Research Article

International Journal of Advanced Technology and Engineering Exploration, Vol 9(89)
ISSN (Print): 2394-5443 ISSN (Online): 2394-7454
http://dx.doi.org/10.19101/IJATEE.2021.874647

A hybrid adaptive grey wolf Levenberg-Marquardt (GWLM) and nonlinear autoregressive with exogenous input (NARX) neural network model for the prediction of rainfall

Sheikh Amir Fayaz[1], Majid Zaman[2*] and Muheet Ahmed Butt[1]
Department of Computer Sciences, University of Kashmir, Jammu and Kashmir, India[1]
Directorate of IT&SS, University of Kashmir, Jammu and Kashmir, India[2]

Received: 21-August-2021; Revised: 06-April-2022; Accepted: 09-April-2022

Abstract
Rainfall prediction, a type of weather forecasting, has a big impact on agriculture and farming, as well as other industries like natural disaster management. One of the most crucial aspects of today's climate is accurate and timely rainfall prediction. Such issues could be avoided if worst-case weather scenarios could be predicted ahead of time and timely warnings issued. The "nonlinear autoregressive (AR) with exogenous inputs" (NARX) neural network (NN) prediction model has been introduced in this paper for the prediction of rainfall using historical geographical data from the Kashmir province of the union territory of Jammu & Kashmir, India. The methodology was developed using six years of historical-geographical data from three different substations in Kashmir. Four explanatory independent variables like maximum temperature, minimum temperature, humidity measured at 12 a.m., and humidity measured at 3 p.m. as well as a target variable indicating the amount of rainfall were considered. For a better computational time and performance accuracy, the proposed algorithm is trained using the grey wolf optimizer (GWO) and the Levenberg-Marquardt (LM) algorithms. The grey wolf Levenberg-Marquardt (GWLM) and NARX implementation methodology was deemed one of the best-fit models. The obtained values for the mean squared error (MSE) and regression value (R) predictions are 3.12% and 0.9899% in the case of training. The values are 0.144% and 0.9936% in validation, and 0.311% and 0.9988% in testing. The suggested model was then compared to a number of traditional and ensemble machine learning (ML) methods, and it was determined that the proposed model performs better with less processing time. The grey wolf Levenberg-Marquardt nonlinear AR with external inputs (GWLM-NARX) model is found to be a more practical neural network model to use.

Keywords
NARX model, Grey wolf optimizer, Geographical data, Rainfall prediction, Levenberg-marquardt algorithm.

1.Introduction
The climate change study looks at how the weather has changed over time. Rainfall forecasting is one of the most significant climatic change activities. It involves specific features like humidity, temperature, etc. to predict rainfall in a certain region. Predicting rainfall has been the most difficult task in weather analysis for every researcher in the contemporary climate. Precision, accuracy, and timely rainfall prediction have been impossible due to the complexity of the measures.

But precise and real-time rainfall prediction can cause a variety of operational security measures in ongoing system implementation, agricultural responsibilities, travel cases, and alarming flood conditions, among other things [1].

A variety of soft computing models have been used to forecast rainfall. Novel techniques such as artificial neural networks (ANN) and hybrid models are being utilized to predict rainfall. Although these procedures give precise and rapid results, their architecture is complex, which is one of their major disadvantages. To convey their results, these models use a black box method, in which the user can only look at the input and output numbers without knowing how the model works internally [2].

*Author for correspondence

Some of the major challenges faced by researchers in predicting rainfall using various machine learning (ML) techniques are mentioned below.

1. The authors mentioned in [3] that future studies are still needed for timely warning of short-term flash flood forecasting, i.e., flash flood prediction accuracy needs to be increased.
2. The proposed model can be extended in [4] to predict rainfall in advance and evaluate the occurrence of floods, allowing for the development of an early warning system.
3. With the local features, this model works well for local rainfall forecasting (temperature, humidity, pressure, etc.). Other mining methods with neural networks (NN) can thus be used to improve the study [5].

The above-mentioned challenges act as objectives of the present study and in this paper. We try to fulfil these objectives by introducing a hybrid adaptive grey wolf Levenberg-Marquardt (GWLM) and nonlinear autoregressive (AR) with exogenous input (NARX) NN model for the prediction of rainfall. The GWLM algorithm is a combination of grey wolf optimization (GWO) and Levenberg-Marquardt (LM) algorithms, which is hybridized with the NARX model and is used for the effective and adaptive prediction of rainfall.

In this paper, we have used the GWLM algorithm with the NARX model for the prediction of rainfall based on the weather data in Kashmir province at three different locations. Various performance parameters are evaluated in this study, and finally, a brief comparison has been made between these algorithms.

Traditional and ensemble algorithms, including decision trees (DT), distributed decision trees (DDT), random forest (RF), logistic model trees (LMT), and gradient boosting (GB), have been implemented on the historical geographical dataset of Kashmir province and the performance has been calculated by all the algorithms [6]. In order to check the performance of the historical geographical data using a NN approach, a mixed adaptive model called the grey wolf Levenberg-Marquardt nonlinear auto regression with external inputs (GWLM-NARX) model has been used in this study. The author [7] proposed a simple NARX model of work where they exclusively validate the data based on the mean squared error (MSE) and root mean square error (RMSE) values, and this motivated us to work on the same algorithm with the GWLM model to check the overall performance on this data. This mixed adaptive model performs better because of its good fitting capability.

This paper is organised as follows: Section 2 introduces the state of knowledge of the work and the various challenges faced by the many researchers in predicting of the rainfall. Section 3 introduces the methodologies which is adopted in this paper, and also in this section, the basic architecture of the NARX model is elaborated. Section 4 deliberates the result along with the experimental setup, dataset description, database description, experimental results and performance metrics were evaluated. A comparative analysis was evaluated in section 5. Section 6 briefly explains the conclusion of the paper.

2.Literature review

2.1State of knowledge

Various scholars have done a significant amount of effort in the past few years to predict rainfall using various data mining and ML techniques. The following are the few of them:

Malerba et al. [8] published a study based on the step-wise model tree induction (SMOTI), which has two types of nodes: regression and splitting nodes. This model was compared to other induction systems and rated based on the empirical investigation. The benefit of this innovative model is that it builds model trees (MT) with multiple regressions at the leaves, and the results were compared to other models on six different datasets, proving its usefulness. The latter was tested on 12 benchmark datasets, resulting in more precise MT. SMOTI outperforms other MT such as the hybrid regression tree learner (HTL), scalable linear regression tree algorithm (SECRET), generalized, unbiased interaction detection and estimation (GUIDE), autoregressive integrated moving average (ARIMA) and others on the benchmark dataset for both MT and regression trees.

Aswini et al. [9] suggested a statistical technique to predict rainfall based on historical data received from the Tamil Nadu government of the Tamil Nadu region in India, containing 7 variables employed in prediction. Naive Bayes (NB), K-nearest neighbor (KNN), DT’s, fuzzy logic, and NN were among the methods employed in the implementation. The DT’s and KNN fared better in terms of accuracy, according to the findings of the experiments. Hemalatha [10] applies the DT approach to 20–30 instances of past data using C4.5 and the iterative dichotomizer 3 (ID3) algorithm. Petre [11] uses the classification and regression techniques (CART) algorithm to create the DT methodology on 48 instances of data available from 2002 to 2005, and Ji et al. [12] uses the C4.5 and CART algorithms to implement the DT technique on previous data. When compared to other algorithms, it was determined that an individual's DT had a 93–99 percent accuracy. Zainudin et al. [13] presented a comparison analysis in which the authors looked at numerous multiple classifiers for rainfall prediction, including NB, support vector machine (SVM), DT, NN, and RF. The data used in the trial were gathered in Malaysia between 2010 and 2014. Temperature, rainfall, relative humidity, water level, and other variables were used to make predictions. Sharaff et al. [14] proposes a study for rainfall prediction in which the influence of machine learning classification approaches such as NN, DT algorithms on rainfall prediction was investigated, and it was discovered that DT surpasses the others through experimental analysis.

Using a non-linear autoregressive model NN technique, Nourani et al. [15] and Le et al. [16] forecasted daily rainfall. This research was carried out in the Hoa Binh City, Vietnam, using eight years of time series meteorological data. Temperature, relative humidity, solar radiation, wind speed, and other variables were used in the research. The coefficient of correlation (R), RMSE, and the mean absolute error (MAE) were used to validate the generated model. With an R value of 0.8846, RMSE = 5.3793, and MAE = 3.0218, this investigation produced some encouraging findings.

Mohd et al. [7], proposes a time series prediction model based on the GWLM-NARX model. This model was used as an adaptive prediction model, with rainfall data from the previous period as input and results computed using the NARX model's GWLM algorithm.

Adnan et al. [17], offered a comparison analysis of four ML techniques for rainfall modelling. The capabilities of the optimum pruning extreme learning machine (OPELM), multi-variate adaptive regression splines (MARS) is the subject of this work. It was discovered that accuracy improves significantly, with RMSE and MAE improvements of more than 90% in most cases.

Patil and Bedekar [18] suggest a research on a few typical data mining approaches for rainfall prediction. In this study, several ML approaches such as DT, NB, and others were compared. The suggested model is tested using rainfall data gathered in Delhi, India, between 1997 and 2016. The strategy might be put to the test to determine if it increases rainfall forecasting accuracy. According to experimental data, the DT classifier is successful in extracting rainfall forecasts. To assist local heavy rain emergency services, DT Classifier apps produce accurate and rapid rainfall predictions.

Paudel et al. [19] proposed a comparison of various ML techniques for rainfall prediction. Long short term memory (LSTM), NARX, Back-propagation neural network (BPNN), and nonlinear autoregressive (NAR) models are examples of these techniques, which work on the stochastic time-series dataset to check the effect of rainfall on the radio signal degradation. On the basis of error comparison, they concluded that the LSTM works much better than the rest of the ML approaches used in their research.

Fayaz et al. [6] implement a stepwise ML approach for the prediction of the rainfall in the Kashmir region of India. In their study, they implemented a LMT algorithm where the leaf node predicts the model functions based on the logistic regression approaches. The dataset used in their study was of Kashmir province, collected from the Indian meteorological department (IMD) Pune, and it contains data from the years 2012 to 2017. Season, temperatures at different intervals, humidity at 12 am and 3 pm, and rainfall is some of the parameters used in the study. This study concludes with a comparative analysis where various traditional and ensemble approaches are compared with the performance of the LMT, and they prove that the accuracy measure of the LMT is much better than the other models used in the study.

Based on the literature reviewed in this paper, we learned that the NARX model can be utilized for a variety of time-series prediction applications, but it has yet to become a standard method in the artificial intelligence sector due to a lack of research. We need to analyze the model's performance using the multiple threshold datasets and, as a result, it needs to be re-evaluated. Furthermore, no Python library has ever included this model.

3.Methods

In this section, the main goal is to study and discuss the rainfall prediction using the NARX NN approach. In this method, the historical geographical data of rainfall are given as the input to the model, and it will automatically to be used to forecast the rainfall. As compared to other traditional and ensemble models, NARX performs better on the data when using a recurrent feed-forward neural network (FFNN). These predictions are one of the major roles where this model shows better performance than other ML methods by providing an effective learning rate and an optimal output solution.

3.1Autoregressive (AR) models

The AR model attempt to anticipate a series entirely on the basis of prior data or lags. The AR models of order 1, often known as AR models, are those that rely only on prior values or lags. Mathematically, as shown in Equation 1:

$$Y_t = \omega + \phi(Y_{t-1}) + e_t \quad (1)$$

Where, Y_t = Target
Y_{t-1} = Lagged Target
e_t = error
ω & ϕ are the intercept and coefficient factors

AR models implements a recursive approaches from the beginning of the data and such models are also called as long term models. These models are mathematically defined as:

$$Y_t = \omega + \phi(Y_{t-1}) + e_t \quad (2)$$
$$Y_{t-1} = \omega + \phi(Y_{t-2}) + e_{t-1} \quad (3)$$

Since in Equation 2 and Equation 3 $Y_t$depends on the value of Y_{t-1} and Y_{t-1}depends on Y_{t-2} and so on, it sorts of recursively plugs into a general mathematical equation by combining these two equations as shown in Equation 4 and Equation 5:

$$Y_t = \omega + \phi(\omega + \phi(Y_{t-2}) + e_{t-1}) + e_t \quad (4)$$
$$Y_t = \frac{\omega}{1-\phi} + \phi^t Y_t + \phi^{t-1} e_2 + \phi^{t-2} e_3 + \cdots + e_t \quad (5)$$

Here in this case the first observation($\phi^t Y_t$)does matter. The first observation in a roundabout way still has a small effect. The whole point of stationery is that this effect goes away. So the effect of these shocks that happened will have only a little effect if the coefficient in absolute value is less than 1 i.e. $|\phi|$ < 1. That is why they are called long-term models.

3.2Architecture of NARX model

There are two different designs of NARX NN model, which are: open loop model and closed loop model (*Figure 1*) which are given below by two function mappings Equation 6 and Equation 7:

$$\hat{y}(t+1) = F(y(t), y(t-1), \ldots y(t-n_y), x(t+1), x(t), x(t-1), \ldots, x(t-n_x)) \quad (6)$$

$$\hat{y}(t+1) = F\left(\hat{y}(t), \hat{y}(t-1), \ldots \hat{y}(t-n_y), x(t+1), x(t), x(t-1), \ldots, x(t-n_x)\right) \quad (7)$$

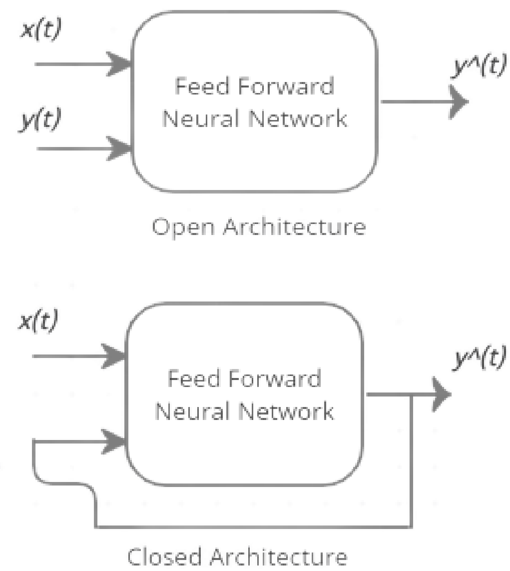

Figure 1 Basic architectures of NARX NN model

Here, in this study open model architecture has been used for the training phase based on the availability of historical geographical data. There are two major advantages of using this model viz:

- Use of the true values as input in FFNN is more precise and the normal training algorithms for multilayer perceptron (MLP) networks can be used.
- Furthermore, after the training phase, this model is converted into closed architecture, which is more valuable for multi-step-ahead prediction [20, 21].

This section depicts the proposed model's architecture, which is made up of three layers: input, output, and hidden layers. Exogenous input vector, delayed exogenous input vector, and regressed output vector are the three types of vectors in the input layer that are used to set arbitrary weather factors. The NARX model's mathematical problem definition is in Equation 8:

$$y(t) = f\left(x(t-1), \ldots x(t-d), y(t-1), \ldots . y(t-d)\right) \quad (8)$$

where, y(t) is the output of the rainfall prediction using NARX model and (t − d) is the rainfall time series data. This model performs much better and predicts the rainfall in an effective manner by determining the optimal weights of the input exogenous vector. Afterwards, these weights are tuned as per the proposed GWLM algorithm based on the error produced between the predicted output and the actual output.

GWLM is used to obtain the optimal weights of the NARX model where it is hybridized between GWO and LM algorithm for permitting the adaptive prediction. The architecture of the GWO algorithm is simple & takes less storage and thus easy to implement and fast in process as compared to the Gauss-Newton (GN) algorithm. There are many advantages of using GWLM–NARX NN model as this approach is used for the effective learning as compared to other models and the convergence rate is much faster.

The comprehensive stepwise approach used in this study is depicted in *Figure 2*. The data collection process is followed by data pre-processing and data cleaning, which is done using various analysis tools such as principal component analysis (PCA). After that ANN approach was implemented in which testing, training and validation methodologies were implemented and then GWLM-NARX model was implemented on the time series data and later performance of the algorithm has been calculated.

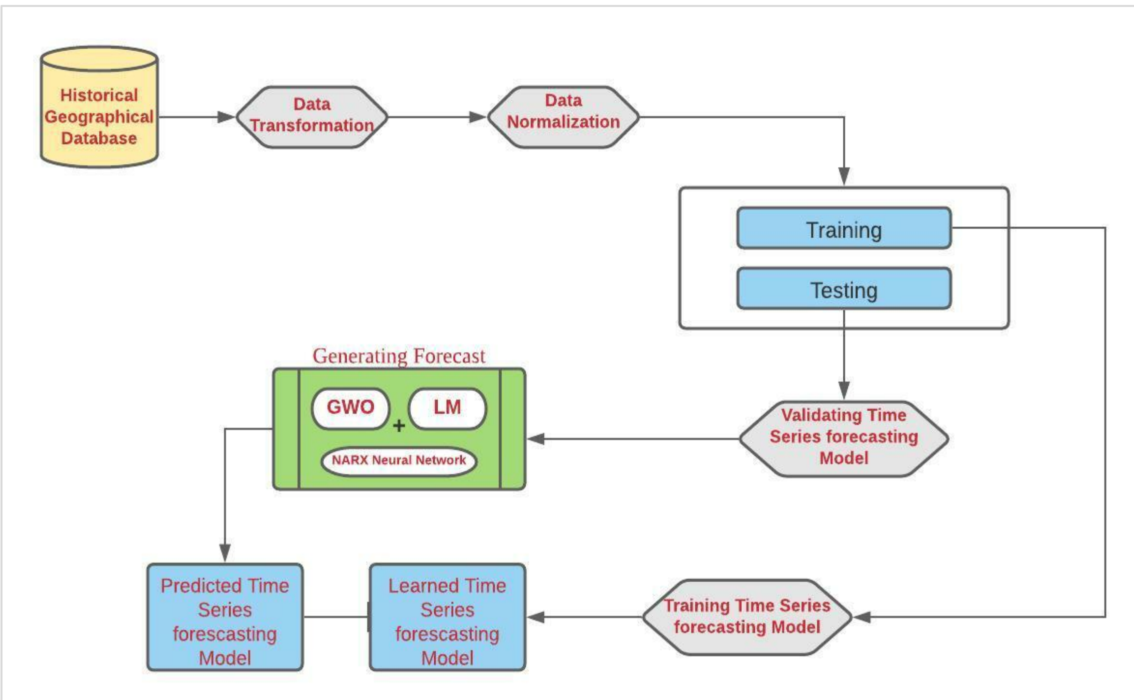

Figure 2 Implemented workflow methodology

4.Experimental analysis

The proposed model for rainfall prediction has been described in this section. Various other criteria, such as experimental setup and results, dataset & database description, and so on, are also discussed.

4.1Experimental setup

The implementation of the proposed model is carried out using the visualization tool matrix laboratory (MATLAB) on windows operating system.

4.2Dataset description

Historical weather data [20−22] has been used for prediction purposes using various ML algorithms. Here, in this study, the research has been carried out on past weather data of Kashmir province, India. The data were collected from the national data centre (NDC), Pune, India. It is the primary organisation in charge of weather observations, forecasting, and other related tasks. The NDC is one of the World Meteorological Society's six provincially designated meteorological data centres.

This data contains a number of parameters that are important in rainfall forecast. These parameters come from the Kashmir province's three zones: the north zone (Gulmarg), the south zone (Qazigund), and the central zone (Srinagar) [23]. The dataset contains four independent attributes, including maximum temperature, minimum temperature, and humidity at 12 am and 3 pm. The target attribute rainfall contains discrete values of yes or no, which depicts the presence and absence of rainfall. The dataset contains approximately 5951 individual entries with three different stations as discussed. The snapshot of the dataset used in this study is shown in *Table* 1.

Table 1 Geographical dataset

Max_temp	Min_temp	Humidity@12	Humidity@3	Rainfall
12.5	11.3	88	97	**N**
16.5	17.2	95	81	**N**
19.5	13.4	75	86	**Y**
18.7	13.5	76	95	**N**
24.6	17.8	96	72	**Y**
30.7	15.8	86	98	**N**
15.5	12.8	94	88	**Y**
20.5	16.4	73	98	**Y**

Clouds can have a cooling effect during the day because they block some of the sun's energy that otherwise would get to the earth's surface, but especially at night, they have a warming effect. It is a kind of like switching gears from location to location and the effect of clouds on the temperature changes. During a clear day and an overcast day, the temperatures are reasonably different because the clouds are blocking some of the incoming heat from the sun [24, 25]. So what are the temperature ranges? We would say that for an overcast day, it is less and there is less fluctuation from the hottest to the coolest as compared to a clear day (*Figure 3*).

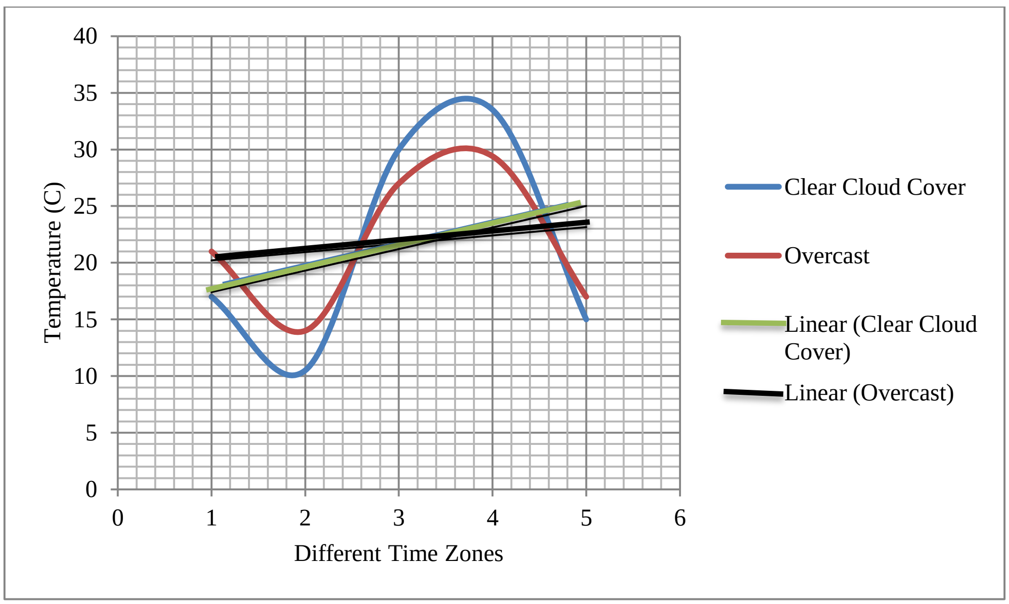

Figure 3 Average temperature on clear cloud cover and overcast days

Location and the season also play a major role in the cloud cover. As we know, Kashmir belongs to the temperate zone and usually the winter season remains for around 4-5 months. Thus, the cloud cover has a major effect on the temperature, and it will automatically affect the quantum of the rainfall. The net effect of the cloud cover is to measure the kind of dampness in the range throughout the annual temperature. The figure below depicts the overall distribution of the single variables as well as relationships between them in the dataset (*Figure 4*).

It defines the relationship between the single attribute and the other attributes, including the quantum of the rainfall, and it essentially defines how the dataset is balanced in its own form.

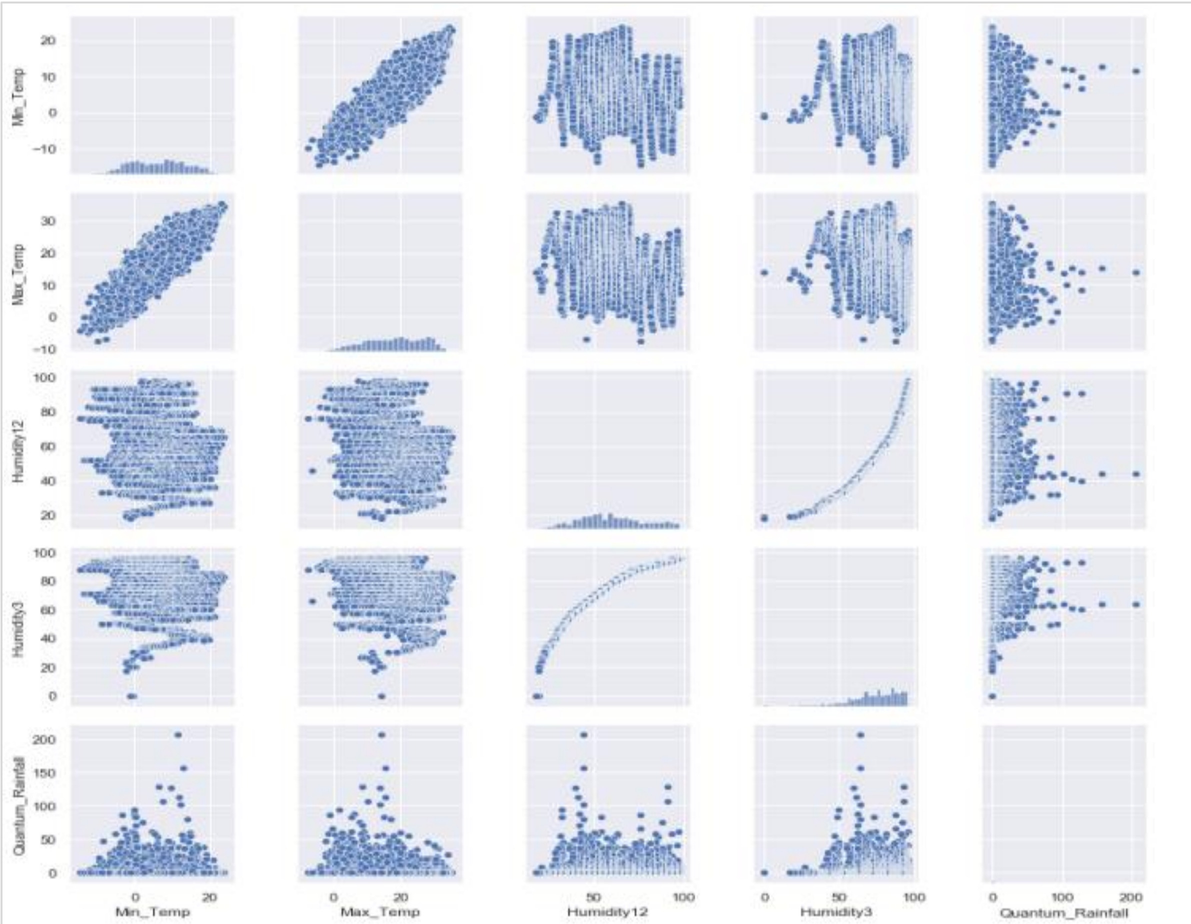

Figure 4 Distribution of single variables and relationships between other variables in the dataset

For example, in *Figure 4*, we can depict the relationship between the humidity at 3 pm and the humidity at 12 am as directly proportional, i.e., with the increase in the value of one attribute means it will also increase the values of the other attributes as well.

4.3Database description

To carry out the analysis of the historical geographical dataset, 5 M×N dimensional databases are used which contains the database of the original dataset, 3 different databases for different zones with their respected station Id's and 1 database for voting purpose which checks the overall prediction of the data. Details of such databases are given below:

4.3.1Zone data tables

There are 3 different zones and each zone contains the same set of fields. The zone data table includes some of the following parameters: ctmax, ctmin, chumid12, chumid3, crfall, cpred. ctmax and ctmin specifies the maximum and minimum temperatures respectively, recorded in degree celsius (°C). Chumid12 and chumid3 respectively specifies the humidity values at 12 am and 3 pm. Crfall field defines the quantum of the rainfall measured in mm.

The table structure of one of the zones is shown below (*Table 2*). Similarly, all other zones follow the same structure with the same set of parameters.

Table 2 Table structure of the zone with station id 42026

S. No.	Field	Type	Null	Default
1	rno	Int(11)	Yes	Null
2	season	Varchar(6)	Yes	Null
3	ctmax	Char(2)	Yes	Null
4	ctmin	Char(2)	Yes	Null
5	chumid12	Char(2)	Yes	Null
6	chumid3	Char(2)	Yes	Null
7	crfall	Char(1)	Yes	Null
8	cpred	Char(1)	Yes	Null

4.3.2Combined voting classifier data table

The combined voting data table contains d26, d27, d44 and rf extra fields than zone table (*Table 2*) which indicates the respective prediction values of all the zones (42026, 42027, 42044) and rf field is used as the voting field which chooses the majority value of the 3 zones. The table structure is shown below (*Table 3*) where all the attributes except rno contains character values and rno stores the integer value.

Table 3 Combined voting_classifier data table structure

S. No.	Field	Type	Null	Default
1.	rno	Int(11)	Yes	Null
2.	season	Varchar(6)	Yes	Null
3.	ctmax	Char(2)	Yes	Null
4.	ctmin	Char(2)	Yes	Null
5.	chumid12	Char(2)	Yes	Null
6.	chumid3	Char(2)	Yes	Null
7.	crfall	Char(1)	Yes	Null
8.	d26	Char(1)	Yes	Null
9.	d27	Char(1)	Yes	Null
10.	d44	Char(1)	Yes	Null
11.	rf	Char(1)	Yes	Null

4.4Experimental results

This section shows the various experimental results based on rainfall prediction rate. *Figure 5* shows the NN training performance with MSE value as the best validation performance of 0.14933 at epoch value 7. *Figure 6* shows the NN training state at epoch value 13 with gradient = 27.3959, Mu = 0.1 and validation checks = 6. *Figure 7* shows NN error histogram with 20 bins and the value of zero error should be between 0 & 1 for less error and good performance. *Figure 8* shows the time-series response of target & output and error in the predicted and actual values. *Figure 9* shows the individual regression values of training, test, validation and all. The overall regression values in all the cases lie between 98.9% to 99.9% and *Table 2* shows the results with MSE values and R values.

4.5Illustration: result evaluation

Figure 5 shows how the NN model got trained. We can see how the error decreased as the number of epochs increased in all three datasets. Furthermore, we can see that the validation stops after epoch 7 and it checks up to epoch 13 if it starts again to decrease, but it didn't happen to generalize the model. Thus, there is a constant linear line after epoch 7, with the best validation performance of 0.14933.

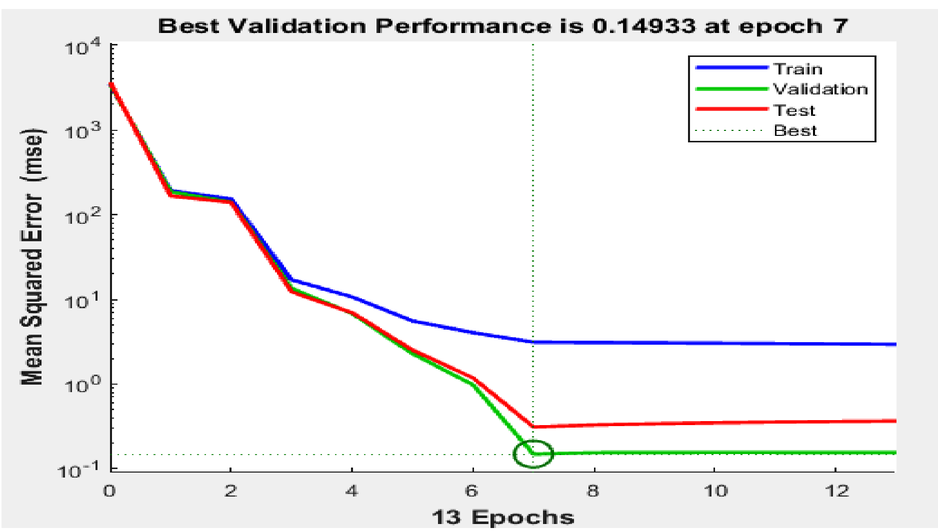

Figure 5 Best validation performance metrics

Training state tells us how the validation check is being done. Once the samples have been trained, we need to check the validation samples. We have a particular minimum target that can be compared whether our particular network is working or not and after that we can go for testing samples. Here *Figure 6* shows that the validation check has a value of 6 at epoch 13 has a better performance.

An error histogram is used to compile the amount of error present in the target outputs. If we want a better performance, then we can retrain the model and, accordingly, at each iteration, we can check the value in the error histogram. *Figure 7* shows the NN error histogram with 20 bins, and the value of zero error should be between 0 and 1 for less error and good performance.

After preparing a FFNN, an error histogram is used to check the error between target values and predicted values. The value of the error histogram can show the negative values because it indicates how these values differ from each other. The total error range is divided into "bins," which are smaller vertical bars. The number of samples from your dataset that fall into each bin is represented on the Y-axis. For example, we have a bin corresponding to the error of -3.923 to 0.5531, and the height of that bin for the training dataset is below, but near 1400, and the height of that bin for the validation and test dataset is between 1400 and 1600. It means that a large number of samples come from various datasets and have an error rate in the range below. On the error axis, the zero error line corresponds to the zero error value (i.e., the X-axis). *Figure 8* shows the response graph between the output and target of training, validation, and testing with the error. We can see that the model fits best as the error is the very least in predicting the outputs in all the cases. Thus, it best fits the model on the time-series data. By trial and error, it was found that 2-time delays and 15 hidden neurons are the most influential parameters for the algorithm.

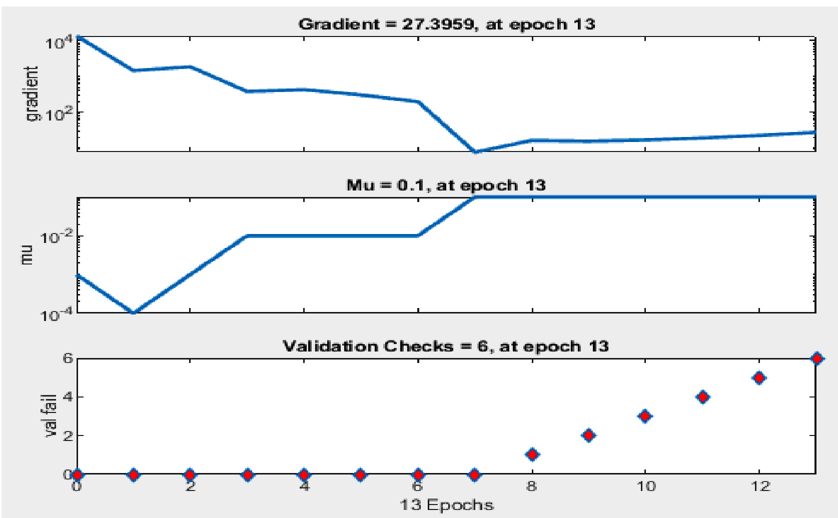

Figure 6 Validation checks with gradient values

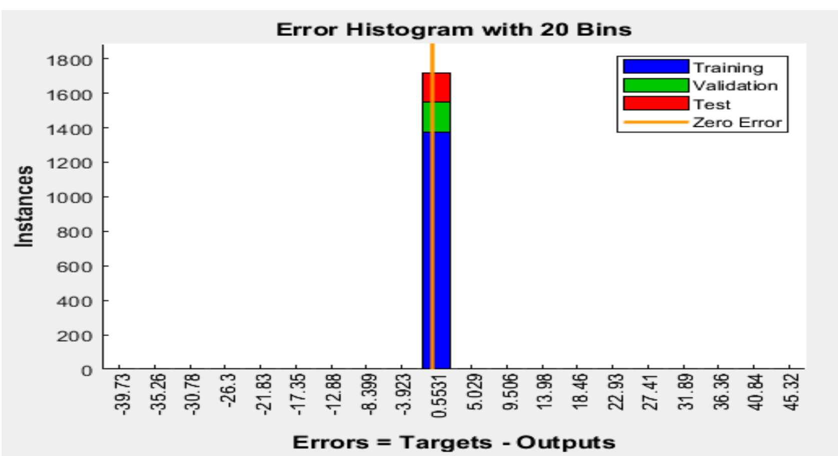

Figure 7 Error histogram

When we work with ML, we divide the dataset into 3 separate sub-datasets, including the training set—used to train the ML model; the validation set—usually used for improving a ML model; and the test set—used for testing the accuracy of the model. The regression values of the training, validation, and testing sets are shown in *Figure 9* and all the values are greater than 0.98, i.e., the data is well fitted for the target.

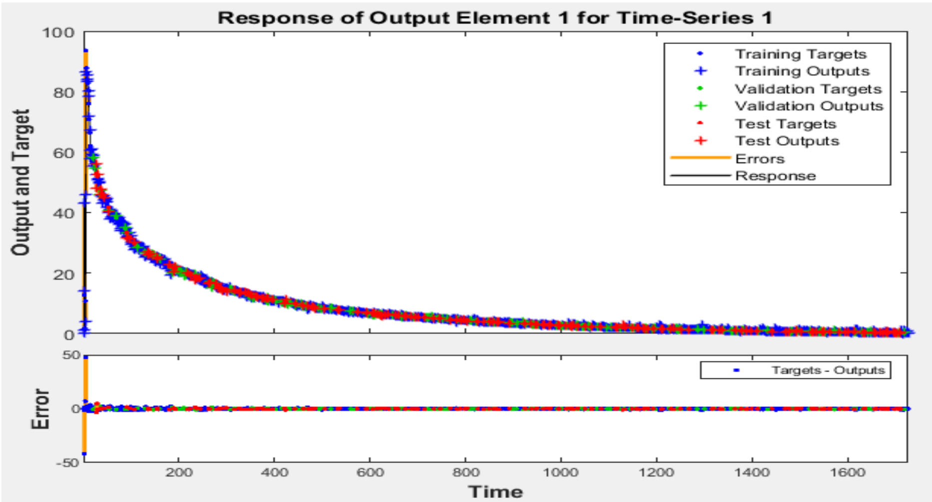

Figure 8 Observed and target outputs of training, testing and validation sets

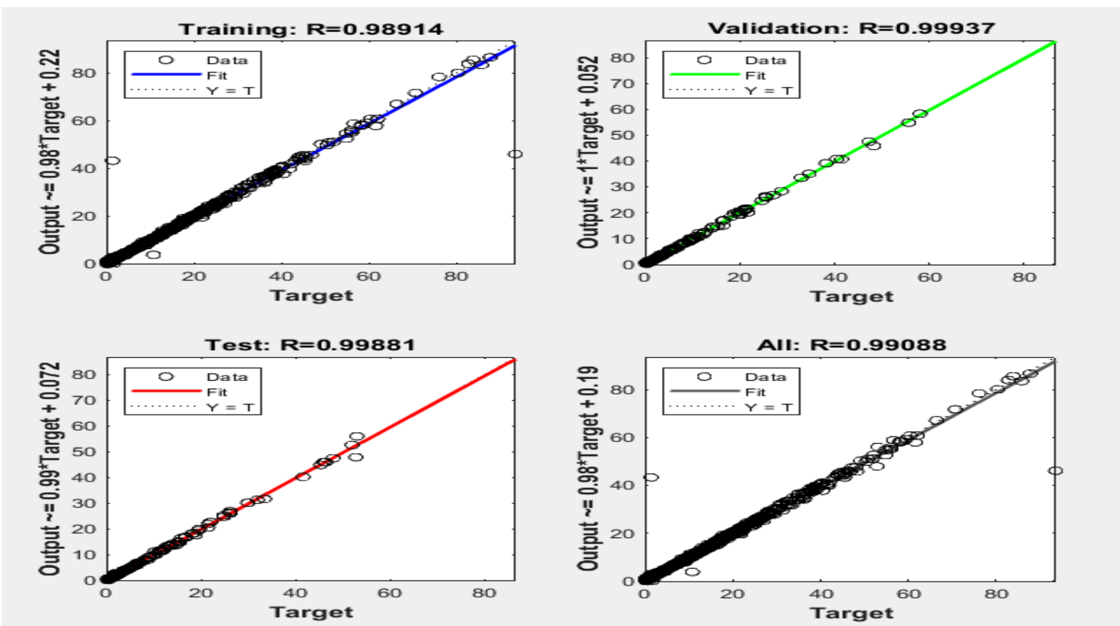

Figure 9 Regression values obtained from experimental analysis using GWLM-NARX model for the prediction of rainfall

4.6Performance metrics

In this study, MSE and R-values of the proposed model were analyzed and for effective results the value of error should be minimized. It was observed that the MSE value of the training data is 3.123% with target values equal to 1382 and similarly MSE value of validation and testing is 0.149% and 0.311% respectively. Furthermore, the regression values in all the three cases remain between 98-99%. The overall accuracy statistics are shown in *Table 4*.

Table 4 Accuracy statistics with MSE and regression values

	Target Values	MSE	R
Training	1382	3.12385	0.989136
Testing	173	0.311559	0.998810
Validation	173	0.149334	0.999365

4.7Result: implications and impact

Following the application of the GWLM-NARX model to the historical geographical dataset, prediction has significantly improved while the mean square error rate has been reduced marginally.

In MATLAB, the execution performance of GWLM-NARX is slightly better than the execution performance of traditional algorithms. This has a significant impact on the outcome obtained when GWLM-NARX is taken into account. However, this is dependent on how the specific dataset is executed, and it may differ for other threshold datasets.

4.8Comparative analysis

This section defines the comparison of the NN model with other traditional and ensemble methods like DT's, DDT's, RF, LMT and M5 model trees (*Table 5*). The same data of Kashmir province were implemented using these algorithms and the overall performance was calculated. It was observed that these ensemble methods take more data and time to train the model as compared to GWLM-NARX model. Moreover, the overall performance in these traditional & ensemble ML algorithms ranges between 66-87% with the mean square error of 6.726% in case of M5 model trees, whereas the MSE value in case of this proposed algorithm ranges between (3.123-0.144) %, thus defining the effectiveness of this proposed hybrid model.

Table 5 Comparative analysis of GWLM-NARX and other algorithms for geographical data set of Kashmir

Algorithm	**ID3**	**SVM**	**NB**	**KNN**	**Fuzzy DT**	**DDT**	**C4.5**	**LMT**	**M5 Model Tree**	**GWLM-NARX**
Model	30-70	30-70	30-70	30-70	30-70	30-70	30-70	30-70	R^2= 0.478	**Regression**
Accuracy	80.12	81.07	76.70	78.94	77.75	78.46	66.99	**87.23**	**MAE** = 1.689	**(R)**
No. of Rules	51	---	---	---	---	21	55	**10**	**MSE** = 6.726	**Testing:** 0.9988%
Error	19.87	18.92	23.29	21.05	22.24	21.54	33.01	**12.77**	**RMSE** = 2.593	**Validation:** 0.9936%
Precision	0.812	0.845	0.818	0.848	0.841	---	0.880	0.892	**MSD** = 0.844	**Training:** 0.9899%
Recall	0.938	0.897	0.864	0.857	0.846	---	0.726	0.973		**MSE**
Cohen Kappa	0.456	0.519	0.411	0.485	0.458	---	-0.022	0.102		**Testing:** 0.311%
F-measure	0.87	0.871	0.84	0.853	0.844	---	0.796	0.931		**Validation:** 0.144%
Specificity	0.938	0.897	0.864	0.857	0.846	---	0.238	0.098		**Training:** 3.12%
Sensitivity	0.938	0.897	0.864	0.857	0.846	---	0.726	0.973		

5.Discussion

The implementation of GWLM-NARX actually divides time series data into conclusive sequence and error sequence. In this study, we propose an adaptive GWLM-NARX model to fit the conclusive sequence of the model. Several researchers concluded that when various traditional and ensemble approaches, such as DT, DDT, RF, LMT [26−29], are used, performance stagnates. When these models are taken into account, the performance accuracy remains somewhere in the middle (66–87%). But after the implementation of the GWLM-NARX model, it was observed that the accuracy statistics increases by having MAE marginally reduced. As a result, the implementation concluded that resolution stagnation is incorrect, and the GWLM-NARX model's results were accordingly improved.

5.1Limitations

Although GWLM-NARX has better simulation and forecasting precision, it was discovered after its implementation that the model reaches the best accuracy measure in a reasonable amount of time. This model takes a long time to construct, and we must consider seasonal aspects while selecting the optimal option. Furthermore, a simple NARX model can handle a wide range of data types, and this study did not look into the impact of GWLM-NARX on different datasets. The findings in this work are necessarily limited, and they do not take into account several other research trends. As a result, testing the same method on many datasets is strongly recommended.

A complete list of abbreviations is shown in *Appendix I.*

6.Conclusion and future work

The GWLM-NARX technique has been proposed using historical-geographical rainfall data from Kashmir province, which includes four explanatory independent variables: maximum temperature, minimum temperature, humidity measured at 12 a.m., and humidity measured at 3 p.m., as well as a target variable indicating the amount of rainfall measured in mm. One of the best fit models built employing (70-

15-15)% training, validation, and test ratios, respectively, had MSE and R of 3.12% and 0.9899% in training, 0.144% and 0.9936% in validation, and 0.311% and 0.9988% in testing. Normal regression and classification can never predict values outside the range of the trained model, and GWLM-NARX confirms the model's efficacy by increasing MSE and R values.

Because the study only looked at fundamental factors like humidity, temperature, and season. We must also evaluate the model's overall performance on variables such as vapor pressure, wind density, and cloud radiation. This study will open the road for the use of hybrid technologies, such as LSTM and NAR models, to examine the overall impact of a dataset with meteorological conditions comparable to those in Kashmir and India's Shimla area. We also need to look at the cross-performance of temperature zones that are opposite to the Kashmir region, such as Rajasthan or India's capital city (New Delhi), after evaluating the model's performance on the same set of data.

Acknowledgment
None.

Conflicts of interest
The authors have no conflicts of interest to declare.

Author's contributions statements
Sheikh Amir Fayaz: Framed the main idea of the work, implementation, interprets the results, data curation, data collection and paper writing. **Majid Zaman:** Study plans with all authors. Provides the basic idea of the work, design and draft of the manuscript and supervision. **Muheet Ahmed Butt:** Study conception and Investigation, testing of the results, editing of the manuscript and co-supervision.

References

[1] Mohamed MA, El AGS, El-mahdy ME. Impact of climate change on rainfall variability in the Blue Nile basin. Alexandria Engineering Journal. 2022; 61(4):3265-75.

[2] Malerba D, Appice A, Bellino A, Ceci M, Pallotta D. Stepwise induction of model trees. In congress of the Italian association for artificial intelligence 2001 (pp. 20-32). Springer, Berlin, Heidelberg.

[3] Zhang S, Lu L, Yu J, Zhou H. Short-term water level prediction using different artificial intelligent models. In international conference on agro-geoinformatics (Agro-Geoinformatics) 2016 (pp. 1-6). IEEE.

[4] Balan MS, Selvan JP, Bisht HR, Gadgil YA, Khaladkar IR, Lomte VM. Rainfall prediction using deep learning on highly non-linear data. International Journal of Research in Engineering, Science and Management. 2019; 2(3):590-2.

[5] Casas DM, González JÁ, Rodríguez JE, Pet JV. Using data-mining for short-term rainfall forecasting. In international work-conference on artificial neural networks 2009 (pp. 487-90). Springer, Berlin, Heidelberg.

[6] Fayaz SA, Zaman M, Butt MA. An application of logistic model tree (LMT) algorithm to ameliorate prediction accuracy of meteorological data. International Journal of Advanced Technology and Engineering Exploration. 2021; 8(84):1424-40.

[7] Mohd R, Butt MA, Baba MZ. GWLM–NARX: Grey Wolf Levenberg–Marquardt-based neural network for rainfall prediction. Data Technologies and Applications. 2020.

[8] Malerba D, Esposito F, Ceci M, Appice A. Top-down induction of model trees with regression and splitting nodes. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2004; 26(5):612-25.

[9] Aswini R, Kamali D, Jayalakshmi S, Rajesh R. Predicting rainfall and forecast weather sensitivity using data mining techniques. International Journal of Pure and Applied Mathematics. 2018; 119(14):843-7.

[10] Hemalatha P. Implementation of data mining techniques for weather report guidance for ships using global positioning system. International Journal of Computational Engineering Research. 2013; 3(3):198-202.

[11] Petre EG. A decision tree for weather prediction. Petrol-Gas University of Ploieşti, Bd. Bucureşti 39, Ploieşti, Department of Informatics. 2009; 61(1):77-82.

[12] Ji SY, Sharma S, Yu B, Jeong DH. Designing a rule-based hourly rainfall prediction model. In international conference on information reuse & integration 2012 (pp. 303-8). IEEE.

[13] Zainudin S, Jasim DS, Bakar AA. Comparative analysis of data mining techniques for Malaysian rainfall prediction. International Journal on Advanced Science, Engineering and Information Technology. 2016; 6(6):1148-53.

[14] Sharaff A, Ukey K, Choure R, Ujee V, Tripathy G. Remodeling rainfall prediction using artificial neural network and machine learning algorithms. In intelligent data engineering and analytics 2022 (pp. 253-9). Springer, Singapore.

[15] Nourani V, Davanlou TA, Molajou A, Gokcekus H. Hybrid wavelet-M5 model tree for rainfall-runoff modeling. Journal of Hydrologic Engineering. 2019; 24(5).

[16] Le VM, Pham BT, Le TT, Ly HB, Le LM. Daily rainfall prediction using nonlinear autoregressive neural network. In micro-electronics and telecommunication engineering 2020 (pp. 213-21). Springer, Singapore.

[17] Adnan RM, Petroselli A, Heddam S, Santos CA, Kisi O. Comparison of different methodologies for rainfall–runoff modeling: machine learning vs conceptual approach. Natural Hazards. 2021; 105(3):2987-3011.

[18] Patil R, Bedekar G. Comparative analysis of machine learning algorithms for rainfall prediction. In innovative data communication technologies and application 2022 (pp. 833-42). Springer, Singapore.
[19] Paudel B, Sarkar N, Chaktraborty S. A comparison of two popular deep learning methods for nowcasting of rainfall. In advances in communication, devices and networking 2022 (pp. 297-305). Springer, Singapore.
[20] Buitrago J, Asfour S. Short-term forecasting of electric loads using nonlinear autoregressive artificial neural networks with exogenous vector inputs. Energies. 2017; 10(1):1-24.
[21] Lingaraju N, Mohan HS. A novel weather parameters prediction scheme and their effects on crops. International Journal of Electrical and Computer Engineering. 2022; 12(1):639-48.
[22] Zhao Y, Meng X, Qi T, Li Y, Chen G, Yue D, et al. AI-based rainfall prediction model for debris flows. Engineering Geology. 2022.
[23] Zaman M, Kaul S, Ahmed M. Analytical comparison between the information gain and Gini index using historical geographical data. International Journal of Advanced Computer Science and Applications. 2020; 11(5):429-40.
[24] Ferreira AA, Ludermir TB, De ARR. Comparing recurrent networks for time-series forecasting. In the international joint conference on neural networks 2012 (pp. 1-8). IEEE.
[25] Burrows K, Marc O, Remy D. Establishing the timings of individual rainfall-triggered landslides using Sentinel-1 satellite radar data. Natural Hazards and Earth System Sciences Discussions. 2022:1-24.
[26] Chen C, Zhang Q, Kashani MH, Jun C, Bateni SM, Band SS, et al. Forecast of rainfall distribution based on fixed sliding window long short-term memory. Engineering Applications of Computational Fluid Mechanics. 2022; 16(1):248-61.
[27] Geethalakshmi V, Kokilavani S, Ramanathan S, Dheebakaran G, Sathyamoorthy N, Maragatham N. Monthly rainfall hind-cast using machine learning algorithms for Coimbatore, Tamil Nadu. MAUSAM. 2022; 73(1):19-26.
[28] Samadianfard S, Mikaeili F, Prasad R. Evaluation of classification and decision trees in predicting daily precipitation occurrences. Water Supply. 2022; 22(4):3879-95.
[29] Lee S, Bae JH, Hong J, Yang D, Panagos P, Borrelli P, et al. Estimation of rainfall erosivity factor in Italy and Switzerland using Bayesian optimization based machine learning models. Catena. 2022.

Sheikh Amir Fayaz is a PhD student at the University of Kashmir in J&K, India. He earned his master's degree in computer science from the University of Kashmir's Department of Computer Science. Data Mining and ML are two of his research areas.
Email: skh.amir88@gmail.com

Dr. Majid Zaman earned his Ph.D. from the University of Kashmir in 2009 and currently works as a Scientist in the University of Kashmir's, Department of Information Technology. India's state of J&K Data Analytics, Data Mining, Data Science, and ML are some of his research interests.
Email: zamanmajid@gmail.com

Dr. Muheet Ahmed Butt obtained his Ph.D. from the University of Kashmir in 2009 and currently works as a Scientist at the University of Kashmir's Department of Computer Science in J&K, India. Data Analytics, Data Mining, Data Science, ML, and Intrusion Detection are among his research interests.
Email: ermuheet@gmail.com

Appendix I

S. No.	Abbreviation	Description
1	ANN	Artificial Neural Network
2	AR	Autoregressive
3	ARIMA	Autoregressive Integrated Moving Average
4	BPNN	Back-Propagation Neural Network
5	CART	Classification and Regression Trees
6	DT	Decision Tree
7	FFNN	Feed-Forward Neural Networks
8	GN	Gauss-Newton
9	GUIDE	Generalized, Unbiased Interaction Detection and Estimation
10	GWLM-NARX	Grey Wolf Levenberg-Marquardt Nonlinear Auto Regression with External Inputs
11	GWO	Grey Wolf Optimization
12	HTL	Hybrid regression Tree Learner
13	ID3	Iterative Dichotomizer 3
14	IMD	Indian Meteorological Department
15	KNN	K-Nearest Neighbour
16	LMT	Logistic Model Trees

17	LSTM	Long Short Term Memory
18	MARS	Multi-Variate Adaptive Regression Splines
19	MATLAB	Matrix Laboratory
20	ML	Machine Learning
21	MLP	Multilayer Perceptron
22	MT	Model Trees
23	NARX	Non Linear Autoregressive Models with Exogenous Inputs
24	NB	Naïve Bayes
25	NDC	National Data Centre
26	NN	Neural Networks
27	OPELM	Optimum Pruning Extreme Learning Machine
28	PCA	Principal Component Analysis
29	RF	Random Forest
30	Rfall	Rainfall
31	RMSE	Root Mean Square Error
32	SECRET	Scalable Linear Regression Tree Algorithm
33	SVM	Support Vector Machine
34	Tmax	Maximum Temperature
35	Tmin	Minimum Temperature