Journal of Engineering Science and Technology Review 16 (1) (2023) 101 - 108

Research Article

JOURNAL OF
Engineering Science and Technology Review

Rainfall Prediction using Machine Learning Techniques for Sabarmati River Basin, Gujarat, India

Anant Patel[1,*], Neha Keriwala[1], Nisarg Soni[2], Unnati Goel[2], Ruchita Bhoj[2], Yakshi Adhyaru[2] and S M Yadav[1]

[1]Sardar Vallabhbhai National Institute of Technology-Surat, SVNIT-Surat
[2]Civil Eng. Dept., Institute of Technology, Nirma University, Ahmedabad, Gujarat, India

Received 11 October 2022; Accepted 25 December 2022

Abstract

Rainfall has a direct effect on agriculture, hydroelectric generation, and water resources management, etc. Many natural catastrophes are also closely linked to rainfall intensity and duration, including flood and drought. Therefore, it is essential to have fast and reliable technique of forecasting rainfall intensity and duration for regional water resource management. Timely rainfall forecasting is also required to avoid and mitigate potential harmful impacts of natural catastrophes such as landslides, floods, and droughts. Rainfall prediction is usually based on numerical weather models combined with meteorological radar data. Such models have been used extensively in studies, including multiple regressions and climatology averaging techniques, numerical methods, and empirical formulations. Forecast accuracy depends on uncertainty, and probabilistic forecasting handles the challenge of unpredictability better than deterministic predictions. The behavior of Random Forest, Gradient Boosting, and Decision tree models has been studied to optimize the results generated from data fed into them. Gradient Boosting was found to work best among those tested, with features related to and affecting rainfall predictability giving an accuracy of 93%. Random forest and Decision tree method having 90% and 78.5% of accuracy was achieved respectively. It was also observed that Mean Absolute Error (MAE) is 1.54, Mean Squared Error (MSE) is 24.94, Root Mean Squared Error (RMSE) is 4.99 for the 40 year time period data. This prediction will be useful for the Meteorological Department, State Disaster Management Department, Water Resources Management Department of State including Dam and Reservoir inflow management.

Keywords: Rainfall, machine learning, prediction, River Sabarmati, Random Forest, Gradient Boosting, Decision Tree classifier

1. Introduction

Predicting rainfall is difficult task due to the complexity of the systems involved [1]. Many topographic, hydrologic, and atmospheric factors influence rainfall, but rainfall prediction is critical in managing surface and subsurface water resources, irrigation projects and hydropower production [2]. The Indian summer monsoon brings a huge amount of rain from June to September, and its unpredictability has a huge effect on the country's economy, making its correct forecasting very important [2,3]. Because of the unpredictability, probabilistic predictions are being used increasingly on both a monthly and seasonal basis.

Statistical forecasting methods commonly provide predictions lacking understanding of uncertainty, but most of those available until recently have also lacked statistical prediction reflecting the potential for change [4,5]. Deterministic predictions provide the likeliest outcomes. Failure to forecast exceptional years has occurred largely because the variability was not captured adequately. Due to India's heavy dependence on agriculture, it is important to depict and convey uncertainty in seasonal forecasts, to help the agricultural agencies make good use of limited resources [6,7]. Decision-makers have greater faith in probabilistic projections, which quantify uncertainty in a way that deterministic forecasts do not.

Extreme rainfall is one of the biggest causes of major natural catastrophes [8,9]. Heavy rainfall events have poor prediction accuracy, making both scientific study and practical forecasting more difficult. In tropical and subtropical coastal Asia, days when 100 mm of rain or more falls occur regularly, and rainfall of 300 mm or more sometimes [10]. Although tropical cyclones are often responsible, many such events arise with monsoon wind patterns. Heavy rainfall, while having large-scale characteristics, is controlled by a number of localized factors [11,12]. It is hard to know whether the most intense rain will fall over the flat lowland, where most people live and with the biggest flood risk [13].

For long-term forecasting, the system must be configured to fulfil all requirements [16,17]. The modeling is based on the assumption that the process of rainfall and stream flow is linear. Nonlinearities are often small and have time-varying parameter trajectories, which reduce their visibility [18,19]. New ideas emerge from the support to the residuals. Introducing non-stationarities (such as shifts or trends) into the historical data is an approach that may be used to distinguish the model from the "future" (scenario) data. Based on these findings, the main characteristics that affect the long-term behavior of the system and highlight significant sub processes may be discovered [21-24]. For weather reports, meteorological agencies use supercomputers operating on both mesoscale and local forecast models.

*E-mail address: anant.patel14@gmail.com
ISSN: 1791-2377
doi:10.25103/jestr.161.13

There are both geographical and temporal effects that come into play when estimating rainfall. These may be particularly important in areas with varied topography, such as the world's mountain ranges [34-37].

The four different geographically distributed rainfall erosive datasets provided by the technique of choice all provide varying degrees of confidence. Analysis revealed that better performance was delivered using the custom model and the SCIA-ISPRA dataset together, with ERA5-Land also outperforming other models, due to major sources of error including improper R-factor estimation, using an empirical model in place of a rigorous model, and using gridded rainfall data instead of point-scale rainfall observations. A critical function in water resources planning and management is making predictions about precipitation [38-42].

1.1. Research Gap

Various research work carried out by the researchers for rainfall prediction across the world and few researchers have carried out the work for Sabarmati River Basin but that was only rainfall runoff modelling work. No researcher has work carried out for rainfall prediction for the Sabarmati River basin. There is various Rainfall runoff model study was carried out using HEC-HMS model and ANN technique. An ANN is a powerful technique that may be used to discover hidden non-linear connections between input and output variables. This included daily rainfall, average temperature, and evaporation data as inputs, and discharge as an output. One researcher carry out the modelling of rainfall and runoff in the Sabarmati River watershed took five years (2001–2005). The neural network was developed, and its simulation results demonstrate an excellent correlation of 0.82 with the observed data. A root mean square error (RMSE) value was 0.11 for the Sabarmati basin's 5-year data was also reported for the aforementioned stations. [25-26, 48]. The limitation of the study was it only focus on Rainfall runoff modeling but author has not carried out the work for Rainfall prediction. Large-scale, highly predictable midlatitude atmospheric patterns are specifically targeted by dynamic numerical weather prediction models. Thermodynamically driven warm-season rainfall events, such afternoon airmass thunderstorm development in subtropical summers, are difficult for these models to reproduce [14]. Ensemble numerical weather prediction simulations are one current approach to this problem; in these, the average of several separate simulations using different model physics is employed as a more accurate forecast than that of any one member.

Although these methods improve results, they still only achieve a moderate level of accuracy. With this in mind, the present study will use machine learning to aggregate information from several ensemble members into a single forecast for the wetter months of the year [15,16]. In order to develop a better prognosis of warm-season rainfall that can be implemented in operational meteorology forecasts, we will develop a support vector machine classification scheme on multiple warm-season rainfall days, using members of a 30-member ensemble as predictors and observed rainfall patterns as a predictand [18-20].

Parameters may be referenced against system signals or represented by a time-based function. In order to find an accurate (nonlinear) system description, these parameter trajectories will be specifically represented and followed to arrive at the final goal of accurately predicting system change. Supercomputers are not accessible widely so neural network-based weather forecasting techniques have been explored using standard computers [26, 27]. To depict error distribution in weather prediction models, the input data may be made artificially noisy so that error can be well-characterized [28, 29]. Model parameterizations may be modified in many ways (e.g. cumulus processes, microphysics, and boundary layer processes), leading to the generation of multiple model solutions, each corresponding to a specific ensemble parameterization scheme.

Better climate risk management may be achieved via the use of probabilistic predictions rather than deterministic ones because of the former's capacity to communicate the forecast's inherent uncertainty. Recent years have seen a rise in interest in the use of probabilistic forecasts based on GCM output. Both parametric (using the Gaussian distribution) and non-parametric (using the counting technique) approaches may be used to construct the probabilistic prediction [30,31]. This research focuses on a non-parametric method for predicting Indian summer monsoon rainfall, which makes no assumptions regarding the shape of the forecast distribution (ISMR). The method of reconstruction included several rain gauges. Choosing the model structure is the most common challenge in black-box modelling.

Methods of combining different forecasts have improved but the ensemble mean continues in use for forecasting. Machine learning tools have opened a new door for weather forecasting to advance and enhance decision-making tools. Until recently, machine learning has not been used to combine precipitation predictions for mountainous areas [32, 33], but such work has started. Climate change makes it important to be able to predict rainfall accurately. This may produce more heavy rainfall (or fewer rainy days) that may bring catastrophes like floods, storms, and landslides. Effective rainfall is rain that feeds the surface system actively.

The main focus of this paper is on how to create long-term data-based models that anticipate the behavior of hydrologic systems and their components. The core belief is that the system will change gradually based on a time-series modelling method [49-52].

It has been shown that better predictions arise if forecasts are made using various techniques that are then combined. The combining method calls for the use of several forecasting and predicting models to provide an overall forecast, which is often preferred over a single one.

Rainfall controls these dispersed systems, which have their own unique features like catchment vegetation or subterranean hydrological processes.

Main objective is to use a combination of machine learning and forecasting methods to predict rainfall. The method needs to predict the total amount of rain that will fall each day, in this case of Sabarmati River, Ahmedabad. This paper shows the behavior of Random Forest, Gradient Boosting, and Decision tree models and also shows the optimized results for given data out of these models.

The primary goal of the work is to obtain a statistically significant improvement of predictive skill over currently utilized ensemble member approaches.

Machine learning tools have opened a new door for weather forecasting to advance and enhance decision-making tool. Using several data-driven models in the area of precipitation prediction has made machine learning more applicable.

This research examines the usefulness of predicting monthly precipitation and also compares a semi-empirical model with machine learning techniques. The stochastic approach shown here provides one-step-ahead prediction characteristics for monthly precipitation estimations.

The algorithms are assessed by how well they forecast monthly precipitation and this is done by using statistics from the connection between weather and topographical conditions. Semi-empirical, matrix factorization methods (NMF), examine how various predictor variables affect precipitation.

2. Study Area

The study area for this research study is the Sabarmati River basin. The river flows year-round, although in the summer it shrinks to a tiny trickle. The three seasons — summer, monsoon, and winter — are marked by different weather patterns. From March to June it is very hot with the temperature varying between 43 and 23°C. Maximum temperatures from November to February range, on average, from 36 to 15°C, and minima from 15 to 5°C. During that time of year, the climate is very dry. January is harsh and cold, with winds from the north [30, 32]. The rainy season begins on 15 June and lasts until 15 September. Average annual rainfall is 832 mm. The record high is 50°C, while the record low is 5°C. Sabarmati is the largest river in Gujarat [29-31].

Ahmedabad city is the most important city in the state of the Gujarat. The city of Ahmedabad, which had a population of about 8.25 million in 2021, is managed by the Ahmedabad Municipal Corporation. The district covers 7,170 km^2.

The urban and peri-urban zones of Ahmedabad include highly populated urban conglomerates, in the west, as well as agricultural settings in the west, industrial districts in the center and east, and a central urban conglomerate. The study area is shown in Figure 1.

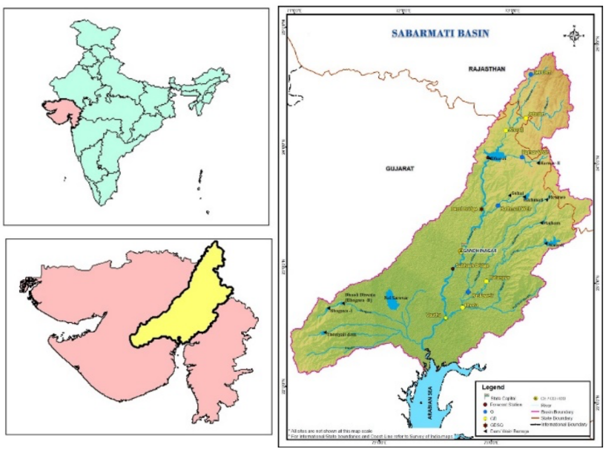

Fig. 1. Study area of Sabarmati River basin

3. Methodology

The Figure 2 depicts the process flow for the rainfall prediction. The first step is the identification of the project, where the project "Rainfall Prediction using Machine Learning techniques for Sabarmati River basin, Gujarat, India" is chosen. The next step involves the research study about the Sabarmati Basin about its location, topology and weather conditions. The next step is to collect the daily rainfall and weather data of past 40 years from January 1982 to June 2021. The data has been collected via two sources, that is, one is satellite-based weather data collection using POWER regional data access from NASA and other involves rainfall data collection using government database WRIS. After the data is collected, it is further preprocessed and compiled and analysis is performed on the data using data visualizations. The next step is to feed the data into three different machine learning models such as Random Forest Regressor, Gradient Boosting Algorithm, Decision Tree Classification. Using different algorithms, the rainfall is predicted and the models are compared based on their performance. After the comparison, the model with best performance is chosen.

3.1. Data Collection & Analysis

The dataset used in the research paper contains daily rainfall data of past 40 years from January 1982 to June 2021. The dataset contains 14426 rows and 12 columns collected from POWER regional data access from NASA's website and Rainfall data is collected from WRIS website for the given dates. Rain is the desired variable. The data set consists of various features such as latitude, longitude, year, month, day, precipitation, relative humidity, surface pressure, frost point, temperature, wind speed and rainfall. – see Table 1. The dataset used contains daily rainfall data from January 1982 to June 2021 – see Table 1.

Exploratory data analysis (EDA) is used to analyze datasets to summarize their main characteristics, and is frequently done with statistical graphics and other data visualization methods. EDA's primary goal is to look beyond the formal modelling or hypothesis testing task. The aim in this study was to analyze each feature and try to understand its variation through time. Graphs represent features such as wind speed, precipitation, relative humidity, surface pressure, frost point, temperature and rainfall – see Figure 3 – and help deduce the weighting and importance of each feature in relation to rainfall prediction. The average rainfall received by Sabarmati basin each year is 736mm. The maximum and minimum rainfall received by the Sabarmati Basin was in the year 2006 with rainfall of 1304mm and in the year 1987 with rainfall of 284mm respectively. The rainfall data has standard deviation of 242mm with mean 736mm.

Fig. 2. Methodology flowchart

Table 1. Dataset description

Features	Description
Latitude	23°49'59.34"N
Longitude	73° 8'20.87"E
Relative humidity	%
Surface Pressure	kPa
Frost Point	°C
Temperature	°C
Wind Speed	m/s
Rainfall	mm (Target variable)

3.2. Random Forest Model

The Random Forest model is a supervised ensemble learning algorithm. Ensemble is a method that combines a number of weak learners to work together to form a single strong predictor. It is a collection of decision trees, known as a forest. New objects are classified and 'voted' for by each tree on the basis of attributes [43, 44]. The classification having the most votes is chosen by the forest. The Working process can be explained in the below steps and diagram:

- Step-1: Random K data points to be selected from the training set.
- Step-2: The decision trees associated with the selected data points (Subsets) is to be built.
- Step-3: The number N for decision trees that you want to build is to be chosen.
- Step-4: Repeat Step 1 & 2.
- Step-5: For new data points, find the predictions of each decision tree, and assign the new data points to the category that wins the majority votes.

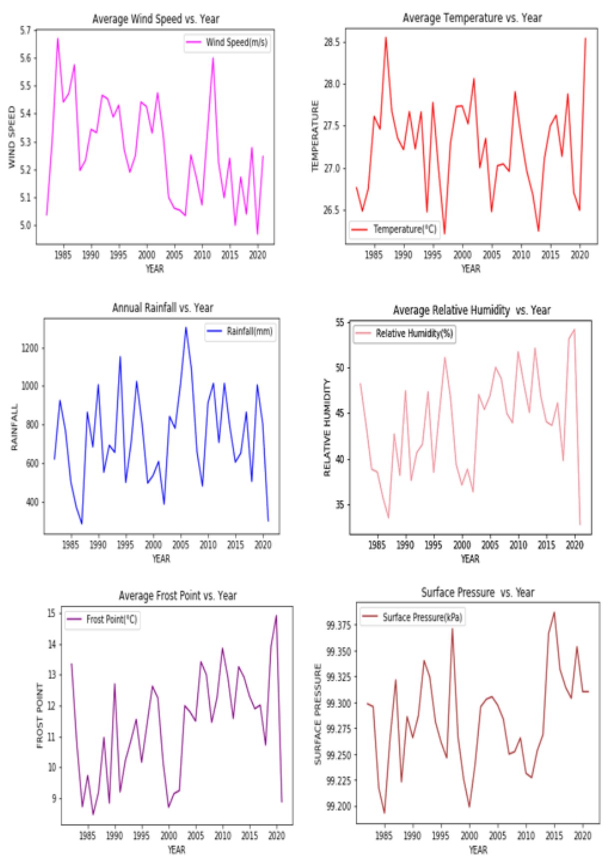

Fig. 3. Wind speed, temperature, rainfall, humidity, Frost point and pressure for the period 1982 to 2021

3.3. Decision Tree Classifier

Decision Tree is a supervised learning technique that can be used for both classification and regression problems, but is generally preferred for the former. The complete process can be better perceived using the below algorithm:

- Step-1: Begin the tree with the root node which contains the total dataset.
- Step-2: Find the best attribute in the dataset utilizing Attribute Selection Measure (ASM).
- Step-3: Divide the root node into subsets that contain potential values for the best attributes.
- Step-4: With the best attribute generate the decision tree node.
- Step-5: Recursively make new decision trees utilizing the subsets of the dataset created in step -3. Precede this process until a stage is reached where you cannot further group the nodes and call the final node as a leaf node.

3.4. Gradient Boosting

In Gradient Boosting, many models are trained sequentially, in a gradual and additive manner, to convert weak learners into strong ones. Trees are trained by assigning equal weight to each observation and modifying each new tree by increasing the weights of difficult observations to classify and decreasing the weights of those that are easier. As steps are repeated, each tree tries to minimize the residual error. Subsequent trees help classify observations not classified by the previous tree. The number of iterations depends on how well the trees fit the model. The final prediction is the sum of the predictions [45-47]. There are few important steps in boosting the algorithm as follows:

- Step-1: A dataset having different data points is considered and initialized.
- Step-2: Now, equal weight to each of the data points is given.
- Step-3: Assume this weight as an input for the model.
- Step-4: The data points that are incorrectly classified are identified.
- Step-5: Increase the weight for data points in step 4.
- Step-6: If you get appropriate output then terminate this process else follow steps 2 and 3 again.

3.5. Performance Evaluators

The performance metrics used below helped understand the performances of the model to better tune the hyper-parameters used in the modelling of the data. The most frequently used metrics for evaluation of binary classification were:

- Precision: the actual values taken as 1 against the total values modelled.

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \tag{1}$$

- Recall: a measure of the model's robustness.

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \tag{2}$$

- F1-Score: the balance between precision and recall, used a measure of the model's accuracy. The value is always between 0 and 1, a higher value signifying better performance.

$$\text{F1-Score} = \frac{2*Precision*Recall}{Precision+Rcall} \quad (3)$$

- Support: the number of samples showing actual occurrences of a class in a dataset.

- Accuracy: a measure of the model's fitness that helps in identifying the features determining the relationships and the patterns between them.

$$\text{Accuracy} = \frac{True\ Positive+True\ Negative}{Total\ Samples} \quad (4)$$

- Mean Absolute Error (MAE): the average difference between the actual and predicted values. Lower MAE signifies a better model but it does not give an idea of the direction of error as different features may have very different range of values.

$$\text{MAE} = \frac{1}{n}\sum_{i=0}^{n}|y_i - \hat{y}_i| \quad (5)$$

where, y_i = actual value, $\hat{y}_i$= predicted value, n= total number of actual values

- Mean Squared Error (MSE): helps to determine the gradient. The square of an error makes it more pronounced, increasing the focus on it.

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (6)$$

where, y_i = actual value, $\hat{y}_i$= predicted value, n= total number of actual values

- Root Mean Squared Error (RMSE): the standard deviation of the residual, which indicates how far the model data points are from the actual values.

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

4. Results & Discussion

To accomplish the objective of the present study, machine learning models were executed such as Random Forest, Gradient Boosting and Decision Tree Classifier and the most efficient technique was selected after comparing the outcomes of all the three models based on the performance indicators to project monsoon streamflow over the Sabarmati River Basin, Gujarat, India. Rainfall intensity is defined as the ratio of the total amount of rain falling during a given period to the duration of the period. Daily rainfall from January 1981 to June 2021 were collected and studied to determine which factor has the greatest influence. The factors considered were wind speed 50 meters above ground level (m/s), precipitation (mm/day), relative humidity 2 meters above ground level (%), surface pressure (kPa), and frost point (°C) and temperature at 2 meters (°C).

4.1. Random Forest results

Random Forest generates multiple decision trees, which are then combined to make a more accurate prediction. A random forest regression algorithm with 1,000 decision trees was used along with classification to predict the future rainfall trends. The training and test dataset ratio was kept at 70:30 as, according to the empirical studies the best results are obtained. The full rainfall data dataset was used in classification and, on that basis, the possibility of rainfall was predicted. As per the results achieved using Random forest method, 90% of accuracy was achieved as shown in - Table2.

Table 2. Random Forest classification scores

	Precision	Recall	F1-score	support
0	0.91	0.95	0.93	2933
1	0.88	0.81	0.84	1395
Accuracy			0.90	4328
Macro Average	0.90	0.88	0.89	4328
Weighted Average	0.90	0.90	0.90	4328

The above table shows the classification scores of the random forest model along with the acquired accuracy. A random forest regression algorithm was used to predict rainfall and 1,000 decision trees used.

Random Forest regression based on time-series forecasting was also used to predict future trends using past data. The results achieved are reported in Table 3.

Table 3. Expected vs. Predicted Values for next 30 days by Time Series Forecasting

Expected	Predicted	Expected	Predicted	Expected	Predicted	Expected	Predicted
0.0	3.0	2.3	6.1	9.8	0.5	4.5	0.1
0.1	0.1	0.0	3.4	1.7	10.3	0.5	3.2
1.0	0.1	0.0	1.6	0.7	2.3	0.6	0.0
0.1	3.4	0.0	0.6	6.6	1.9	12.1	1.2
0.0	0.8	0.0	0.1	2.7	9.6	3.7	10.6
0.0	0.0	0.0	0.5	6.8	30.8	0.5	6.5
0.0	0.2	0.0	0.1	13.8	11.2		
0.0	0.2	0.0	0.0	14.1	15.9		

4.1.1. Error Analysis

MAE, MSE/RMSE: Error analysis is performed to identify the outliers in the data and to fit them so as to adjust with the model. It was also observed that Mean Absolute Error (MAE) is 1.54, Mean Squared Error (MSE) is 24.94, Root Mean Squared Error (RMSE) is 4.99 for the 40 year time period data.

4.2 Decision tree classifier results

When the Decision Tree Classifier was used for rainfall prediction, the training/validation ratio was 70:30 again, and the full dataset (1982 to 2021) was used. Many parameters were involved, as noted at the head of the results section.
Using the Decision Tree Classifier, the model resulted in an accuracy of 78.53%.

The actual and. predicted values for some entries using the Decision Tree classification algorithm are shown in Table

4.The accuracy achieved in the Decision Tree Classifier was 78.53%.

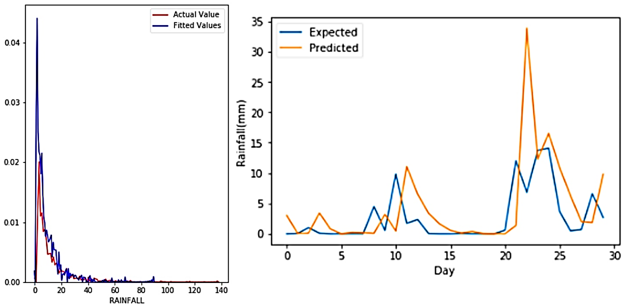

Fig. 4. (a) Actual vs. Fitted Values (b) Expected vs. Predicted Values

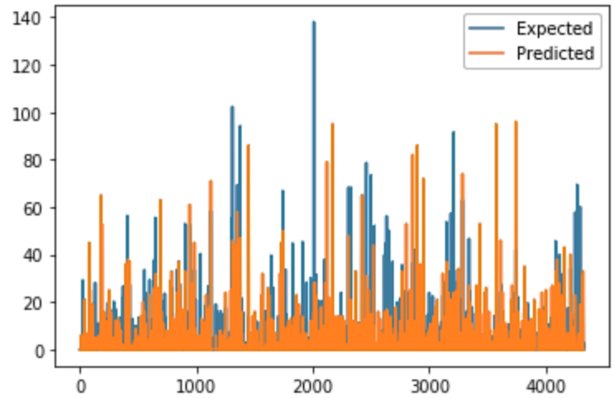

Fig. 5. Expected vs Predicted Rainfall using decision tree classifier

Table 4. Actual and predicted values from the Decision Tree classifier

Date	Actual	Predicted	Date	Actual	Predicted
31-01-1986	0.02	0	02-01-1999	0	0
30-08-1986	7.63	6	05-09-1999	10.14	9
06-02-1995	2.86	3	03-03-2003	0	0
18-06-1995	2.73	4	21-06-2006	2.26	3
09-09-1996	7.86	6	09-05-2007	1.24	2
16-08-1997	29.35	23	26-11-2009	0	0
19-05-1998	0	0	07-07-2020	0.56	1

4.3 Gradient Boosting

Training and validation for Gradient Boosting classification was done using the same 70:30 ratio as for the other models. The Gradient Boosting model achieved an accuracy of 92% – Table 5. The actual and predicted values for some entries with Gradient Boosting classification are shown in Table 6:

Table 5. Gradient Boosting accuracy

	Precision	Recall	F1-score	Support
0	0.91	0.95	0.93	2933
1	0.88	0.80	0.84	1395
Accuracy			0.90	4328
Macro average	0.89	0.87	0.88	4328
Weighted average	0.90	0.90	0.90	4328

Table 6. Actual vs Predicted values for Gradient Boosting classification

Date	Actual	Predicted	Date	Actual	Predicted
1st Jan, 1982	0	0	10th Jan, 1982	1	0
5th Jan, 1982	0	0	12th Jan, 1982	1	1
22nd Mar, 1999	0	0	28th Apr, 1995	1	1
19th Mar, 2003	0	0	22nd Jul, 2007	1	1
21st Nov, 2004	0	0	7th Jun, 2021	1	1

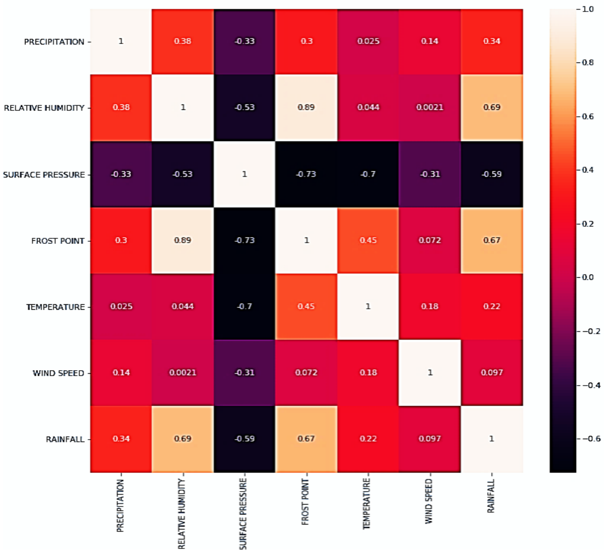

Fig. 6. Correlation matrix of the correlation coefficients for various parameters for rainfall prediction

5. Concluding remarks

A machine learning method was described in this article to forecast the following day's accumulated precipitation. The method predicts the total amount of rain that will fall each day, in this case in Ahmedabad. The new method's performance has been compared to that of other approaches. Many people have expressed concern about rainfall forecasts, as rainfall impacts on so many human activities. As it is the most closely related variable to natural occurrences like landslides, floods, etc, prediction is critical. If prediction is reliable, it enables mitigation and prevention actions to be taken. Organizations managing catastrophe mitigation may use forecasts based on meteorological time-series to aid decision-making. The method presented in this paper is based on machine learning and used to forecast the following day's total precipitation. It provides better results than previous designs.

The project's goal was to use a combination of machine learning and forecasting methods to predict rainfall. Even though actual rainfall depends on numerous factors, high levels of classification accuracy can be achieved using just a few. Even with eight distinct rainfall categorization groups, the study has shown that accuracy is still respectable. Random Forest, Gradient Boosting, and Decision Tree models have been studied, and attempts made to optimize them for dataset used. This showed that Gradient Boosting works best of the three, with features related to and affecting rainfall predictability giving an accuracy of 93%. Random forest and Decision tree method having 90% and 78.5% of accuracy was achieved respectively. It was also observed that Mean Absolute Error (MAE) is 1.54, Mean Squared Error (MSE) is 24.94, Root Mean Squared Error (RMSE) is 4.99 for the 40 year time period data. This prediction will be useful for the Meteoritical Department, State Disaster Management Department, Water Resources Management Department of State including Dam and Reservoir inflow management.

References

1. Abhishek, K., Kumar, A., Ranjan, R., & Kumar, S. A “Rainfall Prediction Model using Artificial Neural Network” (ICSGRC), 2012, pp.82–87.
2. Acharya, N., Mohanty, U. C., & Sahoo, L. “Probabilistic multi-model ensemble prediction of Indian summer monsoon rainfall using general circulation models: A non-parametric approach”. *Comptes Rendus - Geoscience*, 345(3) 2013, pp.126–135. https://doi.org/10.1016/j.crte.2013.01.006
3. Bilgili, M., & Sahin, B. “Prediction of long-term monthly temperature and rainfall in Turkey”. *Energy Sources, Part A: Recovery, Utilization and Environmental Effects*, 32(1) 2010, pp.60–71. https://doi.org/10.1080/15567030802467522
4. Cowan, T., Stone, R., Wheeler, M. C., & Griffiths, M. “Improving the seasonal prediction of Northern Australian rainfall onset to help with grazing management decisions”. *Climate Services*, 19(September), 2020. https://doi.org/10.1016/j.cliser.2020.100182
5. Fernando, D. N., Chakraborty, S., Fu, R., & Mace, R. E.. “A process-based statistical seasonal prediction of May–July rainfall anomalies over Texas and the Southern Great Plains of the United States”. *Climate Services*, 16(August), 2019, 100133. https://doi.org/10.1016/j.cliser.2019.100133
6. George, J., Letha, J., & Jairaj, P. G. “Daily Rainfall Prediction Using Generalized Linear Bivariate Model – A Case Study”. *Procedia Technology*, 24, 2016, pp. 31–38. https://doi.org/10.1016/j.protcy.2016.05.006
7. Haq, D. Z., Rini Novitasari, D. C., Hamid, A., Ulinnuha, N., Arnita, Farida, Y., Nugraheni, R. D., Nariswari, R., Ilham, Rohayani, H., Pramulya, R., & Widjayanto, A. “Long Short-Term Memory Algorithm for Rainfall Prediction Based on El-Nino and IOD Data”. *Procedia Computer Science*, 179, 2019, pp.829–837. https://doi.org/10.1016/j.procs.2021.01.071
8. Hasanah, Y., Herlina, M., & Zaikarina, H. “Flood Prediction using Transfer Function Model of Rainfall and Water Discharge Approach in Katulampa Dam”. *Procedia Environmental Sciences*, 17, 2013, pp.317–326. https://doi.org/10.1016/j.proenv.2013.02.044
9. He, J., Qiu, H., Qu, F., Hu, S., Yang, D., Shen, Y., Zhang, Y., Sun, H., & Cao, M. Corrigendum to “Prediction of spatiotemporal stability and rainfall threshold of shallow landslides using the TRIGRS and Scoops3D models” 2021, 197, https://doi.org/10.1016/j.catena.2020.105074
10. Hern, E. “Rainfall Prediction : A Deep Learning Approach”, 2016, pp.151–162. https://doi.org/10.1007/978-3-319-32034-2
11. Jaydeep Sagathia, Kotecha, N., Patel, H., & Patel, A. “Impact Assessment of Urban Flood in Surat City using HEC-HMS and GIS” *SSRN, Elsevier*, 2020, pp.1–9. https://ssrn.com/abstract=3558360
12. Kajewska-Szkudlarek, J. “Clustering approach to urban rainfall time series prediction with support vector regression model”. *Urban Water Journal*, 17(3), 2020, pp.235–246. https://doi.org/10.1080/1573062X.2020.1760319
13. Karimnezhad, A., Niazi, S., & Parsian, A. “Bayes and robust Bayes prediction with an application to a rainfall prediction problem”. *Journal of the Korean Statistical Society*, 43(2), 2014, pp.275–291. https://doi.org/10.1016/j.jkss.2013.10.001
14. Kashiwao, T., Nakayama, K., Ando, S., Ikeda, K., Lee, M., & Bahadori, A. “A neural network-based local rainfall prediction system using meteorological data on the Internet: A case study using data from the Japan Meteorological Agency”. *Applied Soft Computing Journal*, 56, 2017, pp.317–330. https://doi.org/10.1016/j.asoc.2017.03.015
15. Keesman, K., & Jakeman, T. “Identification for Long-Term Prediction of Rainfall-Streamflow Systems”. *IFAC Proceedings*, 30(11), 1997, pp.1445–1450. https://doi.org/10.1016/s1474-6670(17)43046-1
16. Khalil, A. B. J. A. B. “Short-term SPI drought forecasting in the Awash River Basin in Ethiopia using wavelet transforms and machine learning methods”. *Sustainable Water Resources Management*, *2*(1), 2016, pp.87–101. https://doi.org/10.1007/s40899-015-0040-5
17. Lansford, J. W., Walsh, T. H., Hromadka, T. V., & Rao, P. “Doppler radar rainfall prediction and gauge data” *BMC Research Notes,* 13(1), 2020, pp.10–12. https://doi.org/10.1186/s13104-020-05311-y
18. Lu, T., Yamada, T., & Yamada, T. “Fundamental Study of Real-time Short-term Rainfall Prediction System in Watershed: Case Study of Kinu Watershed in Japan”. *Procedia Engineering*, 154,2016, pp.88–93. https://doi.org/10.1016/j.proeng.2016.07.423
19. Mandal, T., & Jothiprakash, V. “Short-term rainfall prediction using ann and mt techniques”. *ISH Journal of Hydraulic Engineering,* 18(1), 2012, pp.20–26. https://doi.org/10.1080/09715010.2012.661629
20. Mannan, M. A., Chowdhury, M. A. M., & Karmakar, S. “Application of NWP model in prediction of heavy rainfall in Bangladesh”. *Procedia Engineering*, 56, 2013, pp.667–675. https://doi.org/10.1016/j.proeng.2013.03.176
21. Mercer, A., Dyer, J., & Zhang, S. “Warm-season thermodynamically-driven rainfall prediction with support vector machines”. *Procedia Computer Science*, 20, 2013, pp.128–133. https://doi.org/10.1016/j.procs.2013.09.250
22. Mislan, Haviluddin, Hardwinarto, S., Sumaryono, & Aipassa, M. “Rainfall Monthly Prediction Based on Artificial Neural Network: A Case Study in Tenggarong Station, East Kalimantan – Indonesia”. *Procedia Computer Science*, 59, 2015, pp.142–151. https://doi.org/10.1016/j.procs.2015.07.528
23. Nonki, R. M., Lenouo, A., Tshimanga, R. M., Donfack, F. C., & Tchawoua, C. “Performance assessment and uncertainty prediction of a daily time-step HBV-Light rainfall-runoff model for the Upper Benue River Basin, Northern Cameroon”. *Journal of Hydrology: Regional Studies*, 36(May), 2021, 100849. https://doi.org/10.1016/j.ejrh.2021.100849
24. Omogbai, B. E. “Prediction of Northern Nigeria Rainfall Using Sea Surface Temperature”. *Journal of Human Ecology*, 32(2), 2010, pp.127–133. https://doi.org/10.1080/09709274.2010.11906331
25. Osman, Y. Z., & Abdellatif, M. E. “Improving accuracy of downscaling rainfall by combining predictions of different statistical downscale models”. *Water Science*, 30(2), 2016, pp.61–75. https://doi.org/10.1016/j.wsj.2016.10.002
26. Padulano, R., Rianna, G., & Santini, M. “Datasets and approaches for the estimation of rainfall erosivity over Italy: A comprehensive comparison study and a new method”. *Journal of Hydrology: Regional Studies*, 34(November 2020), 2021, 100788. https://doi.org/10.1016/j.ejrh.2021.100788
27. Pandya, U., Patel, A., & Patel, D. P. “Application of Open Source Google Image For River Bathymetry Delineation For 1D Hydrodynamic”, *6*(1), 2019, pp.1051–1058.
28. Patel, A. and Yadav, S.M. “Stream flow prediction using TIGGE ensemble precipitation forecast data for Sabarmati river basin”. *Water Supply*, 22(11), 2022, pp.8317-8336.

29. Patel, A. "Design of Optimum Number of Rain Gauge Network over Sabarmati river basin" *I-Manager's Journal on Future Engineering & Technology*, *14*(1), 2018, pp.26–31.
30. Patel, A. "Rainfall – Runoff modelling and simulation using Remote Sensing and Hydrological model for Banas River" in *Lecture Notes in Civil Engineering, Springer, Singapore* (p. 7) 2018.
31. Patel, A., Panchal, U. and Keriwala, N., "A Novel Plan for Gujarat to Mitigate the Effect of Flood, Drought and Salinity Using Interlinking of Canal and Rivers". *In Environmental Restoration: Proceedings of F-EIR Conference 2021, Cham: Springer International Publishing*. 2022, *pp. 155-167.*
32. Patel, A. D., & Patel D., Prakash I., Dholakiya M. "Analysis of Optimum Number of Rain Gauge in Shetrunji River Basin, Gujarat – India", *2*(11), 2016, pp.380–384.
33. Patel, A. D., Patel, D., & Prakash, I. "Flood Modelling Using HEC-RAS And Geo - Informatics Technology Flood Modelling Using Hec-Ras And Geo - Informatics Technology In Lower Reaches Of Shetrunji River, Gujarat, India", (October) 2016.
34. Patel, P., Bablani, S., & Patel, A. "Impact assessment of urbanization and industrialization using Water Quality Index on Sabaramati River, Ahmedabad", 2020, pp.1–7.
35. Persiano, S., Ferri, E., Antolini, G., Domeneghetti, A., Pavan, V., & Castellarin, A. "Changes in seasonality and magnitude of sub-daily rainfall extremes in Emilia-Romagna (Italy) and potential influence on regional rainfall frequency estimation". *Journal of Hydrology: Regional Studies,* 32(July), 2020, 100751. https://doi.org/10.1016/j.ejrh.2020.100751
36. Praveen, B., Talukdar, S., Mahato, S., & Mondal, J. "Analyzing trend and forecasting of rainfall changes in India using non- parametrical and machine learning approaches", 2020, pp.1–21. https://doi.org/10.1038/s41598-020-67228-7
37. Rachmawati, R. N. "Estimation of Extreme Rainfall Patterns Using Generalized Linear Mixed Model for Spatio-temporal data in West Java, Indonesia". Procedia Computer Science, 179, 2020, pp.330–336. https://doi.org/10.1016/j.procs.2021.01.013
38. Rachmawati, R. N., Sungkawa, I., & Rahayu, A. "Extreme rainfall prediction using Bayesian quantile regression in statistical downscaling modeling". *Procedia Computer Science,* 157, 2019, pp.406–413. https://doi.org/10.1016/j.procs.2019.08.232
39. Richard, M., & Rao, K. G. "Artificial neural networks in temporal and spatial variability studies and prediction of rainfall". *ISH Journal of Hydraulic Engineering,* 20(1), 2014, pp.1–6. https://doi.org/10.1080/09715010.2013.806400
40. Ridwan, W. M., Sapitang, M., Aziz, A., Faizal, K., Najah, A., & El-shafie, A. "Rainfall forecasting model using machine learning methods : Case study". *Ain Shams Engineering Journal*, 2020. https://doi.org/10.1016/j.asej.2020.09.011
41. Roman, U. C., Patel, P. L., & Porey, P. D. "Prediction of missing rainfall data using conventional and artificial neural network techniques". *ISH Journal of Hydraulic Engineering*, 18(3), 2012, pp.224–231. https://doi.org/10.1080/09715010.2012.721660
42. Sahai, A. K., Soman, M. K., & Satyan, V. "All India summer monsoon rainfall prediction using an artificial neural network". *Climate Dynamics*, 16(4), 2000, pp.291–302. https://doi.org/10.1007/s003820050328
43. Shah, U., Garg, S., Sisodiya, N., Dube, N., & Sharma, S. "Rainfall Prediction : Accuracy Enhancement Using Machine Learning and Forecasting". *Fifth International Conference on Parallel, Distributed and Grid Computing (PDGC)*, 2018, pp.776–782.
44. Sharma, A., & Goyal, M. K. "Bayesian network for monthly rainfall forecast: A comparison of K2 and MCMC algorithm". I*nternational Journal of Computers and Applications,* 38(4), 2016, pp.199–206. https://doi.org/10.1080/1206212X.2016.1237131
45. Singh, A., Acharya, N., Mohanty, U. C., & Mishra, G. "Performance of Multi Model Canonical Correlation Analysis (MMCCA) for prediction of Indian summer monsoon rainfall using GCMs output". *Comptes Rendus - Geoscience*, 345(2), 2013, pp.62–72. https://doi.org/10.1016/j.crte.2012.11.003
46. Suparta, W., & Samah, A. A. "Rainfall prediction by using ANFIS times series technique in South Tangerang, Indonesia". *Geodesy and Geodynamics*, 11(6), 2020, pp.411–417. https://doi.org/10.1016/j.geog.2020.08.001
47. Thai, B., Minh, L., Le, T., Bui, K. T., & Minh, V. "Development of advanced artificial intelligence models for daily rainfall prediction". *Atmospheric Research*, *237*, 2019, 104845. https://doi.org/10.1016/j.atmosres.2020.104845
48. Vathsala, H., & Koolagudi, S. G. "Closed Item-Set Mining for Prediction of Indian Summer Monsoon Rainfall A Data Mining Model with Land and Ocean Variables as Predictors". *Procedia Computer Science*, 54, 2015, pp.271–280. https://doi.org/10.1016/j.procs.2015.06.032
49. Verma, R. "ANN-based Rainfall-Runoff Model and Its Performance Evaluation of Sabarmati River Basin, Gujarat, India". *Water Conserv Sci Eng*. 2022. https://doi.org/10.1007/s41101-022-00160-1
50. Wang, N., Lombardo, L., Gariano, S. L., Cheng, W., Liu, C., Xiong, J., & Wang, R. "Using satellite rainfall products to assess the triggering conditions for hydro-morphological processes in different geomorphological settings in China". *International Journal of Applied Earth Observation and Geoinformation*, 102, 2021, 102350. https://doi.org/10.1016/j.jag.2021.102350
51. Xian-Mei, L. "An Effective Approach for Improving the Real-Time Prediction of Summer Rainfall over China". *Atmospheric and Oceanic Science Letters,* 4(2), 2011, pp.75–80. https://doi.org/10.1080/16742834.2011.11446907
52. Yu, W., Nakakita, E., & Jung, K. "Flood Forecast and Early Warning with High-Resolution Ensemble Rainfall from Numerical Weather Prediction Model". *Procedia Engineering*, 154, 2016, pp.498–503. https://doi.org/10.1016/j.proeng.2016.07.544