

On the performance analysis of rainfall prediction using mutual information with artificial neural network

Shilpa Hudnurkar, Neela Rayavarapu

Department of Electronics and Telecommunication Engineering, Symbiosis Institute of Technology,
Symbiosis International (Deemed University), Pune, India

Article Info

Article history:

Received Feb 20, 2022

Revised Sep 26, 2022

Accepted Oct 20, 2022

Keywords:

Artificial neural network

Climate variables

Feature selection

Mutual information

Rainfall prediction

ABSTRACT

Monsoon rainfall prediction over a small geographic region is indeed a challenging task. This paper uses monthly means of climate variables, namely air temperature (AT), sea surface temperature (SST), and sea level pressure (SLP) over the globe, to predict monthly and seasonal summer monsoon rainfall over the state of Maharashtra, India. Mutual information correlates the temperature and pressure from a grid of 10° longitude X 10° latitude with Maharashtra's monthly rainfall time series. Based on the correlations, selected features over the respective latitude and longitudes are given as inputs to an artificial neural network. It was observed that AT and SLP could predict monthly monsoon rainfall with excellent accuracy. The performance of the test dataset was evaluated through mean absolute error; root mean square error, correlation coefficient, Nash Sutcliffe model efficiency coefficient, and maximum rainfall prediction capability of the network. The individual climate variable model for AT performed better in all evaluation parameters except maximum rainfall capability, where the combined model 2 with AT, SLP and SST as predictors outperformed. The SLP-only model's performance was comparable to the AT-only model. The combined model 1 with AT and SLP as predictors was found better than the combined model 2.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Shilpa Hudnurkar

Department of Electronics and Telecommunication Engineering, Symbiosis Institute of Technology,

Symbiosis International (Deemed University)

Symbiosis Knowledge Village, Pune-412115, India

Email: shilpa.hudnurkar@sitpune.edu.in

1. INTRODUCTION

The monsoon, an essential and spectacular phenomenon, brings rainfall every year to tropical countries. This rainfall is vital for activities such as farming, water resource management, and power generation. The economy of several countries is dependent on agriculture, and India is one such country. Timely and sufficient rainfall is required mainly for farming-related activities on the rainfed land. India receives most of its annual rainfall during the summer season. This is referred to as summer monsoon rainfall (SMR) or Southwest monsoon rainfall (as it is received because of the Southwest monsoon). The rainfall distribution is not uniform as it depends on the topography of the region, its orography. Hence, based on the rainfall's homogeneity, India is divided into 36 subdivisions as shown in Figure 1. Four of the 36 subdivisions are in Maharashtra; hence, this region was selected for the study. Further, Maharashtra has over 60% of rainfed agricultural land. Thus, SMR is essential for this state.

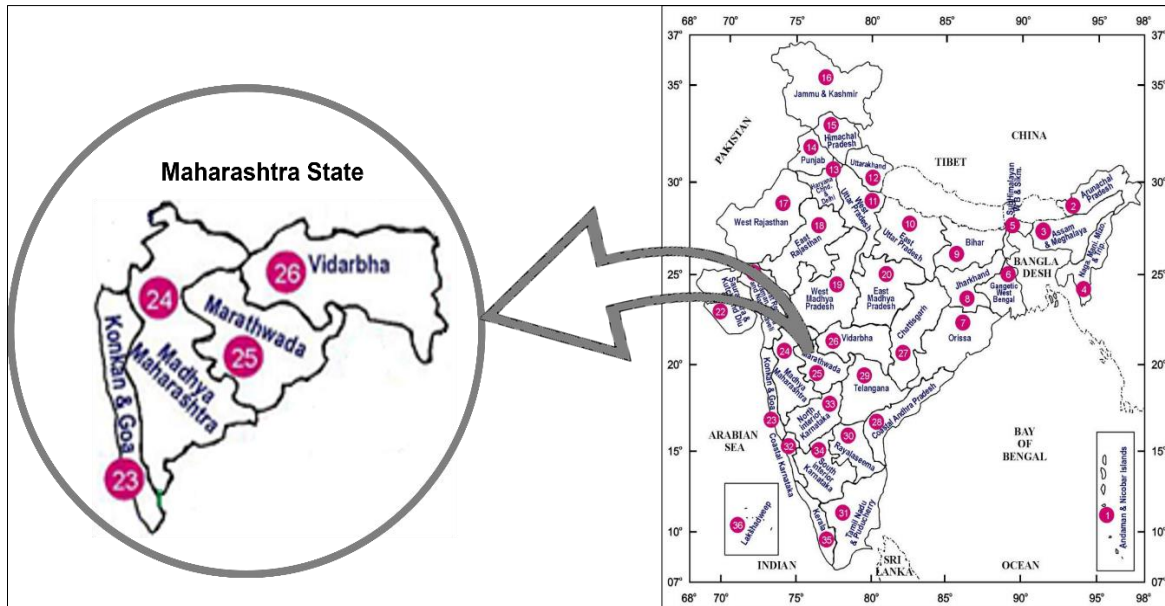


Figure 1. Map showing the 36 meteorological sub-divisions with the study area encircled on the left [1]

The amount of rainfall received during the summer monsoon months, i.e., June to September, is not the same for every region. It shows interannual variability highlighted in earlier studies [2], [3]. Studies show that many factors contribute to this variability [3]–[5]. These factors are important for predicting rainfall and are mainly related to temperature and pressure anomalies over different latitudes and longitudes. To identify the factors that contribute to the interannual variability and the factors that have a significant correlation with rainfall at a particular geographic location, the use of artificial intelligence (AI) tools have been done. While numerous studies use AI for rainfall prediction [6]–[12], limited studies use AI for correlating climate variables with rainfall [13].

In a different application of massive time series forecasting of PM_{2.5}, Jin *et al.* [14] used self-screening layer to remove highly correlated inputs. Various clustering algorithms have been described in the literature, such as support vector machine (SVM), k-nearest neighbor (KNN), and kernel fisher linear discriminant analysis [15]. These are unsupervised algorithms. One other unsupervised machine learning algorithm used for clustering is the autoencoder. Basha *et al.* [12] have used an unsupervised technique, autoencoder, to identify rainfall predictors. The authors have used climate variables to predict the Indian summer monsoon rainfall (ISMR). The authors studied climate variables over a grid of latitudes and longitudes of 10° X 20°. They used a regression tree with a bagging algorithm for the seasonal prediction of ISMR.

In the clustering approach, clusters are formed based on specific criteria, and later correlation of the target variable with the clusters must be performed to enable better prediction. On the other hand, a supervised approach leads to a set of predictors that can be fed to the prediction model. Supervised feature selection algorithms identify predictors based on the correlation of the variables, under study, with the predictand. Filtering, wrapping, and embedded technique are among the most widely used supervised techniques for feature selection (or predictor identification). The filter method provides speed and flexibility and has a limited computational cost when a large dataset is to be dealt with [16]. It also has good generality with no dependency on the prediction algorithm. Mutual information (MI) is a filtering technique used for high-volume data. MI can reflect nonlinearity in the data and remains almost unaffected in the presence of outliers [17].

MI has been utilized in various applications for feature selection [18]–[20]. He *et al.* [21] used MI to identify predictors for monthly rainfall prediction. The authors employed ten climate indices, including El Niño–Southern Oscillation (ENSO) indices, Indian Ocean Dipole, and a hybrid prediction model to predict rainfall over Australia. Kim *et al.* [17] used MI to determine the correlation between various rain gauge stations.

In this study, we use air temperature (AT), sea level pressure (SLP), and sea surface temperature (SST) over the globe to predict rainfall over the state of Maharashtra, India. Our significant contributions to this study are as: i) development of a hybrid model for long-range rainfall prediction over a state; ii) use of a

supervised technique to correlate the global climate variables with monthly rainfall to identify the rainfall predictors; iii) identification of rainfall predictors over a smaller grid size of 10° latitude X 10° longitude; iv) despite the high rainfall variability over monthly time scales, the model achieved a very good Nash Sutcliffe model efficiency coefficient (R^2 -score) [22] and accuracy; v) the performance of the prediction model was gauged by mean absolute error (MAE), root mean square error (RMSE), R-value, R^2 -score and the maximum rainfall prediction capability of the network; and vi) the performance of the prediction model was gauged by MAE, RMSE, Pearson correlation coefficient (R-value), R^2 -score, and the maximum rainfall prediction capability of the network.

This paper is organized into six sections, where section 2 details the methods used for carrying out the research. Section 3 describes the supervised correlation technique, namely mutual information, and section 4 provides a brief overview of artificial neural network (ANN). Experimental work is detailed in section 5, discussing the results. Section 6 concludes the paper and briefly covers the future scope of the study.

2. METHOD

The data required for this study is the time series of monthly rainfall over Maharashtra and the monthly mean of climate variables AT, SLP and SST. The climate variable data at the National Atmospheric and Oceanic Administration (NOAA), USA website were freely available. It is a reanalysis dataset for the monthly mean of the respective climate variable prepared over 2.5° latitude X 2.5° longitude over the globe. The dataset of the three climate variables, AT, SLP, and SST, were downloaded. The data available were multidimensional in network common data form (netCDF) format. The four dimensions for AT were latitude, longitude, level of the atmosphere, and time (month). For the SLP and SST, data were available on three dimensions, excluding the level of the atmosphere. Each climate variable was further extracted in Microsoft Excel using the Panoply open-source software developed at the National Aeronautics and Space Administration (NASA) Goddard Institute for Space Studies. The monthly mean values for AT and SLP were available from January 1948 to December 2020 and for SST till February 2020. Each dataset extracted in Excel was individually pre-processed. The monthly rainfall time series over Maharashtra was prepared by adding monthly rainfall over all the subdivisions of Maharashtra, namely, Konkan-Goa, Madhya Maharashtra, Marathwada, and Vidarbha. The dataset of monthly rainfall over subdivisions of India was made available by the Indian Institute of Tropical Meteorology and is freely downloadable [23] These datasets were used for feature selection using a supervised method of filtering. Selected features were fed to the ANN to predict monthly rainfall. The broader view of methodology is shown in Figure 2.

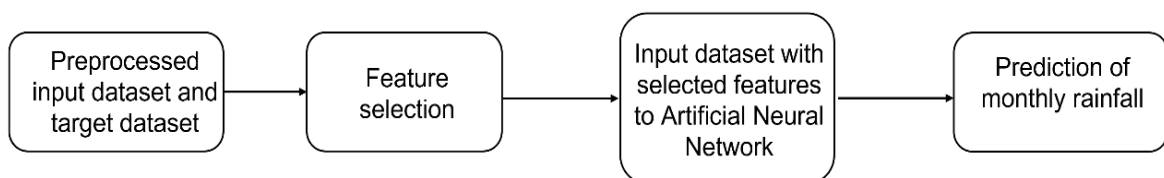


Figure 2. A broad view of the methodology

Climate variables were pre-processed by following the steps given: i) averaging the variable over the grid of 10° X 10° of latitude and longitude, ii) preparing time series of the climate variable over each pair of latitude and longitude, iii) calculating the anomaly by arranging the time series for each month of the year from 1948 to 2020, and iv) preparing the final anomaly time series dataset for each pair of latitude and longitude for all the years.

After pre-processing, MI was employed for feature selection to identify significant predictors. The output of the filtering technique enabled the identification of latitude-longitude-specific predictors of the climate variable. This was carried out for all the three climate variables and lagged correlation. The selected features for respective climate variables were plotted on the world map to understand the geographical distribution of the predictors. The dataset was prepared for the selected features to be used to train and test the prediction model developed using ANN. Three separate networks were trained and tested for the prediction of monthly rainfall. The performance of these networks was evaluated based on MAE, RMSE, R-value, R^2 -score [22], and maximum rainfall prediction capability of the network. The detailed methodology is represented in Figure 3.

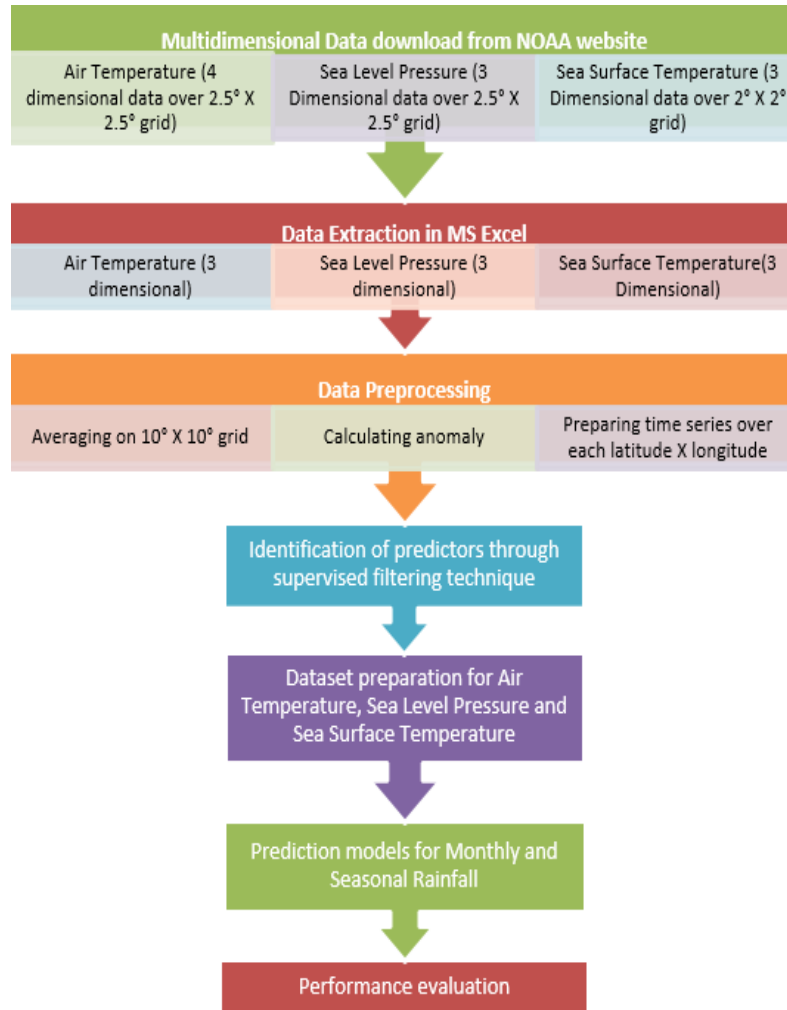


Figure 3. Methodology of the work

3. MUTUAL INFORMATION

Feature selection is essential when large features are present in the dataset [24]. Various feature selection methods to select a subset of features are available such as filter, wrapper [25], and embedded [26]. Feature selection enables noise reduction and helps to remove irrelevant features [26]. The filter method selects features based on the statistical characteristics of the dataset and is efficient as it does not depend on any learning algorithm [27]. MI is a measure of dependence between two variables based on Shannon's entropy [28]. The formula in (1) is Shannon's entropy giving the uncertainty of Y.

$$H(Y) = -\sum_y p(y)\log(p(y)) \quad (1)$$

Conditional entropy is given by (2), where the uncertainty of Y gets reduced by observing X [29]. The reduction in uncertainty is given by (3), implying that X has some correlation with Y. If $I(Y, X)$ has a value greater than zero it indicates that X can provide information about Y [30].

$$H(Y|X) = -\sum_x \sum_y p(x, y)\log(p(y|x)) \quad (2)$$

$$I(Y, X) = H(Y) - H(Y|X) \quad (3)$$

Here, the information gain does not depend on the classifier applied; hence, performance can be varied [31]. In other words, it can be said that it ranks the features in an unbiased manner, and hence any machine learning algorithm can be used later for prediction or classification. Feature selection techniques, including MI, have been reviewed in [32], [33].

4. ARTIFICIAL NEURAL NETWORK

The literature has widely used ANN for classification and regression problems [34]. It processes information the same way as the neuron in the human brain. Hence, each node of the ANN is called a neuron. An ANN consists of input and output neurons or nodes. These are called the input layer and output layer. Another layer, called the hidden layer, also consists of neurons. A simple architecture of ANN with one hidden layer is shown in Figure 4, where x is the input and $x=1, 2, 3, \dots, n$; n is the neuron or node and $n=1, 2, 3, \dots, m$; and O is the output. The number of outputs depends on the application.

Each neuron in the hidden layer has a processing element. The processing element multiplies each input by a certain weight value and sums all inputs. The weight value depends on the significance to be given to that input. The activation function decides the final output of the neuron, which is passed to the neuron/s connected to it in the next layer. During the training phase, the supervised ANN compares the final output of the ANN with the expected result to calculate the error. This error is minimized by using various learning algorithms. The ANN is trained until it achieves minimum error. This trained network consists of fixed weights and biases for each layer. This network is used to test unseen data [35].

The architecture of the ANN depends on how the layers are interconnected. The number of layers, number of neurons, activation functions, and learning rate are some of the parameters that must be selected. The number of hidden layers can be increased if the problem is complex. Various activation functions such as hyperbolic tangent (TanH), Sigmoidal, and rectified linear unit (ReLU) can be used in the neurons [36].

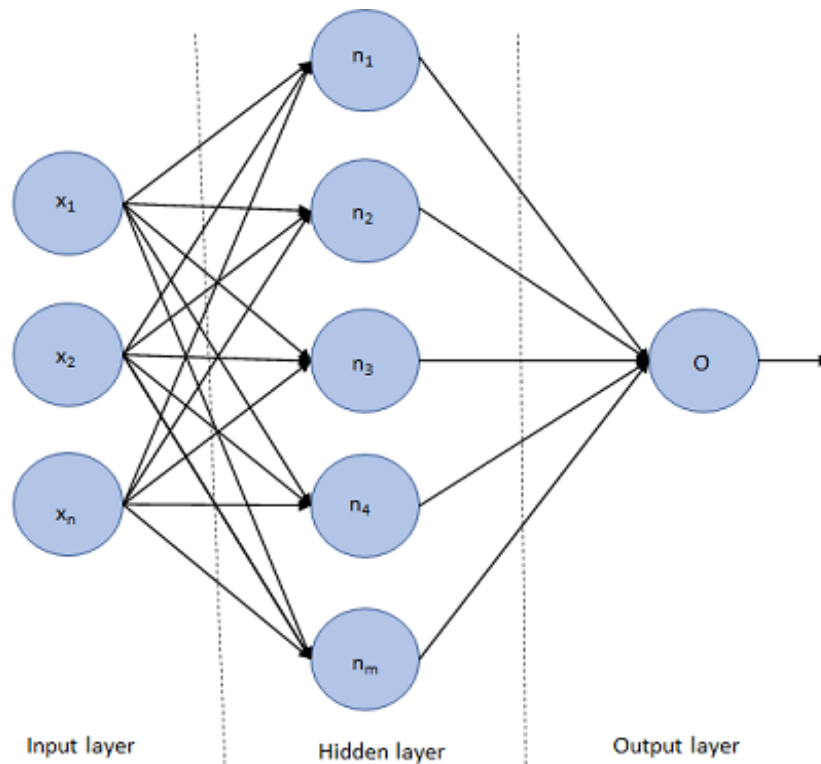


Figure 4. A simple architecture of artificial neural network

5. EXPERIMENTAL RESULTS AND DISCUSSION

As the change in temperature and pressure play a significant role in the formation of rainfall, AT, SLP, and SST over the globe were considered the potential predictors. The Panoply software was used to visualize the data. Representative pictures of the AT (May 2020), SLP (May 2020), and SST (May 2019) are shown in Figures 5, 6, and 7, respectively. AT and SLP are NCEP-NCAR Reanalysis 1 datasets, and SST is NOAA Extended reconstructed SST V3b dataset. References for AT, SLP, and SST are [29]–[37], respectively.

The monthly mean for AT and SLP is available on the grid of 2.5° latitude X 2.5° longitude, and SST is available on the grid of 2° latitude X 2° longitude. The AT and SLP datapoints are for 73 latitudes, 144 longitudes, and 876 months. SST datapoints are over 89 latitudes, 180 longitudes, and monthly data from 1,854 to January 2020.

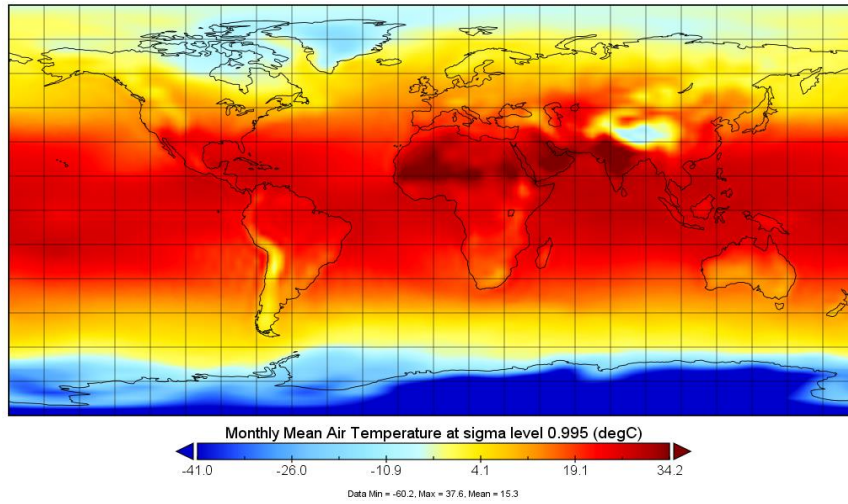


Figure 5. Monthly mean air temperature over the globe of May 2020

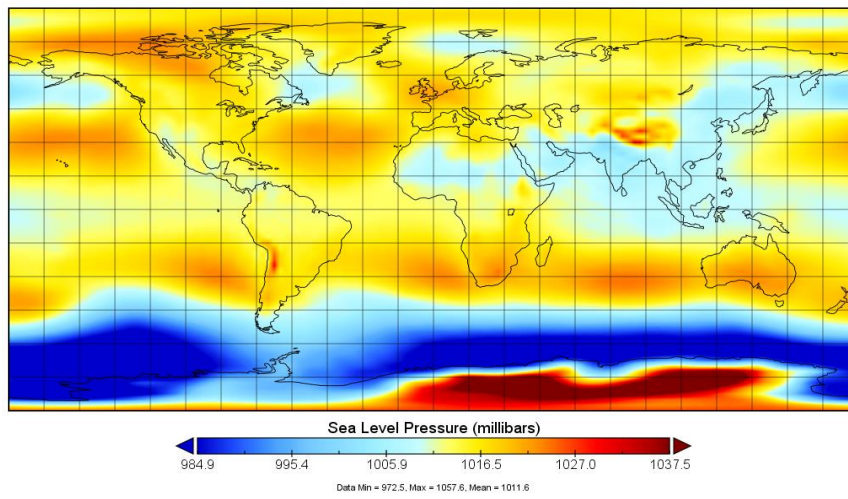


Figure 6. Monthly mean sea level pressure over the globe of May 2020

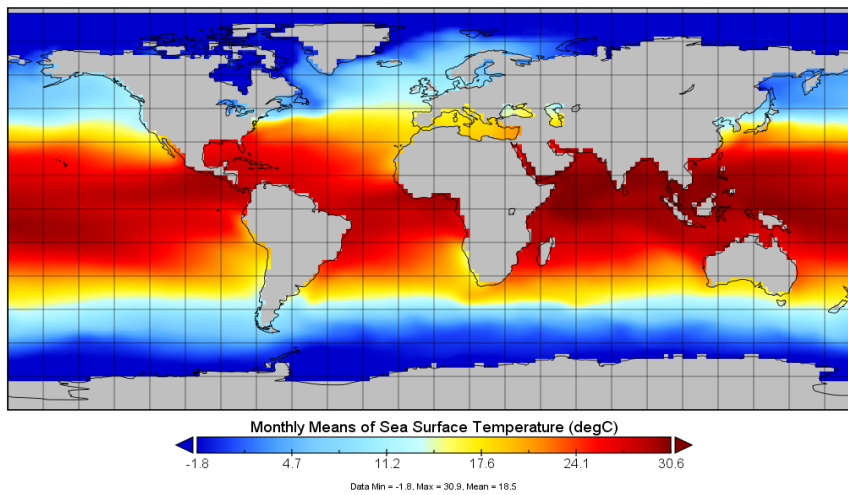


Figure 7. Monthly mean sea surface temperature over the globe of May 2019

These datasets are in netCDF file format and extracted to Microsoft Excel using Panoply software. The pre-processing of this data was done for dimensionality reduction and to determine the anomaly over the years. All the datasets were first averaged over the 10° latitude and 10° longitude. Averaging reduced the climate variable data to 18 latitudes and 36 longitudes. The data of a particular month was available sequentially over all the latitudes and longitudes in a tabular format. Hence, the data was organized for each latitude-longitude pair to prepare the time series of the climate variable. Anomaly time series was obtained by subtracting the mean value from the datapoint as given in (4) [13]:

$$\text{Anomaly}_m = X_m - \text{mean}(X_m) \quad (4)$$

where X_m is the climate variable value for month m, and $\text{mean}(X_m)$ is the average of variable values over all the years under study for month m. For calculation of anomaly, the data were rearranged for each month from the year 1948 to the year 2020. After anomaly calculation, anomaly time series was obtained where the records were then available year-wise, i.e., for each pair of latitude and longitude, climate variable anomaly was arranged from January to December 1948, January to December 1949, and so on till December 2020.

The monthly rainfall data over the Maharashtra time series was prepared from the monthly rainfall over four subdivisions, Madhya Maharashtra, Marathwada, Vidarbha, and Konkan-Goa, available from the years 1871 to 2016. A new dataset was prepared for applying the supervised feature selection method over all the available features to select AT, SLP, and SST at spatial regions correlated with monthly rainfall over Maharashtra (MMR). This dataset consisted of all the records from January 1948 to December 2016. Hence, three such datasets were prepared. The MI feature selection technique was then used to get maximum correlated pressures and temperatures over the 648 latitude longitude pairs (18 latitudes X 36 longitudes=648 features). The algorithm was run over each latitude and longitude pair, and significantly correlated latitudes and longitudes were plotted on the world map by selecting a threshold for maximum correlation.

For AT and SLP, 19 spatial locations were selected based on the threshold value of greater than 0.055; for SST, 13 spatial locations were selected based on the threshold value greater than 0.05. To understand where these locations are situated with respect to Maharashtra, they were plotted on the world map. Figures 8, 9, and 10 show spatial locations of selected features for AT, SLP, and SST, respectively.

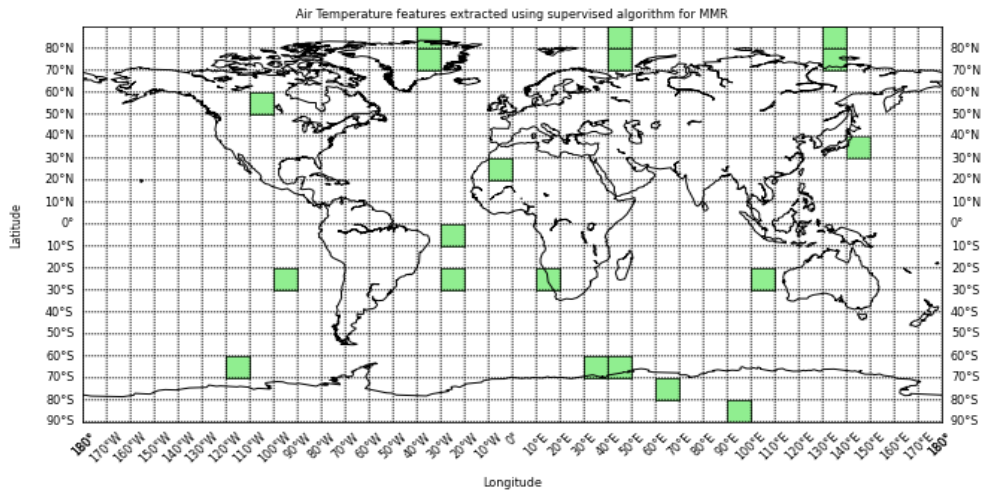


Figure 8. Selected features of air temperature over the globe for monthly rainfall prediction with a 0.055 value as a threshold

It can be observed from Figure 8 that mean temperatures of the Western and Southwest regions are significantly correlated with MMR. In the case of SLP, mean pressure over the Southern hemisphere is more significant than in the Northern hemisphere. In the case of SST, it was observed from Figure 10 that locations in the Southern hemisphere played an important role. Almost half of the locations were from 40°S to 50°S and 20°W to 90°W. To check whether the lagged correlation reveals the same geographical locations, the algorithm was run to plot those. The lag was 12 months to verify the rainfall's correlation with these climate variables.

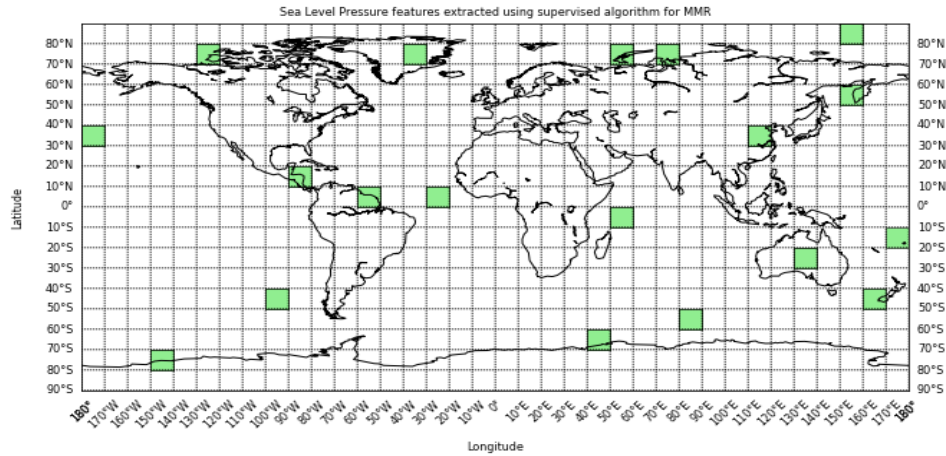


Figure 9. Selected features of sea level pressure over the globe for monthly rainfall prediction with 0.055 value as a threshold

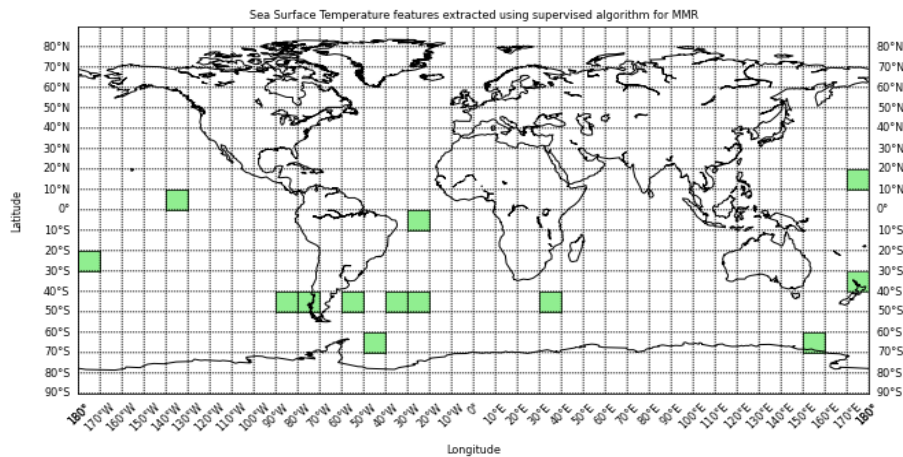


Figure 10. Selected features of sea surface temperature over the globe for monthly rainfall prediction with a 0.05 value as a threshold

The resulting latitudes and longitudes were plotted as shown in Figures 11 to 13 for AT, SLP, and SST, respectively. For lagged correlations, 14 locations for AT, 19 locations for SLP, and 14 locations for SST were selected. The threshold value was determined based on that for the lag 0 cases. The highest 30% MI values were used in each case.

The lagged correlations revealed that the mean temperatures and pressures near the equator and in the Southern Hemisphere were more significant. SLP in the Indian, South Pacific, and South Atlantic Oceans was significant. The AT over the Arabian Sea was significantly related to MMR, whereas the SST of the Indian Ocean was closely related to MMR. AT and SST in the Southern Hemisphere were more significant than in the Northern Hemisphere. In the case of lagged correlations, geographical locations identified as significant predictors of AT, SLP, and SST were closer to India than the lag 0 correlation.

The time series of the temperatures and pressure anomalies at the selected latitudes and longitudes obtained using lagged correlation was prepared as input to the ANN. All records were normalized using min-max normalization. A simple architecture of ANN with $2n+1$ hidden neurons in a single hidden layer was employed for this purpose, where n is a number of inputs [40]. The dataset was divided into training and testing sets. Monthly records from 1948 to 2010 were used for training, and those from 2011 to 2015 were used for testing purposes. Overall, 60 records were separately tested for long-range forecasting. The network's performance was evaluated based on MAE, RMSE, R-value, and maximum rainfall prediction capability of the network. Initially, the networks were separately developed, trained, and tested for AT, SLP, and SST. The results are presented in Table 1.

Table 1. Performance of prediction models trained and tested with ANN

Evaluation parameter	Air temperature	Sea level pressure	Sea surface temperature
MAE (mm)	224.8	428.3	571
RMSE (mm)	326.7	733.3	665.0
Maximum predicted rainfall (mm)	1574.5	1819.2	856.2
Maximum actual rainfall during the test period (mm)	2218.3	2218.3	2218.3
R-value	0.86	0.82	-0.29
R ² score	0.73	0.67	0.09
Percentage of MAE with respect to Maximum rainfall	10.13	19.31	25.7

From the results, it was observed that each climate variable could perform better on at least one evaluation criterion. The best results obtained are highlighted in Table 1. We observed that AT provided minimum MAE, RMSE, maximum R-value, and R² score. SLP provided maximum rainfall capability, and SST resulted in moderate RMSE. Concerning the maximum actual rainfall during the test period, the percentage MAE was minimum in the case of AT. What if all the climate features are combined? Will it improve the prediction performance? The ANN was developed following the same network architecture to get answers to these questions. This time, the inputs were the significant predictors identified by MI of AT, and SLP for the first network combined model 1 and AT, SLP and SST for the second network referred as combined model 2. The performance of the networks is given in Table 2.

The results show that the combined model 2 is better in terms of MAE and maximum rainfall prediction capability. In contrast, combined model 1 is better in terms of RMSE, R-value and R² score. If the R² score is compared, the AT-only and SLP-only provided better results than the combined model 1 and 2. The graphical representation of the prediction results during the training and testing phase of combined model-1 are shown in Figures 14 and 15. Further monthly prediction result analysis revealed that the maximum R-value was obtained for June, July, August, and September. The graph of the performance for the rainfall, in September, for the years 1949 to 2016 is shown in Figure 16. This highlights the effect of climate variables on the summer monsoon months.

Table 2. Performance of prediction model with AT, SLP, and SST as predictors

Evaluation parameter	Air temperature, sea level pressure as predictors (Combined Model-1)	Air temperature, sea level pressure, sea surface temperature as predictors (Combined Model-2)
MAE (in mm)	299	268
RMSE (in mm)	396	408
Maximum predicted rainfall (in mm)	1831	1938
Maximum actual rainfall during the test period (in mm)	2218.3	2218.3
R-value	0.78	0.77
R ² score	0.66	0.53
Percentage of MAE with respect to Maximum rainfall	13.5	12

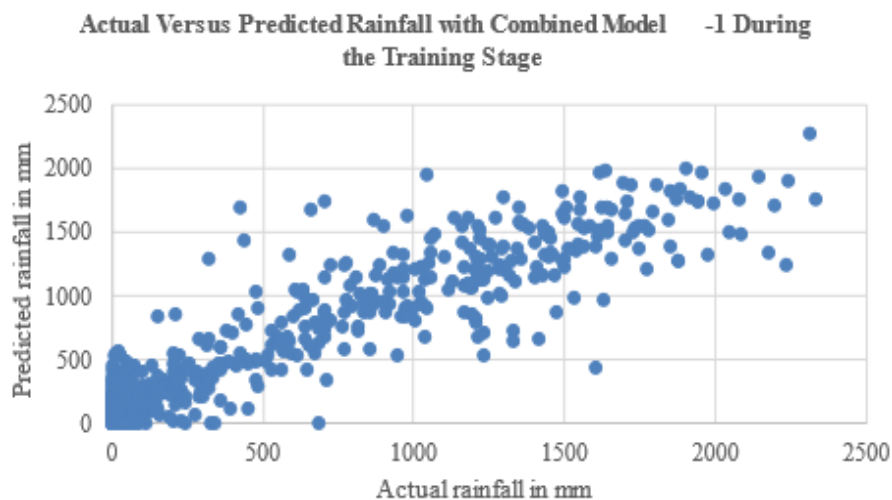


Figure 14. Graph showing actual versus predicted rainfall during the training stage of the combined model 1

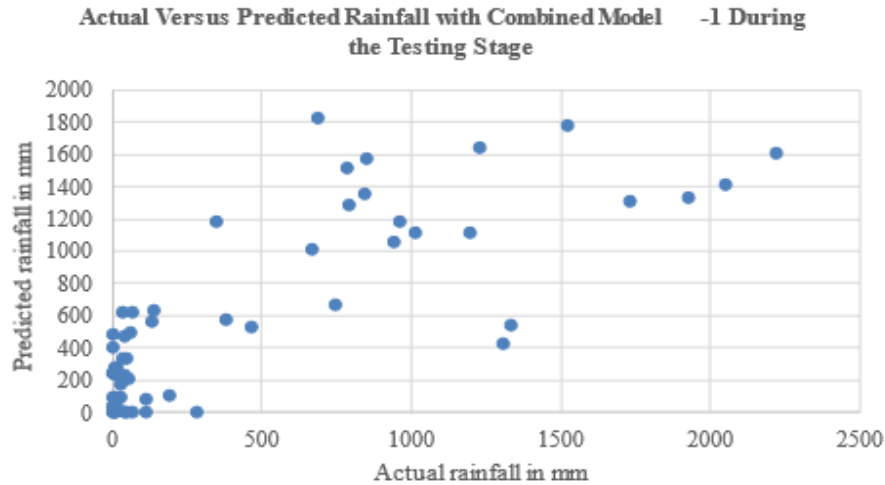


Figure 15. Graph showing actual versus predicted rainfall during the testing stage of the combined model 1

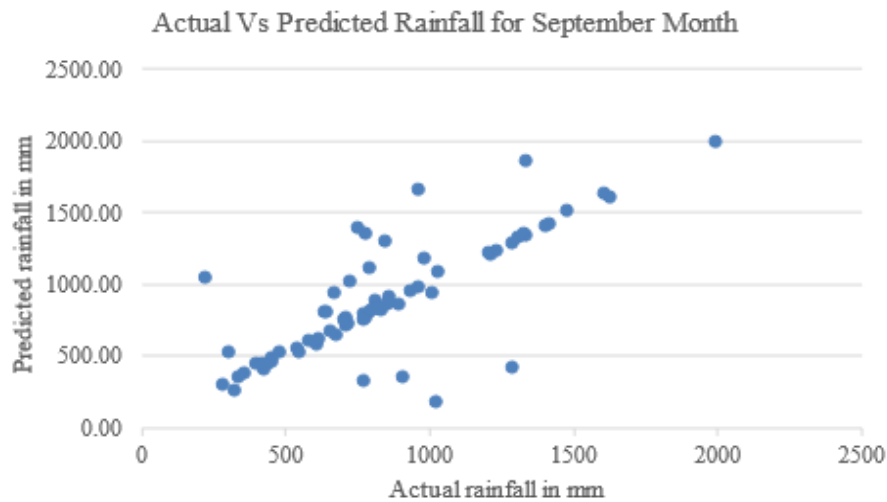


Figure 16. Graph showing actual versus predicted rainfall for the September month rainfall obtained by combined model 2

6. CONCLUSION AND FUTURE SCOPE

Monthly rainfall prediction was undertaken using the climate variables over the globe. AT, SLP, and SST were considered as potential predictors, and the data which is available on the resolution of 2.5° latitude X 2.5° longitude for AT and SLP and 2° latitude X 2° longitude for SST was averaged on the 10° latitude X 10° longitude. MI feature selection technique was used for further dimensionality reduction. ANN was used for the prediction of the long-range forecast rainfall. The results obtained by individual and combined climate variables were evaluated for MAE, RMSE, R-value, R^2 -score, and maximum prediction capability of the prediction model. AT gave a maximum R-value of 0.86, a maximum R^2 -score of 0.73, and a minimum RMSE of 326.7 mm. SLP showed a better capability to predict maximum rainfall than AT and SST, and it also showed a comparable R-value of 0.82. The combined model 2 with AT, SLP, and SST performed moderately, resulting in the best maximum rainfall prediction capability whereas the combined model with AT and SLP as inputs performed better over the combined model 2 for all parameters except the maximum rainfall capability. It showed moderate MAE and RMSE and an acceptable R-value and R^2 -score. The approach used in this work can be utilized for any other state with high rainfall variability.

In this study, three important climate variables were used. AT at various levels of the atmosphere was not considered. Wind and geopotential height were also not considered. More climate variables can be further explored to achieve more accuracy in forecasting. It may also help to understand the interdependence of all climate variables and their significance and capability in predicting monthly rainfall.




REFERENCES

- [1] M. Mohapatra, "Forecaster's guide," Deputy Director-General of Meteorology (Weather Forecasting) and India Meteorological Department, Pune, India, 2012.
- [2] S. Hastenrath, "On climate prediction in the tropics," *Bulletin of the American Meteorological Society*, vol. 67, no. 6, pp. 696–702, Jun. 1986, doi: 10.1175/1520-0477(1986)067<0696:OCPITT>2.0.CO;2.
- [3] G. Di Capua *et al.*, "Long-lead statistical forecasts of the Indian summer monsoon rainfall based on causal precursors," *Weather and Forecasting*, vol. 34, no. 5, pp. 1377–1394, Oct. 2019, doi: 10.1175/WAF-D-19-0002.1.
- [4] V. Moron and A. W. Robertson, "Interannual variability of Indian summer monsoon rainfall onset date at local scale," *International Journal of Climatology*, vol. 34, no. 4, pp. 1050–1061, Mar. 2014, doi: 10.1002/joc.3745.
- [5] R. J. Bombardi, V. Moron, and J. S. Goodnight, "Detection, variability, and predictability of monsoon onset and withdrawal dates: a review," *International Journal of Climatology*, vol. 40, no. 2, pp. 641–667, Feb. 2020, doi: 10.1002/joc.6264.
- [6] R. P. Shukla, K. C. Tripathi, A. C. Pandey, and I. M. L. Das, "Prediction of Indian summer monsoon rainfall using Niño indices: a neural network approach," *Atmospheric Research*, vol. 102, no. 1–2, pp. 99–109, Oct. 2011, doi: 10.1016/j.atmosres.2011.06.013.
- [7] N. Acharya, S. Chattopadhyay, M. A. Kulkarni, and U. C. Mohanty, "A neurocomputing approach to predict monsoon rainfall in monthly scale using SST anomaly as a predictor," *Acta Geophysica*, vol. 60, no. 1, pp. 260–279, Feb. 2012, doi: 10.2478/s11600-011-0044-y.
- [8] A. Nair, G. Singh, and U. C. Mohanty, "Prediction of monthly summer monsoon rainfall using global climate models through artificial neural network technique," *Pure and Applied Geophysics*, vol. 175, no. 1, pp. 403–419, Jan. 2018, doi: 10.1007/s00024-017-1652-5.
- [9] Y. M. Chiang, F. J. Chang, B. J. D. Jou, and P. F. Lin, "Dynamic ANN for precipitation estimation and forecasting from radar observations," *Journal of Hydrology*, vol. 334, no. 1–2, pp. 250–261, Feb. 2007, doi: 10.1016/j.jhydrol.2006.10.021.
- [10] J. Lee, C.-G. Kim, J. Lee, N. Kim, and H. Kim, "Application of artificial neural networks to rainfall forecasting in the Geum River Basin, Korea," *Water*, vol. 10, no. 10, p. 1448, Oct. 2018, doi: 10.3390/w10101448.
- [11] M. Sangiorgio *et al.*, "A comparative study on machine learning techniques for intense convective rainfall events forecasting," in *Theory and Applications of Time Series Analysis*, Springer International Publishing, 2020, pp. 305–317.
- [12] C. Z. Basha, N. Bhavana, P. Bhavya, and S. V., "Rainfall prediction using machine learning deep learning techniques," in *International Conference on Electronics and Sustainable Communication Systems (ICESC)*, Jul. 2020, pp. 92–97, doi: 10.1109/ICESC48915.2020.9155896.
- [13] M. Saha, P. Mitra, and R. S. Nanjundiah, "Autoencoder-based identification of predictors of Indian monsoon," *Meteorology and Atmospheric Physics*, vol. 128, no. 5, pp. 613–628, Oct. 2016, doi: 10.1007/s00703-016-0431-7.
- [14] X.-B. Jin, W.-T. Gong, J.-L. Kong, Y.-T. Bai, and T.-L. Su, "A variational Bayesian deep network with data self-screening layer for massive time-series data forecasting," *Entropy*, vol. 24, no. 3, Feb. 2022, doi: 10.3390/e24030335.
- [15] M. N. M. Sap and A. M. Awan, "Finding spatio-temporal patterns in climate data using clustering," in *International Conference on Cyberworlds (CW'05)*, 2005, vol. 2005, doi: 10.1109/CW.2005.45.
- [16] N. A. Nnamoko, F. N. Arshad, D. England, J. Vora, and J. Norman, "Evaluation of filter and wrapper methods for feature selection in supervised machine learning," in *The 15th Annual Postgraduate Symposium on the convergence of Telecommunication, Networking and Broadcasting*, Liverpool, UK, 2014.
- [17] K. Kim, H. Joo, D. Han, S. Kim, T. Lee, and H. S. Kim, "On complex network construction of rain gauge stations considering nonlinearity of observed daily rainfall data," *Water*, vol. 11, no. 8, Jul. 2019, doi: 10.3390/w11081578.
- [18] W. Wei, X. Fan, H. Song, and H. Wang, "Video tamper detection based on multi-scale mutual information," *Multimedia Tools and Applications*, vol. 78, no. 19, pp. 27109–27126, Oct. 2019, doi: 10.1007/s11042-017-5083-1.
- [19] Y. Wu, B. Liu, W. Wu, Y. Lin, C. Yang, and M. Wang, "Grading glioma by radiomics with feature selection based on mutual information," *Journal of Ambient Intelligence and Humanized Computing*, vol. 9, no. 5, pp. 1671–1682, Oct. 2018, doi: 10.1007/s12652-018-0883-3.
- [20] F. Zhao, J. Zhao, X. Niu, S. Luo, and Y. Xin, "A filter feature selection algorithm based on mutual information for intrusion detection," *Applied Sciences*, vol. 8, no. 9, Sep. 2018, doi: 10.3390/app8091535.
- [21] X. He, H. Guan, and J. Qin, "A hybrid wavelet neural network model with mutual information and particle swarm optimization for forecasting monthly rainfall," *Journal of Hydrology*, vol. 527, pp. 88–100, Aug. 2015, doi: 10.1016/j.jhydrol.2015.04.047.
- [22] M. Sangiorgio and F. Dercole, "Robustness of LSTM neural networks for multi-step forecasting of chaotic time series," *Chaos, Solitons and Fractals*, vol. 139, Oct. 2020, doi: 10.1016/j.chaos.2020.110045.
- [23] "Monthly, seasonal and annual rainfall (in 10th of mm) 1871–2016 (1871–2014 based on 306 stations and 2015–2016 based on IMD subdivisional rainfall)," *Indian Institute of Tropical Meteorology*, <https://www.tropmet.res.in/data/data-archival/rain/iitm-subdivrf.txt> (accessed Jan. 31, 2021).
- [24] S. Galelli, G. B. Humphrey, H. R. Maier, A. Castelletti, G. C. Dandy, and M. S. Gibbs, "An evaluation framework for input variable selection algorithms for environmental data-driven models," *Environmental Modelling & Software*, vol. 62, pp. 33–51, Dec. 2014, doi: 10.1016/j.envsoft.2014.08.015.
- [25] R. Taormina, S. Galelli, G. Karakaya, and S. D. Ahipasaoglu, "An information theoretic approach to select alternate subsets of predictors for data-driven hydrological models," *Journal of Hydrology*, vol. 542, pp. 18–34, Nov. 2016, doi: 10.1016/j.jhydrol.2016.07.045.
- [26] R. Zhang, F. Nie, X. Li, and X. Wei, "Feature selection with multi-view data: a survey," *Information Fusion*, vol. 50, pp. 158–167, Oct. 2019, doi: 10.1016/j.inffus.2018.11.019.
- [27] J. Li and H. Liu, "Challenges of feature selection for big data analytics," *IEEE Intelligent Systems*, vol. 32, no. 2, pp. 9–15, Mar. 2017, doi: 10.1109/MIS.2017.38.
- [28] Y. Ji, J. Hao, N. Reyhani, and A. Lendasse, "Direct and recursive prediction of time series using mutual information selection," in *Lecture Notes in Computer Science*, vol. 3512, Springer Berlin Heidelberg, 2005, pp. 1010–1017.
- [29] M. I. Belghazi *et al.*, "Mutual information neural estimation," *35th International Conference on Machine Learning, ICML 2018*, vol. 2, pp. 864–873, 2018.
- [30] G. Chandrashekar and F. Sahin, "A survey on feature selection methods," *Computers and Electrical Engineering*, vol. 40, no. 1, pp. 16–28, Jan. 2014, doi: 10.1016/j.compeleceng.2013.11.024.
- [31] S. Chormunge and S. Jena, "Correlation based feature selection with clustering for high dimensional data," *Journal of Electrical Systems and Information Technology*, vol. 5, no. 3, pp. 542–549, Dec. 2018, doi: 10.1016/j.jesit.2017.06.004.
- [32] B. Venkatesh and J. Anuradha, "A review of feature selection and its methods," *Cybernetics and Information Technologies*, vol. 19, no. 1, pp. 3–26, Mar. 2019, doi: 10.2478/cait-2019-0001.




- [33] B. Ghoghj *et al.*, "Feature selection and feature extraction in pattern analysis: a literature review," *arxiv.org/abs/1905.02845*, May 2019, [Online]. Available: <http://arxiv.org/abs/1905.02845>.
- [34] M. Sangiorgio *et al.*, "Improved extreme rainfall events forecasting using neural networks and water vapor measures," in *6th International conference on Time Series and Forecasting*, 2019, pp. 820–826.
- [35] S. N. Sivanandam and S. N. Deepa, *Principles of soft computing*. Wiley India Pvt. Limited, 2011.
- [36] C. Nwankpa, W. Ijomah, A. Gachagan, and S. Marshall, "Activation functions: comparison of trends in practice and research for deep learning," Nov. 2018, [Online]. Available: <http://arxiv.org/abs/1811.03378>.
- [37] "National Oceanic and Atmosphere Administration-NOAA," *NOAA Physical Sciences Laboratory (PSL)*. <http://www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis.derived.html> (accessed Jan. 31, 2021).
- [38] "National Oceanic and Atmosphere Administration-NOAA," *NOAA Physical Sciences Laboratory (PSL)*. <http://www.psl.noaa.gov/data/gridded/data.ncep.reanalysis.derived.html> (accessed Jan. 31, 2021).
- [39] "National Oceanic and Atmosphere Administration-NOAA," *NOAA Physical Sciences Laboratory (PSL)*. <https://www.psl.noaa.gov/data/gridded/data.noaa.ersst.html> (accessed Jan. 31, 2021).
- [40] S. Hudnurkar and N. Rayavarapu, "Performance of artificial neural network in nowcasting summer monsoon rainfall: a case study," in *IEEE Punecon*, Nov. 2018, pp. 1–5, doi: 10.1109/PUNECON.2018.8745413.

BIOGRAPHIES OF AUTHORS



Shilpa Hudnurkar    is B.E. Instrumentation and has completed an M. Tech in Electronics and Telecommunication. She is currently an Assistant Professor in the Department of Electronics and Telecommunication Engineering at Symbiosis Institute of Technology, affiliated with Symbiosis International (Deemed University). She is a research scholar at Symbiosis International (Deemed University). Her research interests include artificial intelligence, machine learning, deep learning, and signal processing. She is working on predicting summer monsoon rainfall over a small region. Her teaching experience is over seven years. She can be contacted at email: shilpa.hudnurkar@sitpune.edu.in.



Neela Rayavarapu    received her BE degree in Electrical Engineering from Bangalore University, Bangalore, India, in 1984, MS Degree in Electrical and Computer Engineering from Rutgers, The State University of New Jersey, USA, in 1987, and a Ph.D. degree in Electronics and Communication Engineering in 2012 from Panjab University, Chandigarh. She has been involved in teaching and research in Electrical, Electronics, and Communication Engineering since 1987. Her areas of interest are digital signal processing and its applications and control systems. She can be contacted at: neela.raya27@gmail.com.