Vol. 24, No. 1, pp. 1-8 (2024)
Journal of Agricultural Physics
ISSN 0973-032X
http://www.agrophysics.in

Research Article

Ensemble Machine Learning Techniques for Prediction of Rainfall in India

ANKIT KUMAR SINGH[1], ANITA SARKAR[1], RANJIT KUMAR PAUL[2]*, MD YEASIN[2], H.S. ROY[2], PRAKASH KUMAR[2] AND A.K. PAUL[2]

[1]*The Graduate School, ICAR-Indian Agricultural Research Institute, New Delhi-110012*
[2]*ICAR-Indian Agricultural Statistics Research Institute, New Delhi-110012*

ABSTRACT

Rainfall plays a crucial role in various sectors, particularly agriculture, where accurate forecasting is essential for optimizing crop yields, managing irrigation, and mitigating the impacts of extreme weather events. This study proposed a novel ensemble machine learning approach to predict rainfall in the Central and Eastern coastal regions of India, integrating the strengths of multiple predictive models. Traditional methods, such as stochastic models and autoregressive techniques, often struggle in capturing extreme rainfall events and complex underlying patterns in time series data. The machine learning (ML), including support vector machines (SVR), random forest (RF), and artificial neural networks (ANN), have been effective in addressing these challenges. The present study integrates these machine learning models into an ensemble learning (EL) framework to enhance forecasting accuracy. By combining the forecasts of these models, the ensemble method aims to improve accuracy and robustness in rainfall prediction. The dataset include monthly rainfall data from ten sub-divisions of central and eastern coastal region in India. The study evaluates the performance of the proposed EL model by comparing their predictions with candidate models viz., RF, ANN, and SVR. The results demonstrate that the proposed EL model offers improvement in accuracy in rainfall prediction, compared to candidate ML models. The results reveal the importance of advanced forecasting techniques and suggest that ensemble machine learning models could offer more reliable prediction of rainfall.

Key words: ANN, SVR, RF, Ensemble, PSO, Rainfall prediction

Introduction

Rainfall holds immense importance across diverse sectors, playing a critical role in shaping agricultural practices (Trinh, 2018). It is crucial for cultivating crops like rice, tea, and several fruits and vegetables that are important to the agriculture based economy. Knowing about rain before it happens is really important for many things. For farmers, rainfall forecasts are invaluable in planning crop activities, optimizing yields, and managing irrigation effectively (Abbot and Marohasy, 2017). Several research studies have demonstrated the importance of rainfall for agriculture, and emphasised the effects of climate change on future rainfall patterns (Krishna *et al*., 2004; Gilmont *et al*., 2018; Dharani, 2022). Water resource authorities depend on these forecasts to manage reservoirs effectively, helping to avoid both droughts and floods. Regular flooding can lead to loss of life, property damage, and interruptions to livelihoods. Additionally, accurate rainfall forecasts contribute to disaster preparedness, enabling early

*Corresponding author,
Email: ranjitstat@gmail.com

warnings for potential floods, landslides, or droughts (Bezak *et al*., 2016). Several recent studies have investigated the variations and trends in rainfall patterns across various regions using both parametric and non-parametric approaches (Pal and Al-Tabbaa, 2011). The amounts of rainfall estimation is important in managing agricultural productivity, water resources, public health, infrastructure, and tourism. Therefore, there is need to develop and improve the methods of rainfall prediction to ensure proper planning and management.

Various models are used for predicting rainfall, based on the evidence that the process being studied is a stationary one and relies on independent variables to account for past rainfall events (Pejman and Latif, 2014). Many researchers have proposed and explored the model for predicting rainfall and compared it with other popular forecasting models. Rajeevan *et al*. (2007) have used multiple linear regression model for the prediction of rainfall in India using six relevant predictors based on the experimental data. Vashisth *et al*. (2008) investigated the accuracy of weather forecast of Delhi. Paul (2017) reported significant presence of long memory in minimum and maximum temperature of India. Some of the previous studies explored statistical models including autoregressive integrated moving average (ARIMA), and seasonal autoregressive integrated moving average (SARIMA) model for prediction of climatic variables (Nugroho, and Simanjuntak, 2014; Geetha, and Nasira, 2016; Paul and Anjoy, 2018; Swain *et al*., 2018). Paul and Birthal (2016) used wavelet technique for investigating rainfall trend in India. Paul *et al*. (2020) applied combination of wavelet and ARIMA model for forecasting sub-divisional rainfall in India. Saha *et al*. (2021) assessed the accuracy of district level rainfall in Mizoram. However, the shortcomings of stochastic models such as their failure to account for extreme values and other complexities in rainfall dynamics suggest the use of machine learning or more advanced forecasting models. In recent years, machine-learning models have become increasingly important, demonstrating a remarkable ability to model and simulate the complex and nonlinear patterns found in hydrometeorological time series (Prathibha *et al*., 2023).

Machine learning models offer a more dynamic approach to modelling rainfall data, effectively capturing extreme values. The machine learning methods include Support Vector Machines (SVR), Random Forest (RF), Artificial Neural Networks (ANN), and K-Nearest Neighbour (KNN) models etc. The effectiveness of these algorithms may vary significantly. Numerous machine learning models have been used for rainfall modelling, as highlighted in previous studies. Among them, ANN and SVR are the most commonly employed (Refonaa *et al*., 2019; Laskar and Pushpa, 2023; Ehteram *et al*., 2023). RF is a decision tree-based model that addresses the overfitting problems associated with single decision trees while maintaining high prediction accuracy. RF offers ease of use and excellent computational speed (Schoppa *et al*., 2020). However, both traditional stochastic time series models and ML models often struggle to capture the complex behaviors of these data, making it difficult to rely on a single model for accurate prediction of rainfall. There is no model which could perform equally efficient in all model. A model may perform better in one kind of dataset but may perform poor in other kind of dataset. The performance of model may depends on the underlying pattern and nature of the dataset. In response to this challenge, ensemble machine learning approaches have become essential. The ensemble approach combines the prediction of different candidate models to make improvement in accuracy of forecast. Sani *et al*. (2020) proposed an ensemble learning to improve the rainfall prediction. Ensemble learning approach combined rainfall prediction of several ML model, which include Naïve Bayes, Decision Tree, Support Vector Machine, Random Forest and Neural Network based models. Some of recent work on rainfall prediction like Gu *et al*. (2022) and Danandeh *et al*. (2021) also used ensemble approach to improve forecast accuracy. Attaf *et al*. (2024) proposed ensemble method to enhance performance of ML model for estimate of rainfall. Yeasin *et al*. (2024) investigated the ensemble approach for prediction of tropical cyclones in north India.

This article presents a novel ensemble approach to predict rainfall in Central and Eastern coastal region of India. The ML models i.e., RF, ANN and SVR model are used in this study to forecast time series. The forecast obtained from ML models are combined using fixed weight ensemble methods. The comparison of the proposed ensembled model has been made with candidate ML models. The detailed outline of the paper is mentioned as follows. Section 2 presents the introduction to the candidate ML models and ensemble approached employed in this study. Section 3 presents the empirical studies conducted. Section 4 discusses the results and findings from these studies. Finally, Section 5 presents the conclusions drawn from the research.

Methodology

Candidate model

This study focuses on predicting the amount of rainfall for central and eastern coastal region of India. This section outlines various well-known machine learning models that are considered as candidate models within an ensemble approach.

Artificial Neural Network (ANN)

An Artificial Neural Network (ANN) is a ML model, which simulates the human brain, designed to tackle complex challenges in scientific research. Fundamentally, an ANN comprises neurons arranged in three primary layers: the input layer, the hidden layer, and the output layer. The input layer takes in various parameters, the output layer generates predicted results, and the hidden layer serves as a middle ground, identifying non-linear relationships between the inputs and outputs. The ANN model incorporates a time series sequence ranging from y_{t-1} to y_{t-n} as input where n represents the number of lagged observations. To model time series data, create a nonlinear function, denoted as f, and the mathematical formulation is given by the equation (1).

$$y_t = w_0 + \sum_{j=1}^{h} w_j f(w_{0j} + \sum_{i=1}^{n} w_{ij} y_{t-i}) + e_t \quad \ldots(1)$$

In equation (1), w_{ij}, and w_j represent the weights to model, h represents number of hidden nodes, and e_t represents an error term. The activation function for the hidden layers in the ANN model can take various forms, such as sigmoid and Rectified Linear Unit (ReLU).

Support Vector Regression (SVR)

Support Vector Regression (SVR) effectively manages non-linear relationships between input variables and the target variable by using kernel functions to transform the data into a higher-dimensional space. This capability of SVR make it suitable for regression tasks that involve intricate associations between input variables and the target variable. The estimated function of the dataset, i.e., $h(y)$ in equation (5) is the output from the model. The process involves applying a kernel function $D(.)$ to relocate non-linear data to higher-dimensional feature space, treating this transformed data as if it were linear, and then computing a dot product with a weight vector represented as 'w.' This dot product is combined with a bias term 'b' to produce the final estimation. The SVR model can be formulated as in equation (2).

$$h(y, \boldsymbol{w}) = h(y, \alpha, \alpha^*) = \sum_i (\alpha_i - \alpha_i^*) D(y, y_i) + b \quad \ldots(2)$$

where, $\boldsymbol{y}$ contains sequence of time series from y_{t-1} to y_{t-n} where n is the number of lag. RBF kernel is most popularly used kernel function in SVR model. The hyper-parameters used in the model is tuned using Lagrange method. α, α^* mentioned in the Equation (2) denote the Lagrange multipliers, and should satisfy the equality: $\alpha_j \alpha_j^* = 0$.

Random Forest (RF)

Random Forest (RF) is an ensemble learning method based on decision trees that has become one of the most popular algorithms in machine learning. Its wide usage can be attributed to its effectiveness across various applications. In classification tasks, a decision tree acts as a simple model, with internal nodes representing tests on attributes, branches showing the outcomes of these tests, and leaves containing class labels. Decision trees can also be utilized for regression tasks when the target variable is continuous. RF employs a technique known as bootstrap aggregation, or bagging, in which each decision tree is trained using a randomly selected subsample from the overall training dataset. The forecasted value ($\hat{y}_t$) is determined by the following equation (3):

$$\hat{y}_t = r(y_{t-1}, y_{t-2}, \dots, y_{t-n}) \ , t = n+1, \dots, k. \quad \dots(3)$$

where $r(.)$ be the represent the function derived from the RF model for predicting y_t based on the time series observations employing n lagged variables

Ensemble learning (EL) model

Ensemble Learning (EL) models are a widely used method in forecasting that involve merging the predictions from multiple individual models to enhance overall accuracy and robustness. These approaches are based on the idea that various models may capture different features of the hidden patterns in data, and on combining their predictions, they may yield more accurate and dependable forecasts. The Unweighted ensemble approach gives equal weightage to each candidate model. Here, suppose $\widehat{y_1}$, $\widehat{y_2}$, … $\widehat{y_N}$, are the forecasted value obtained from N number of models respectively. Then unweighted ensemble forecast ($\widehat{y_{uw}}$) will be given by equation (4):

$$\widehat{y_{uw}} = \frac{1}{N}\left(\sum_{i=1}^{N} \widehat{y_i}\right) \quad \dots(4)$$

whereas weights of candidate model in weighted ensemble approach is different but fixed and determined using optimization algorithm. The predictions from the individual models are aggregated using assigned weights. So, the forecast from weighted ensemble method ($\widehat{y_{fw}}$) will be given by equation (5):

$$\widehat{y_{fw}} = \left(\sum_{i=1}^{N} w_i \,\widehat{y_i}\right) \quad \dots(5)$$

where the w_i's are the weight associated to jth candidate model which are determined using particle swarm optimization (PSO) such that $\sum_{j=1}^{N} w_j = 1$. The PSO is used to optimise the weights, which offers the benefit of increasing the chances of identifying the global optimum while reducing the chance of entrapment in local optima (Yeasin and Paul, 2024). Fig. 1 represents the detailed architecture of the proposed EL model based on PSO weight optimisation.

Empirical study

Data description

This study utilises the monthly rainfall data from central and eastern coastal region of India, (specifically 10 subdivisions), has been obtained

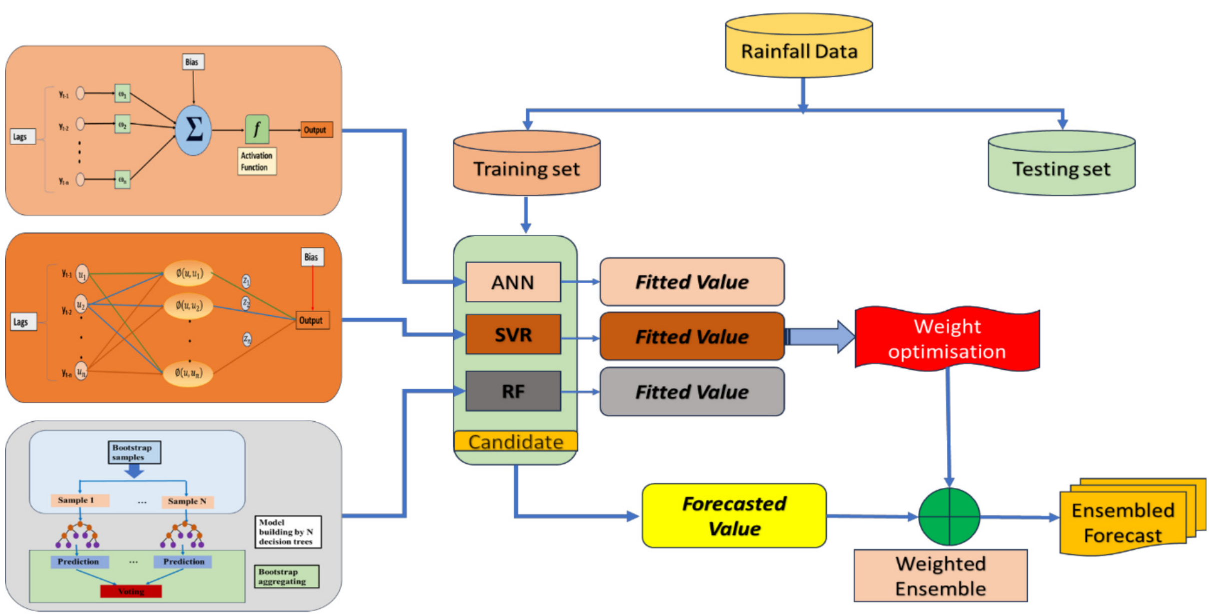

Fig. 1. The detailed flowchart of proposed EL model

from the Indian Institute of Tropical Meteorology (www.tropmet.res.in) in Pune, India. The central region of India plays a crucial role in the agricultural economy, while the eastern coast is highly vulnerable to extreme weather events, such as cyclones. The central regions encompass subdivisions like Bihar (BH), Chhattisgarh (CHH), East Uttar Pradesh (EUP), East Madhya Pradesh (EMP), West Madhya Pradesh (WMP), Vidharba (VDH), and Telangana (TEL). The eastern coastal region includes Coastal Andhra Pradesh (CoAP), Tamil Nadu (TN), and Odisha (OD). The dataset includes monthly rainfall measurements from 1871 to 2021 recorded in millimetres (mm). The dataset comprises 1812 data points, and the series is partitioned into training and testing sets with 80:20 ratio. The training set encompasses the initial 1452 observation, utilized for model fitting, while the remaining 360 of observations in the test set are reserved for validation purpose. The table 1 provides the descriptive statistics of the above mentioned dataset.

Table 1 shows that the coefficient of variation (CV) of selected sub-division ranges from 95.42% to 150.93%. The prediction of rainfall data is not easy due to its complex pattern. The minimum value of rainfall in all the selected subdivisions are 0 mm whereas maximum value ranges from 449.60 mm to 742.40 mm. The positive value of skewness describe that the data corresponding to all the 10 subdivision are positively skewed.

Model implementation

The ML model viz., ANN, SVR, and RF are employed on the given dataset as discussed in section 2. The ReLU activation function in ANN model and the RBF kernel function in SVR model has been used. All the hyper-parameters have been properly fine-tuned toget the optimum value. The fitted value corresponding to first 12 observations are lost as lag number for ML model are taken as 12. The forecast value obtained from the ML models are combined using weighted ensemble approach. The weights of ensemble approach are fixed and determined using PSO weight optimization as mentioned in section 2.1.4. The accuracy metrices mentioned in section 3.3 have been computed for candidate model as well as ensemble approach and the obtained results corresponding to all the candidate models and ensemble approach for each of the sub division.

Accuracy metrices

The performance of EL models are validated using accuracy metrices particularly root mean square error (RMSE), mean absolute error (MAE), and root relative squared error (RRSE) which are formulated as in equation (6), (7) and (8 respectively:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \qquad \ldots(6)$$

Table 1. Descriptive statistics of the subdivision wise rainfall of the central region and the eastern coastal region in India

	Central Region							Eastern Coastal Region		
	BH	CHH	EMP	EUP	TEL	VDH	WMP	OD	TN	CoAP
mean	101.61	112.63	103.54	83.80	74.71	90.53	79.90	122.69	77.29	81.86
Std. dev.	129.98	151.20	148.42	119.08	91.54	121.86	120.58	141.01	73.76	88.09
Median	31.40	26.90	21.70	18.15	28.00	20.55	10.80	53.20	55.65	53.70
Minimum	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Maximum	594.90	697.00	742.40	584.40	510.70	596.90	564.20	688.00	449.60	476.20
Skewness	1.29	1.31	1.53	1.50	1.31	1.37	1.57	1.13	1.57	1.30
Kurtosis	0.67	0.53	1.33	1.26	1.05	0.89	1.48	0.29	2.77	1.59
SE	3.05	3.55	3.49	2.80	2.15	2.86	2.83	3.31	1.73	2.07
CV (%)	127.92	134.24	143.34	142.11	122.52	134.60	150.93	114.93	95.42	107.61

$$\text{MAE} = \frac{1}{N}\sum_{\text{i}=1}^{N}|y_i - \hat{y}_i| \quad \ldots(7)$$

$$RRSE = \sqrt{\frac{\sum_{i=1}^{N}\left(y_i - \hat{y}_i\right)^2}{\sum_{i=1}^{N}\left(y_i - \bar{y}\right)^2}} \quad \ldots(8)$$

Where, y_i is the actual value, $\hat{y}_i$ is the forecasted value, N is the total number of observation.

The obtained value of accuracy metrices are also computed for the candidate models and the performance of EL model has been compared with the candidate models.

Results and Discussion

The value corresponding to accuracy measures of all the candidate ML models for predicting amount of rainfall of different subdivision are mentioned in table 2. The results obtained from table 2 shows that among ML model, the SVR Model is performing better in all the 10 sub division in terms of lower value of RMSE and RRSE however it has higher value in terms of MAE compared to RF model. The higher value of MAE may be due to presence of few larger value of error for SVR model. These differences suggest that no single model can be accurate at all data points. The performance of model may vary significantly in different hidden pattern in the dataset. These limitations in performance of candidate models provide strong evidence to utilise ensemble approach to improve the prediction on rainfall. The proposed EL model is the weighted ensemble approach where weights are determined using PSO optimisation algorithm and it is fixed. The proposed EL model has lowest value of RMSE, MAE, and RRSE among all the candidate models for BH, CHH, CoAP, EUP and VDH subdivision, which shows that proposed EL model significantly outperformed the ML models in terms of lower value of accuracy measures. In case of EMP, OD, TN, TEL, and WMP sub division, the proposed EL model has lowest RMSE and RRSE value but the MAE values are relatively higher than the RF model however the difference between MAE values of proposed EL model and RF model is significantly less than that of SVR and RF model. These results shows that

Table 2. Accuracy measures of different models

Region	Model	RMSE	MAE	RRSE	Region	Model	RMSE	MAE	RRSE
BH	ANN	61.74	40.84	0.49	OD	ANN	67.61	44.69	0.48
	RF	11.37	6.13	0.09		RF	22.84	9.23	0.16
	SVR	8.29	6.29	0.07		SVR	16.28	9.97	0.12
	Ensemble	7.04	4.91	0.06		Ensemble	15.33	9.37	0.11
CHH	ANN	55.01	34.04	0.39	TN	ANN	51.09	34.29	0.65
	RF	12.49	5.79	0.09		RF	15.09	6.17	0.19
	SVR	9.21	7.11	0.07		SVR	12.90	6.79	0.16
	Ensemble	8.25	5.75	0.06		Ensemble	11.62	6.52	0.15
CoAP	ANN	68.14	46.76	0.77	TEL	ANN	52.26	33.81	0.58
	RF	12.94	6.16	0.15		RF	12.32	5.07	0.14
	SVR	12.63	7.29	0.14		SVR	9.09	6.07	0.10
	Ensemble	10.34	5.96	0.12		Ensemble	8.74	5.26	0.10
EMP	ANN	64.61	38.51	0.46	VDH	ANN	56.93	34.51	0.48
	RF	14.05	6.75	0.10		RF	13.25	5.84	0.11
	SVR	10.93	8.46	0.08		SVR	8.26	5.95	0.07
	Ensemble	9.20	6.94	0.07		Ensemble	7.82	5.31	0.07
EUP	ANN	50.23	31.48	0.49	WMP	ANN	57.83	33.73	0.47
	RF	8.98	4.70	0.09		RF	12.72	5.98	0.10
	SVR	6.63	5.41	0.07		SVR	10.04	7.28	0.08
	Ensemble	5.23	3.73	0.05		Ensemble	8.97	6.28	0.07

proposed EL model has improved the forecast accuracy in all the 10 subdivision considered in this study.

Conclusion

This study proposed a novel EL model approach based on PSO optimisation for forecasting of amount of rainfall in central and eastern coastal region of India. The result obtained from this study provide evidence regarding effectiveness of proposed EL model in prediction of amount of rainfall compared to commonly used ML model viz., ANN, SVR, and RF model. As proposed EL model based on weighted ensemble approach harness the strength of the ANN, SVR, and RF model by assigning appropriate weights to the ML models. The PSO technique is used for weight optimisation. The result of empirical study carried out on rainfall data collected for 10 sub divisions of India have reached the conclusion that the proposed EL model outperformed the ML model in terms of lower value of at least two out of three accuracy measures. The accuracy of EL model depends on the performance of candidate models. Therefore, the selection of candidate model should be rational. Use of different optimization algorithms may improve the performance of proposed EL model.

References

Abbot, J. and Marohasy, J. 2017. Skilful rainfall forecasts from artificial neural networks with long duration series and single-month optimization. *Atmospheric Research* **197**: 289-299.

Attaf, Y., Lazri, M., Labadi, K., Mohia, Y., Ouallouche, F. and Absi, R. 2024. Enhancement of ANN performance for remote sensing rainfall estimate in northern Algeria using ensemble learning methods. *Journal of Earth System Science* **133**: 92.

Bezak, N., Šraj, M. and Mikoš, M. 2016. Copula-based IDF curves and empirical rainfall thresholds for flash floods and rainfall-induced landslides. *Journal of Hydrology* **541**: 272-284.

Danandeh Mehr, A. 2021. Seasonal rainfall hindcasting using ensemble multi-stage genetic programming. *Theoretical and Applied Climatology* **143**: 461-472.

Dharani, C., Maragatham, N., Geethalakshmi, V., Ramanathan, S.P. and Kannan, B. 2022. Rainfall variability analysis and trend assessment in Theni District of Tamil Nadu, India: an implication for crop planning. *Journal of Agrometeorology* **24**: 305-308.

Ehteram, M., Ahmed, A.N., Sheikh Khozani, Z. and El-Shafie, A. 2023. Convolutional neural network-support vector machine model-gaussian process regression: a new machine model for predicting monthly and daily rainfall. *Water Resources Management* **37**: 3631-3655.

Geetha, A. and Nasira, G.M. 2016. Time-series modelling and forecasting: Modelling of rainfall prediction using ARIMA model. *International Journal of Society Systems Science* **8**: 361-372.

Gilmont, M., Hall, J.W., Grey, D., Dadson, S.J., Abele, S. and Simpson, M. 2018. Analysis of the relationship between rainfall and economic growth in Indian states. *Global Environmental Change* **49**: 56-72.

Gu, J., Liu, S., Zhou, Z., Chalov, S.R. and Zhuang, Q. 2022. A stacking ensemble learning model for monthly rainfall prediction in the Taihu Basin, China. *Water* **14**: 492.

Krishna Kumar, K., Kumar, R.K., Ashrit, R.G., Deshpande, N.R. and Hansen, J.W. 2004. Climate impacts on Indian agriculture. *International Journal of Climatology* **24**: 1375-1393.

Laskar, S.I., Kotal, S.D. and Bhowmik, S.R. 2014. Analysis of rainfall and temperature trends of selected stations over North East India during last century. *Mausam* **65**: 497-508.

Niksaz, P., and Latif, A.M. 2014. Rainfall events evaluation using adaptive neural-fuzzy inference system. *Journal of Information Technology and Computer Science* **9**: 46-51.

Nugroho, A. and Simanjuntak, B.H. 2014. ARMA (Autoregressive Moving Average) model for prediction of rainfall in regency of Semarang-Central Java-Republic of Indonesia. *International Journal of Computer Science Issues (IJCSI)* **11**: 27.

Paul, R.K. and Birthal, P.S. 2016. Investigating rainfall trend over India using the wavelet technique. *Journal of Water and Climate Change* **7**: 353-364.

Paul, R.K. 2017. Modelling long memory in maximum and minimum temperature series in India. *Mausam* **68**: 317-326.

Paul, R.K. and Anjoy, P. 2018. Modeling fractionally integrated maximum temperature series in India in presence of structural break. *Theoretical and Applied Climatology* **134**: 241-249.

Paul, R.K., Paul, A.K. and Bhar, L.M. 2020. Wavelet-based combination approach for modeling sub-divisional rainfall in India. *Theoretical and Applied Climatology* **139**: 949-963.

Pal, I. and Al-Tabbaa, A. 2011. Assessing seasonal precipitation trends in India using parametric and non-parametric statistical techniques. *Theoretical and Applied Climatology* **103**: 1-11.

Prathibha, K., Rithvik Reddy, G., Kosre, H., Lohith Kumar, K., Rajak, A. and Tripathi, R. 2023. Rainfall Prediction Using Machine Learning. *Machine Intelligence Techniques for Data Analysis and Signal Processing: Proceedings of the 4th International Conference MISP 2022,* **1**: 457-468.

Rajeevan, M., Pai, D.S., Anil Kumar, R. and Lal, B. 2007. New statistical models for long-range forecasting of southwest monsoon rainfall over India. *Climate Dynamics* **28**: 813-828.

Refonaa, J., Lakshmi, M., Dhamodaran, S., Teja, S. and Pradeep, T.N.M. 2019. Machine learning techniques for rainfall prediction using neural network. *Journal of Computational and Theoretical Nanoscience* **16**: 3319-3323.

Saha, S., Das, B., Chakraborty, D., Chawdhury, S., Lalliansanga, S. and Saithantluanga, H. 2021. Quantitative accuracy assessment of district level rainfall forecast in Mizoram. *Journal of Agricultural Physics* **21**, 332-340.

Sani, N.S., Abd Rahman, A.H., Adam, A., Shlash, I. and Aliff, M. 2020. Ensemble learning for rainfall prediction. *International Journal of Advanced Computer Science and Applications* **11**.

Schoppa, L., Disse, M. and Bachmair, S. 2020. Evaluating the performance of random forest for large-scale flood discharge simulation. *Journal of Hydrology* **590**: 125531.

Swain, S., Nandi, S. and Patel, P. 2018. Development of an ARIMA model for monthly rainfall forecasting over Khordha district, Odisha, India. *Recent Findings in Intelligent Computing Techniques* 325-331.

Trinh, T.A. 2018. The impact of climate change on agriculture: findings from households in Vietnam. *Environmental and Resource Economics* **71**: 897-921.

Vashisth, A., Das, D.K., Bhagawati, G. and Sharma, P.K. 2008. Accuracy of weather forecast for semiarid climate of Delhi for Agricultural management practices. *Journal of Agricultural Physics* **8**: 51-58.

Yeasin, M. and Paul, R.K. 2024. Opti Semble Forecasting: optimization-based ensemble forecasting using MCS algorithm and PCA-based error index. *The Journal of Supercomputing* **80**: 1568-1597.

Yeasin, M., Paul, R.K. and Shankar, S.V. 2024. Ensemble machine learning models for forecasting tropical cyclones in North Indian region. *Earth Science Informatics* **17**: 3705-3714.

Received: 2 March 2024; Accepted: 30 May 2024