, Meng YS, Winkler S. A Data-Driven Approach for Accurate Rainfall Prediction. IEEE Transactions on Geoscience and Remote Sensing. 2019 Nov; 57(11): 9323-9331.
[8] Luini L Capsoni C. A Unified Model for the Prediction of Spatial and Temporal Rainfall Rate Statistics. IEEE Transactions on Antennas and Propagation. 2013 Oct; 61(10): 5249-5254.
[9] Biswas AN, Lee YH, Manandhar S. Rainfall Forecasting Using GPS-Derived Atmospheric Gradient and Residual for Tropical Region. IEEE Transactions on Geoscience and Remote Sensing. 2022; 60: 1-10.
[10] Pucheta J, Rodriguez Rivero C, Herrera M, Salas C, Sauchelli V. Rainfall Forecasting Using Sub sampling Nonparametric Methods. IEEE Latin America Transactions. 2013 Feb; 11(1): 646-650.
[11] Harris A, Hossain F. Investigating the Optimal Configuration of Conceptual Hydrologic Models for Satellite-Rainfall-Based Flood Prediction. IEEE Geoscience and Remote Sensing Letters. 2008 July; 5(3): 532-536.
[12] Shi J, Xu C, Guo J, Gao Y. Real-Time GPS Precise Point Positioning-Based Precipitable Water Vapor Estimation for Rainfall Monitoring and Forecasting. IEEE Transactions on Geoscience and Remote Sensing. 2015 June; 53(6): 3452-3459.
[13] Vuyyuru VA, Apparao G, Anuradha S. Rainfall Prediction Using Machine Learning Based Ensemble Model. 2021 5th International Conference on Information Systems and Computer Networks (ISCON). 2021; 1-4.
[14] Gao J, Zhang P, Jia Y, Song W, Leung H. Short-Term Rainfall Forecasting Using Multi-Layer Perceptron. IEEE Transactions on Big Data. 2020 March 1; 6(1), 93-106.
[15] Wu Z, Zhou Y, Wang H. Real-Time Prediction of the Water Accumulation Process of Urban Stormy Accumulation Points Based on Deep Learning. IEEE Access. 2020; 8: 151938-151951.
[16] Khan MI, Maity R. Hybrid Deep Learning Approach for Multi-Step-Ahead Daily Rainfall Prediction Using GCM Simulations. IEEE Access. 2020; 8: 52774-52784.
[17] Liu JNK, Li BNL, Dillon TS. An improved naive Bayesian classifier technique coupled with a novel input solution method [rainfall prediction]. IEEE Transactions on Systems, Man, and Cybernetics (Applications and Reviews). 2001 May; 31(2): 249-256.
[18] Nam K, Wang F. An extreme rainfall-induced landslide susceptibility assessment using autoencoder combined with random forest in Shimane Prefecture, Japan. Geoenviron Disasters. 2020; 7.
[19] Balasubramanian R, Kaliappan M, Venkatesh C. Rainfall prediction using generative adversarial networks with convolution neural network. Soft Computing. 2021; 25: 4725-4738.
[20] Cao W, Zhang P, Li W. Surface and high-altitude combined rainfall forecasting using convolutional neural network. Peer-to-Peer Networks Applications. 2020.
[21] He L, Gou L, Xia S, Wang W, Xiang Y. A SVR–ANN combined model based on ensemble EMD for rainfall prediction. Applied Soft Computing. 2018 December; 73: 874-883.
[22] S OM, JP A. MapReduce and Optimized Deep Network for Rainfall Prediction in Agriculture. The Computer Journal. 2020 Jan; 63(1): 900-912.
[23] Zhang X, Mohanty SN, Parida AK, Pani SK, Dong B, Cheng X. Annual and Non-Monsoon Rainfall Prediction Modelling Using SVR-MLP: An Empirical Study From Odisha. IEEE Access. 2020; 8: 30223-30233.
[24] Guhathakurta P. Long lead monsoon rainfall prediction for meteorological sub-divisions of India using deterministic artificial neural network model. Meteorological Atmospheric Physics. 2008; 101: 93-108.
[25] Pham BT, Le LM, Le T-T, Thi Bui K-T, Le VM, Ly H-B. Indra Prakash Development of advanced artificial intelligence models for daily rainfall prediction. Atmospheric Research. 2020 June; 237.
[26] Kaveh A, Mahdavi VR. Colliding bodies optimization: A novel meta-heuristic method. Computers & Structures. 2014 July; 139: 18-27.
[27] Mirjalili S. Moth-flame optimization algorithm: A novel nature-inspired heuristic paradigm. Knowledge-Based Systems. 2015 November; 89: 228-249.
[28] Iriany A, Mahmudy WF, Wahyuni I. Rainfall prediction in Tengger region Indonesia using Tsukamoto fuzzy inference system, 2016 1st International Conference on Information Technology. Information Systems and Electrical Engineering (ICITISEE). 2016.
[29] Salehin I, Talha IM, Hasan MM, Dip ST, Saifuzzaman M, Moon NN. 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). 2020.

[30] Mohammadi B, Bahrami-Pichaghchi H, Aghelpour P, Mehdizadeh S, Duan Z. A novel hybrid dragonfly optimization algorithm for agricultural drought prediction. Stochastic Environmental Research and Risk Assessment. 2021; 35: 2459-2477.
[31] Peng Z, Zhang S, Yan J, Li D, Liu P. Optimal design of shielding structure in ultra-high voltage GIS bushing using artificial fish-swarm algorithm. 2017 1st International Conference on Electrical Materials and Power Equipment (ICEMPE). 2017; 57-60.
[32] Khosla E, Ramesh D, Sharma RP, yakotey S. RNNs-RT: Flood based Prediction of Human and Animal deaths in Bihar using Recurrent Neural Networks and Regression Techniques. Procedia Computer Science. 2018; 132: 486-497.
[33] Smith A, Yuan H, Deng L, Dixon N. Machine learning prediction of landslide deformation behaviour using acoustic emission and rainfall measurements. Engineering Geology. 2021 November; 293: 106315,
[34] Ambati LS, Narukonda K, Bojja GR. Dave Bishop Factors Influencing the Adoption of Artificial Intelligence in Organizations-From an Employee's Perspective. Adoption of AI in Organization from Employee Perspective. 2020.
[35] Tabjula JL, Kanakambaran S, Kalyani S, Rajagopal P. Balaji Srinivasan Outlier analysis for defect detection using sparse sampling in guided wave structural health monitoring. Structural control and Health Monitoring. 2021 March; 28(3).
[36] Tabjula J, Kalyani S, Rajagopal P. Balaji Srinivasan Statistics-based baseline-free approach for rapid inspection of delamination in composite structures using ultrasonic guided waves. Structural Health Monitoring. 2021; 21(6).
[37] Ambati, Sai L, El-Gayar. Omar Human Activity Recognition: A Comparison of Machine Learning Approaches. Journal of the Midwest Association for Information Systems. 2021; 1.
[38] Kumawat M, Khaparde A. Time-Variant Satellite Vegetation Classification Enabled by Hybrid Metaheuristic-Based Adaptive Time-Weighted Dynamic Time Warping. International Journal of Image and Graphics. 2022; 2450016.