Journal of Soft Computing in Civil Engineering 6-3 (2022) 63-77

Contents lists available at SCCE

Journal of Soft Computing in Civil Engineering

Journal homepage: www.jsoftcivil.com

Combined Standardized Precipitation Index and ANFIS Approach for Predicting Rainfall in the Tropical Savanna Region

Fenil R. Gandhi[1*], Jayantilal N. Patel[2]

1. Ph.D. Student, Department of Civil Engineering, S. V. National Institute of Technology, Surat, India
2. Professor (HAG), Department of Civil Engineering, S. V. National Institute of Technology, Surat, India
Corresponding author: *fenilgandhi15@gmail.com*

https://doi.org/10.22115/SCCE.2022.333365.1412

ARTICLE INFO

Article history:
Received: 10 March 2022
Revised: 07 June 2022
Accepted: 31 August 2022

Keywords:
ANFIS;
Drought;
Forecasting;
SPI;
Surat.

ABSTRACT

Climate change has affected many sectors in the world. Therefore, the Prediction of climatic factors is essential in case to achieve sustainability in human life. Rainfall prediction is also important as the agricultural sector depends on rainfall, and human life depends on agricultural products. This study presents the Standardized Precipitation Index (SPI) prediction using the adaptive neuro-fuzzy inference system (ANFIS). Various models (6 nos.) with different combinations of Rainfall and SPI values are prepared to predict the SPI index. Out of these six models, the M2 model (SPI3 SPI4 R4) performed best in the case of SPI 5. (RMSE value is 0.059, the R^2 value is 0.987, and the value of the coefficient of determination is 0.993. In the case of SPI 6, the M1 model (SPI5 SPI4 SPI3 R5) performed best (RMSE value is 0.042, the R^2 value is 0.992, and the value of the coefficient of determination is 0.996. The outcome may be helpful to the policymakers, scientists, researchers and government authorities in building a policy for sustainable water resources management in the region.

How to cite this article: Gandhi FR, Patel JN. Combined standardized precipitation index and ANFIS approach for predicting rainfall in the tropical savanna region. J Soft Comput Civ Eng 2022;6(3):63–77. https://doi.org/10.22115/scce.2022.333365.1412

1. Introduction

The world is changing at a faster rate than ever before. Even in the agricultural industry, new technology is emerging. Nature, on the other hand, remains unconquerable by humans. For a long, human beings have been searching for security, making attempts to accomplish it. Forecasting techniques are also developed in various sectors to achieve security. Extreme rainfall and drought events are serious for any human society. Hence forecasting the rainfall events is equally important. Drought and extreme rainfall occur in all climatic zones, and it occurs due to uneven precipitation from a few months to many years.

The Artificial Intelligence (AI) technique is commonly utilized to solve hydrological problems, such as forecasting rainfall. The capacity of the software to depict input-output patterns without knowing the factors impacting the prediction parameters is important to the utility of this modeling [1]. This forecast is mostly beneficial to farmers, and it is also possible to effectively utilize water supplies. Prediction of rainfall is a quite difficult task as it possesses uncertainty. However, Machine learning and artificial intelligence techniques have made it a little simpler. To predict rainfall for a particular region, the historical dataset is essential. The soft computing techniques have the ability to compare the modeled dataset and the actual data set with parameters like Root Mean Square Error (RMSE), Mean Squared Error (MSE), Coefficient of Correlation (R^2), coefficient of determination, etc. Different methods provide varied accuracy levels, so it's important to pick the proper algorithm and model it to meet the criteria.

Researchers are using different tools to predict rainfall indices. Researcher [2] used an Autoregressive integrated moving average (ARIMA) statistical model with Long short-term memory (LSTM) using multivariate inputs. Researchers [3] have developed a trained and validated ANFIS model for drought forecasting. The best model has also been tested and trained using a feed-forward Neural Network. ANFIS models provided high accuracy and reliability [4]. Moreover, Author [5] developed ARIMA and Artificial Neural Network (ANN) models. ANN proved a better performance model. Furthermore, scientists[6] developed ANN, Support Vector Regression (SVR), and Wavelet Neural Network (WNN) models to predict SPI values. The WNN model was considered to be effective. [7] developed six different neural networks. Out of those models, wavelet neural network (WNN), Radial Basis Functions (RBF) and General Regression Neural Network (GRNN) had better performance. Researchers [8] developed ANFIS, M5P, and Multilayer Perceptron (MLP) models. MLP had given promising results. Additionally, [9] used Multivariate Adaptive Regression Splines (MARS), Least-Square Support Vector Machine (LSSVM), and M5 Tree models to predict drought index.[10] developed different models using ANN, ANFIS, and SVM. Results displayed that SVM models have high accuracy. Models for Recursive Multi-Layer Perceptron (RMLP) and recursive support vector regression (RSVR) were also used by the researcher and optimized using an Imperialist Competitive Method (ICA) [9]. [11] developed Distributed Lag Nonlinear Model (DLNM) and XGBoost model. Out of these two models, the XGBoost model was better. [12] developed wavelet-ANN, ANN, and Wavelet- ARIMA-ANN models. Wavelet 2A model had better performance. [13] developed wavelet neural network (WNN), ARIMA, and SVM models. Out of these models, the ARIMA model was better. [14] developed ANN, Wavelet ANN, and Wavelet- ARIMA-ANN model, W2A performed best.

The Surat district of Gujarat state falls under the tropical savanna climate region. The climate of the area is driven by the Arabian sea and the bay of Khambhat. The district's 70 % of the land is agricultural [15]. Additionally, the district can be known as the country's economic capital, which has an important role in the GDP of the country. The district has faced various tangible and intangible damages due to the heavy and low rainfall in previous years [16]. Therefore, it is essential to understand the climate condition of the area to arrange the necessary steps for the future. The various studies have been done using the SPI index and ANFIS techniques. However, the long-term record of the dataset for a region is not studied yet in order to forecast rainfall more accurately for the selected study area. Moreover, no effort has been made to forecast rainfall in the study area using SPI and ANFIS with different possibilities and inputs till date. Therefore, the present study aims (1) to analyze the rainfall pattern by SPI index and (2) to model the rainfall by combining the SPI-ANFIS approach with different pairs of inputs for the Surat district. The research will assist the policymakers, government authorities, and scientists in developing an action plan for mitigating the uneven rainfall events to manage water resources in the region.

2. Methodology

2.1. Study area and data collection

Surat district is located on the western coast of the Deccan Peninsula and in the southern portion of Gujarat state (India) (Fig. 1). The district is located between 21° 30′ N and 21° 35′ N (latitude) and 72° 35′ E and 74° 20′ E (longitude). The four major zones in the region are hilly areas, piedmont slopes, alluvial plains, and coastal plains [17]. The Tapi River, which is perennial and flows westward into the Arabian Sea, is the area's largest significant water supply. The district has a tropical savanna climate moderated by the Arabian Sea and the Gulf of Cambay. Winter (November-February), Summer (March-June), and Monsoon (July-August) are the three seasons of the district (July-October). The hottest months are April and May, with average high temperatures of 37 degrees Celsius. Furthermore, during the monsoon season, the average rainfall is around 1140 mm. Winter lasts from December until late February, with temperatures averaging approximately 23 degrees Celsius and minimal rain. The rainfall data is used to fulfill the objective of the study. Here, monthly data is collected from two sources (1) year 1901-2002: downloaded from the India water portal (2) year 2003-2016: collected from the State water data center, Gandhinagar. Sample data table is shown in Table 1.

Table 1
Sample data.

YEAR	Jan	Feb	Mar	Apr	May	June	July	August	Sept	Oct	Nov	Dec
1901	2.714	0	2.115	0.776	1.595	77.731	445.8	194.179	20.291	19.766	0.196	0
1902	2.275	0	0	0.029	0.411	55.08	581.097	387.121	389.041	10.537	4.281	7.703
1903	0.008	0	0	0	20.701	96.372	724.489	206.228	262.72	10.658	0.196	0.048
1904	0	0.804	2.363	0.329	0.405	63.648	191.236	74.714	211.709	10.66	0.196	0.904
1905	0	0.184	0	0.001	0.095	29.427	501.574	47.41	49.12	3.354	0.369	0

2.2. Standard precipitation index (SPI)

The SPI is the most widely used indicator globally for identifying and monitoring meteorological droughts. [18] developed the SPI indicator, which compares observed total precipitation quantities for a particular accumulation period (e.g., 1, 3, 4, 6, 12, 48 months) with the long-term historic rainfall record for that time. The SPI is calculated by dividing the difference in precipitation over a particular time frame by a standard deviation and subtracting the difference from the mean,

$$\mathbf{SPI} = {\mathbf{y_{jk}} - \mathbf{y_j}}\Big/\mathbf{SD_j} \tag{1}$$

SD_j is Standardized deviation for the jth station, y_{jk} = rainfall for the jth location and kth observation and y_j= Mean rainfall for the jth location.

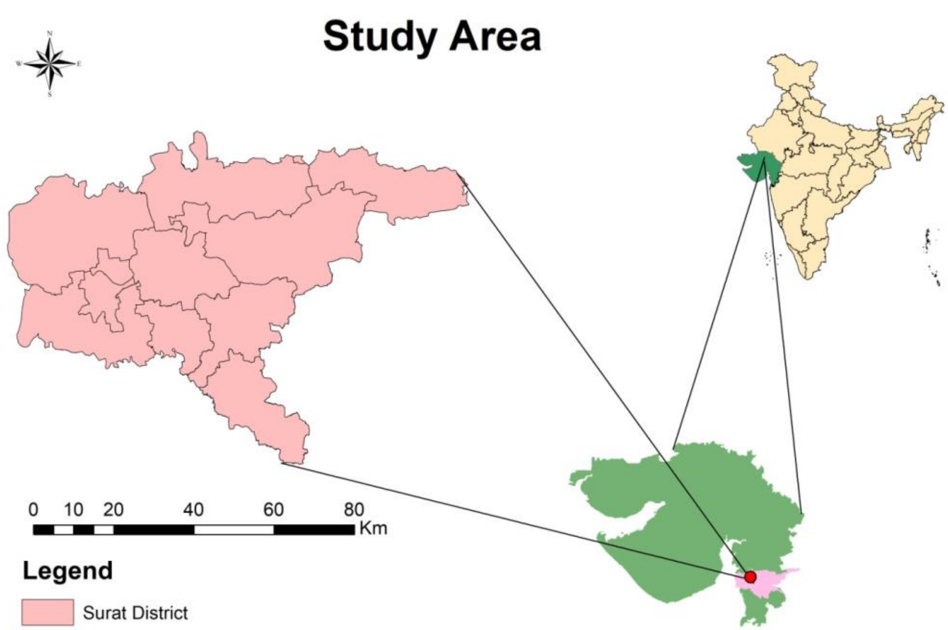

Fig. 1. Study area.

As in India, generally, a monsoon season is between June to September (sometimes it lasts till October and November). Therefore, the authors have computed the SPI value for 3,4, 5 and 6 months, starting from June. Here, SPI 3 = SPI value calculated from June to Aug, SPI 4 = SPI value calculated from June to September, SPI 5 = SPI value calculated from June to October, and SPI 6 = SPI value calculated from June to November. Here the calculation of SPI value can be done under two types of distribution, such as Gamma Distribution and Log-Normal Distribution. However, in the study, Gamma distribution has been considered for the calculation of SPI value. The observed SPI values can be classified with a table given by [18].

2.3. Adaptive neuro-fuzzy inference system (ANFIS) modeling

ANFIS technique is being investigated for rainfall prediction. The Sugeno first order is used in the ANFIS model. The ANFIS fuzzy inference algorithm is a method for generating a new approximate fuzzy set resolution using fuzzy rules and a fuzzy set as a premise. The fuzzy inference system (FIS) is typically used when systems are difficult to describe accurately or when the description of the research problems is imprecise and ambiguous with the use of an ANFIS; input characteristics are mapped to input membership functions (MFs), input MFs to a set of if-then rules, rules to a set of output characteristics, output characteristics to output MFs, and output MFs to a single-valued output or a decision associated with the output [19,20].

The ANFIS-GUI is designed to assist users in predicting the output. ANFIS uses Sugano FIS to forecast data and tune membership using backpropagation or a hybrid technique. Using the input-output Relationship as a preliminary step, the fuzzy rule was created using this study's fuzzy grid partition functions. The hybrid approach was utilized to compute the nonlinear input and linear output parameters. This method offered the learning technique as well as the rule building. The ANFIS structure is made up of four layers of neurons, each with its own unique behavior. Layers 1 and 4 have changing parameters, allowing the network to be trained. The behavior of all nodes within the same layer is identical, subject to the influence of altering factors (layers 1 and 4). Fig. 2 depicts the structure of ANFIS.

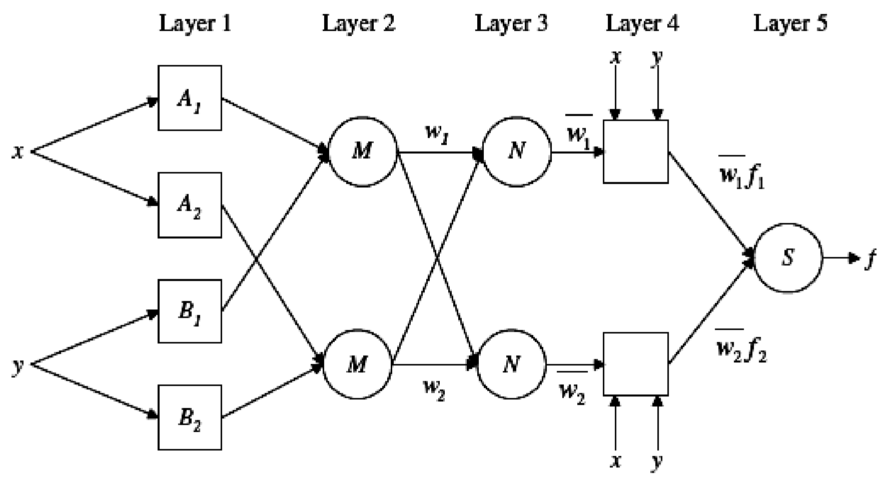

Fig. 2. ANFIS architecture.

The fundamental issue with fuzzy logic is that there is no standardized process for defining membership function parameters and creating fuzzy sets. However, by successfully utilizing the learning potential of ANN for automated fuzzy rule generation and optimization techniques, ANFIS solves the fundamental challenge in fuzzy technique development (identifying the MF parameters and constructing fuzzy if-then rules). FIS is considered to have two input parameters, SPI(T-1) and R(T-1), and one output parameter, SPI(T-1). The rule can be written,

Rule 1: IF SPI$(\boldsymbol{T}-\mathbf{1})$ be $\boldsymbol{X_1}$, $\boldsymbol{R}(\boldsymbol{T}-\mathbf{1})$ be $\boldsymbol{Y_1}$ THEN $\boldsymbol{f_1}=\boldsymbol{a_1}\times\mathbf{SPI}(\boldsymbol{T}-\mathbf{1})+\boldsymbol{b_1}\times\boldsymbol{R}(\boldsymbol{T}-\mathbf{1})+\boldsymbol{c_1}$

Rule 2: IF SPI$(\boldsymbol{T}-\mathbf{1})$ be $\boldsymbol{X_2}$, $\boldsymbol{R}(\boldsymbol{T}-\mathbf{1})$ be $\boldsymbol{Y_2}$ THEN $\boldsymbol{f_2}=\boldsymbol{a_2}\times\mathbf{SPI}(\boldsymbol{T}-\mathbf{1})+\boldsymbol{b_2}\times\boldsymbol{R}(\boldsymbol{T}-\mathbf{1})+\boldsymbol{c_2}$

Layer 1 possesses several input value nodes, each node generates a membership function for each given input, and the output will be computed by,

$$\boldsymbol{O_n^1 = \mu_{X_n}\mathbf{SPI}(T-1)} \text{ for } \boldsymbol{n = 1, 2} \text{ and } \boldsymbol{O_n^1 = \mu_{Y_{n-2}}R(T-1)} \text{ for } \boldsymbol{n = 3, 4} \tag{2}$$

where $\mathrm{SPI}(T-1)$ indicates the value of SPI at the time$(T-1)$ for the network node n, $R(T-1)$ depicts the rainfall occurrence at the time $(T-1)$, $\mathrm{SPI}(T)$ is the notation for SPI value at a time (T) for the network node n, the linguistic labels are denoted by X_n and Y_n, μ_{X_n} and μ_{y_n} are defined as membership functions (MF) of variables X_n and Y_n (linguistic variables). Here, the study is performed on each membership function: triangular mf, trapezoidal mf, gaussian mf, gaussian 2d mf, generalized bell MF, Sigmoid membership, and pimf, etc., in order to get a better output model. If the gaussian mf is employed,

$$O_n^1 = \mu_{X_n}(x) = \mathrm{e}^{\frac{-(\mathrm{SPI}[T-1)-c)^2}{2\sigma^2}} \tag{3}$$

Layer 2 possesses the product of the corresponding degree obtained from layer 1, which is denoted as O_n^2 and can be written as,

$$\boldsymbol{O_n^2 = w_n = \mu_{X_i}\mathbf{SPI}(T-1)\mu_{Y_i}R(T-1), n = 1, 2} \tag{4}$$

The basic aim is to find the ratio of the firing strength of each nth rule to the total firing strength of all rules (Layer 3). This layer's firing strength can be normalized as follows:

$$\boldsymbol{O_n^3 = \bar{w}_n = \frac{w_n}{\sum_i w_n}, n = 1, 2} \tag{5}$$

Eq. 6 calculates the n^th rule's contribution to the overall output, model output, and function specified (Layer 4)

$$O_n^4 = \bar{w}_n f_n = \bar{w}_n(a_n \times \mathrm{SPI}(T-1) + b_n \times R(T-1) + c_n), n = 1,2 \tag{6}$$

Layer 5 contains the output nodes, where a single node adds all incoming signals to produce the total output.

$$\boldsymbol{f(x, y) = \frac{w_1 f_1 + w_2 f_2}{w_1 + w_2}} \tag{7}$$

$$\boldsymbol{Q_n^5 = f(\mathbf{SPI}(T-1), R(T-1)) = \textstyle\sum_n \bar{w}_n \cdot f_n = \bar{w}_n f_1 + \bar{w}_n f_2 = \frac{\sum_n w_n f_n}{\sum_n w_n}} \tag{8}$$

As seen in the third layer, w_n Represents the preceding layer's n^th node output. The hybrid learning technique, which combines the "gradient descent" and "least-squares" techniques, is used by ANFIS to choose the input and output model parameters. The nonlinear antecedent parameters are assigned using the gradient descent approach, while the linear consequence parameters are identified using the least-squares method.

2.4. Input variables and model structure

Selecting input variables is one of the most crucial phases in building a successful forecasting model. Because the factors influence the weighted coefficient and the model's outcomes, they define the architecture of the prediction model. Various estimation models were created in the study. The models for the region are created in this study using various combinations of antecedent values of actual precipitation and SPI values. The weighted coefficient and the model's outcomes are affected by the specified variables, which form the structure of the forecasting model. In this study, rainfall and SPI values of years 1901 to 2016 were used to construct the ANFIS forecasting models.

Table 2
ANFIS model structure.

Model	Input	Output	Remark
M_1	SPI5 SPI4 SPI3 R(5)	SPI6	Combination of rainfall and SPI Value as input
M_2	SPI4 SPI3 R(4)	SPI5	
M_3	SPI5 SPI4 SPI3	SPI6	SPI values as input
M_4	SPI4 SPI3	SPI5	
M_5	R(5) R(4) R(3) R(2) R(1)	SPI6	Rainfall value as input
M_6	R(4) R(3) R(2) R(1)	SPI5	

The six models were constructed using different inputs. As shown in Table 2, In this research work model is created for SPI 5 and SPI 6. Yearly rainfall and SPI values from the year 1901 to 2016 are employed to build ANFIS forecasting models. Various models have been generated using combinations of SPI and Rainfall. First, the Model uses only the SPI value. Secondly, the Model uses only antecedent rainfall, and another model is created by combining both SPI value and antecedent rainfall. The obtained data set is bifurcated into two parts such as training and checking. While constructing the model, different epoch values, membership functions, and the value of membership functions are decided. The training data set contains data records from the years 1901 to 1981. Each model is trained and checked for 3, 4, 5, and 6 number of membership functions at epoch value 50. Models are tested using a data set that was not used during the training phase to obtain a more accurate evaluation and comparison. Data records from 1982 to 2016 make up the checking data set.

2.5. Model evaluation criteria

The system's output will be checked to see if it meets the goal that was set for it. If this is not the case, the learning method will be enhanced by implementing a new network design or learning. Root Mean Square Error (RMSE), Coefficient of Correlation (r), and Coefficient of Determination (r) are being used to assess the performance of ANFIS models (R^2). These indices can be used to assess the strength of the linear relationship between observed and estimated data [17]. RMSE can be computed from equation 9,

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(x_{forecasted} - x_{\mathrm{observed}}\right)^2} \tag{9}$$

The correlation coefficient (R) may be calculated using observation (training and testing), input parameters such as learned meteorological data, and network performance input-output structure.

$$R = \sqrt{1 - \frac{\sum_{j=1}^{n} (x_{observed} - x_{\text{forecasted}})^2}{\sum_{j=1}^{n} (x_{\text{forecasted}})^2}} \tag{10}$$

The coefficient of determination, commonly known as R^2, is a statistical technique for determining and evaluating a statistical model's capacity to explain and predict future events. It can be formulated by equation 11

Coefficient of Determination (R^2) = Explained Variation / Total Variation (11)

3. Results and discussion

3.1. General statistics

General statistical analysis is given in Table 2 of rainfall from 1991 to 2018. South Gujarat region and Surat district fall under three seasons. But for analysis, we have taken four groups of seasons, i.e., December to February (winter), March to May (PreM - Premonsoon), June to September (SWM - South-west monsoon) and October – November (PoM - Postmonsoon). The General analysis shows the different parameters for annual and seasonal (Table 3). The average annual precipitation is 1142 mm from 1901 to 2018.

Table 3
Statistical analysis of rainfall (the year 1901-2016).

Parameter	Winter	PreM	SWM	PoM	Annual
Average (mm)	2.12	5.74	1093.50	41.04	1142.41
Median	0.46	1.11	1057.27	21.52	1135.28
Mode	0.00	0.00	-	0.43	-
SD	3.84	11.04	316.28	49.35	316.75
CV	180.78	192.35	28.92	120.23	27.73
Kurtosis	7.66	12.51	-0.13	4.27	-0.09
Skewness	2.68	3.20	0.11	2.01	0.07
Min (mm)	0.00	0.00	309.82	0.00	321.38
Max (mm)	20.10	71.16	1955.42	253.06	1978.82
Confidence Level (95.0%)	0.70	2.02	57.91	9.04	58.00

Maximum and minimum rainfall is 1978 mm and 321 mm, respectively. Here, the coefficient of variance for winter, pre-monsoon, southwest monsoon, post-monsoon and annual are 180.78, 192.35, 28.92, 120.23 and 27.73. The lower 27.73 percent COV value shows that the inter-annual rainfall variability is lower. High COV values (> 100 percent) represent high variability between October and May, which are non-monsoon seasons. Generally, in the south Gujarat region, 95% of rainfall is contributed by the June to September months of annual rainfall. Likewise, pre-monsoon and post-monsoon contribute less than 0.20%, 0.50%, and less than 4% in the winter season, respectively.

The annual and seasonal rainfall series are shown graphically in Fig.3. The fact that the annual and Southwest monsoon rainfall is highly correlated reflects the time series. The mean seasonal precipitation values are 2.12, 5.74, 1093.50 and 41.04 mm, respectively, in winter, pre-monsoon, southwest monsoon and post-monsoon. Time-series analysis shows that there might be the lowest precipitation of the century in 1917-1918, 1973-1974 and 2011-2012. And in 1975-1976 and 2004-2005 show the highest precipitation from the year 1901 to 2018.

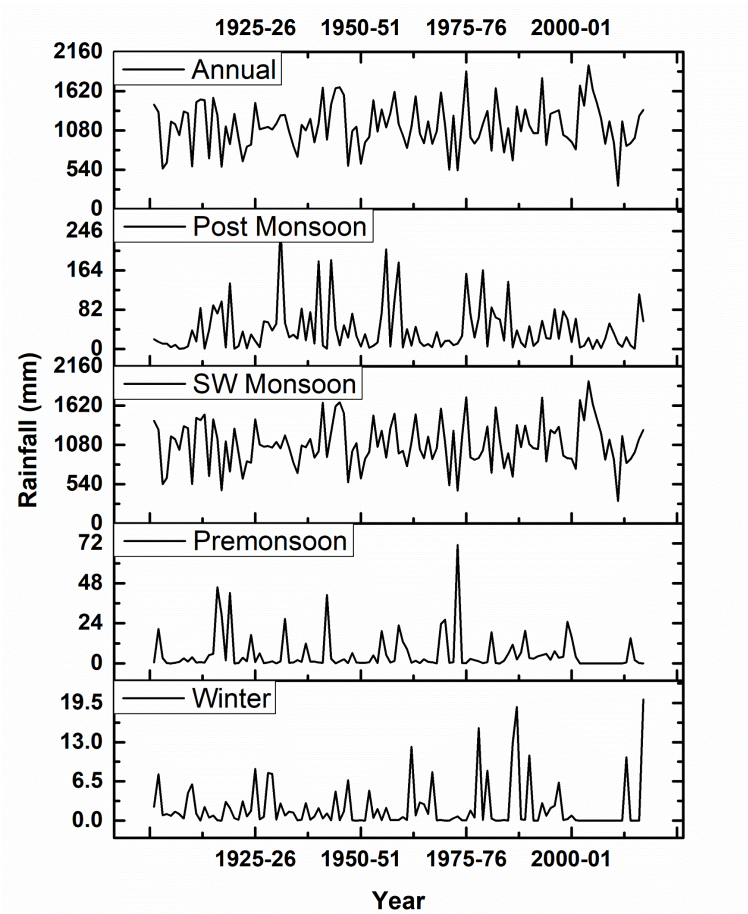

Fig. 3. Time Series of Monthly and Seasonal Rainfall in Surat District (1901 to 2018).

3.2. Trend analysis and SPI value

The Standardized Precipitation Index (SPI) is calculated for various time scales. The SPI for one month is not considered herein study as it does not reveal much information about the dry and wet periods. Usually, the rainfall starts in the Surat district and south Gujarat region in June and ends in September or October. Sometimes, the region also experiences rainfall in the month of November and December. Therefore, the 3-month (June-Aug) and 4-month (June-Sept) SPI represent the moisture conditions in the Southwest monsoon season. Although, the 6-month SPI (June-Nov) shows trends in the post-monsoon season, which can also be said seasonal rainfall. At the same time, the 12-month SPI (June-May) indicates yearly rainfall, including winter and pre-monsoon season. Here the SPI value representation is framed in (Fig. 4) for the last 118 years.

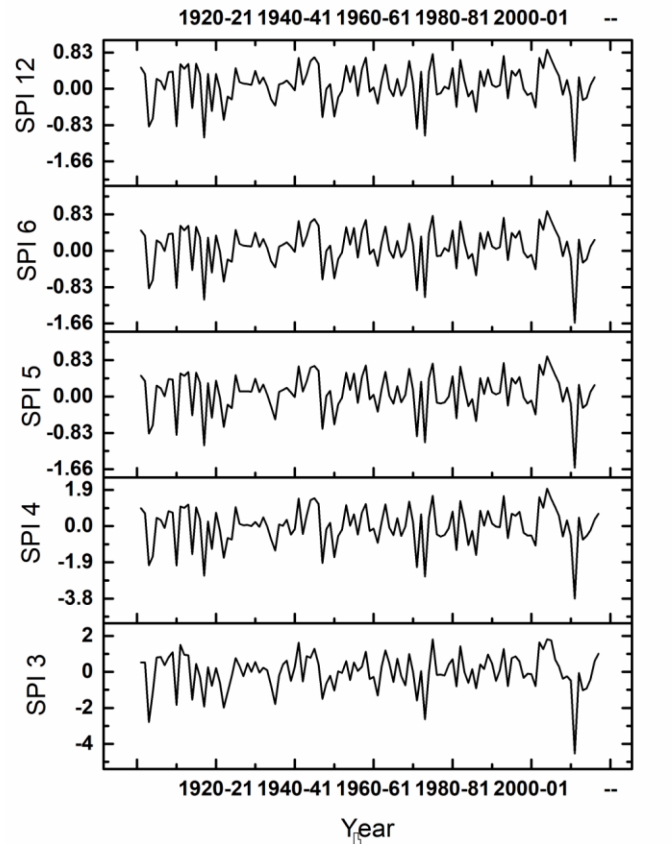

Fig. 4. Timeseries of SPI values at different time scales.

Table 4
Rainfall classification.

Condition	SPI @			
	3-month	4-month	5-month	6-month
Extreme drought	1924,1949,1975,2013	1926,1972,1974,1985 ,2012	-	-
Severe drought	1937,1973	1924,1949,1952,1988	2013	2013
Moderate Drought	1924,1951,1962,2014	1920,1936,1982,2002	1918,1974	1918,1974
Moderate Wet	1946,1959,1964,1970 ,1983,1994,2004,201 6	1942,1945,1946,1947 ,1954,1959,1964,197 0,1983,2006,2007	-	-
Very Wet	1942,1976,2003,2005 ,2006	1976,1994,2003,2005	-	-
Near Normal	Remaining years	Remaining years	Remaining years	Remaining years

The years involved in the different rainfall conditions are furnished in Table 3. Apparently, many of the years are falling under the normal rainfall category. Despite that, the years 1923, 1926, 1948, 1951, 1972, and 1987 have faced extremely dry conditions in the Southwest monsoon season. Moreover, the severely dry condition was observed in 1923, 1936, 1948, 1951, 1972, and 1978 during the Southwest monsoon, and the Years 2011 and 2012 have undergone post-monsoon season. On the contrary, many years have also experienced moderately wet and very wet conditions, which appear in Table 4.

3.3. Forecasting of rainfall using ANFIS

In this section, the forecast results of the six models are present. The results are from training and validation data sets for SPI 5 and SPI 6. The forecasted values for SPI are also statistically evaluated and presented in this section. The adaptive neuro-fuzzy inference system was developed using the Fuzzy Logic Toolbox graphical user interface (GUI) tools in MATLAB R2014a. SPI values from the previous time scale, a combination of previous SPI and antecedent rainfall conditions, and antecedent rainfall were used as input data for SPI 5 and SPI 6 prediction (Fig. 5).

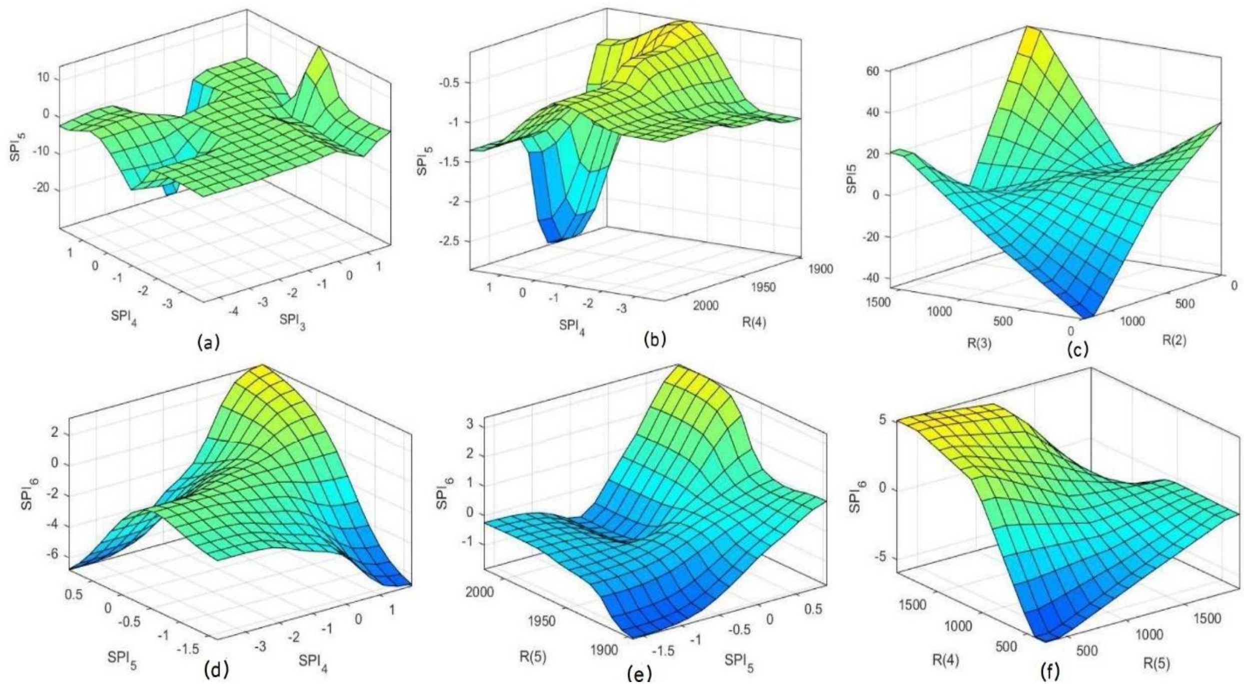

Fig. 5. Surface viewer of SPI 5 with input as (a) SPI value (b) SPI and rainfall (c) rainfall; SPI 6 with input as (d) SPI value (e) SPI and rainfall (f) rainfall.

The Surface Viewer displayed the mapping from various inputs to a single output for SPI value prediction. When the results of the ANFIS models are examined, it confirms that models comprised of precipitation values from the previous time step perform worse than the other models. The findings of models that employ SPI demonstrate that they perform better than others. When just the precipitation values are included in the models, the findings show that in the training data set, the model is showing good performance; however, the application of this model on the checking dataset shows a comparatively higher value of error. The M2 model, whose output is SPI 5, performed the best (Table 5). For the validation set, the RMSE value is 0.059, the R^2 value is 0.987, and the value of the coefficient of determination is 0.993. In the case of SPI 6, the M1 model has performed best. For the validation set the RMSE value is 0.042, R^2 value is 0.992, and the value of the coefficient of determination is 0.996.

Table 5
Performance analysis of ANFIS models.

Forecasted Entity	Model Type	Model Type	Nos. of Membership Function	Model Performance Evaluation Parameter					
				Training Set			Testing Set		
				r	R^2	RMSE	r	R^2	RMSE
SPI 5	M4	gbellmf	6	0.954	0.909	0.044	0.982	0.964	0.089
	M2	gauss2mf	3	0.992	0.984	0.034	0.993	0.987	0.059
	M6	trimf	6	0.979	0.959	0.011	0.659	0.434	0.524
SPI 6	M3	gaussmf	3	0.867	0.752	0.225	0.997	0.994	0.043
	M1	gbellmf	3	0.999	0.999	0.015	0.996	0.992	0.042
	M5	gauss2mf	6	0.967	0.935	0.006	0.584	0.341	0.452

Fig. 6 and Fig. 7 show the linearity between observed and forecasted SPI 5 and SPI 6 values in the training and validation period. Model M1 and M2 both have inputs rainfall and SPI values. Other models have either rainfall value or SPI value as input. From statistical evaluation, it has been seen that the ANFIS model forecasts better if the input is rather than the SPI value. If only SPI or rainfall value is taken as input, the ANFIS model does not give good results. With rainfall and SPI values, it can give promising results. Fig. 6 shows a Comparison of Observed and Forecasted SPI 5 Values in the validation Period 2003-2016. Fig. 7 shows a Comparison of Observed and Forecasted SPI 6 Value in the validation Period 1982-2016. From both figures, it has been proved that the ANFIS technique is good for predicting SPI values.

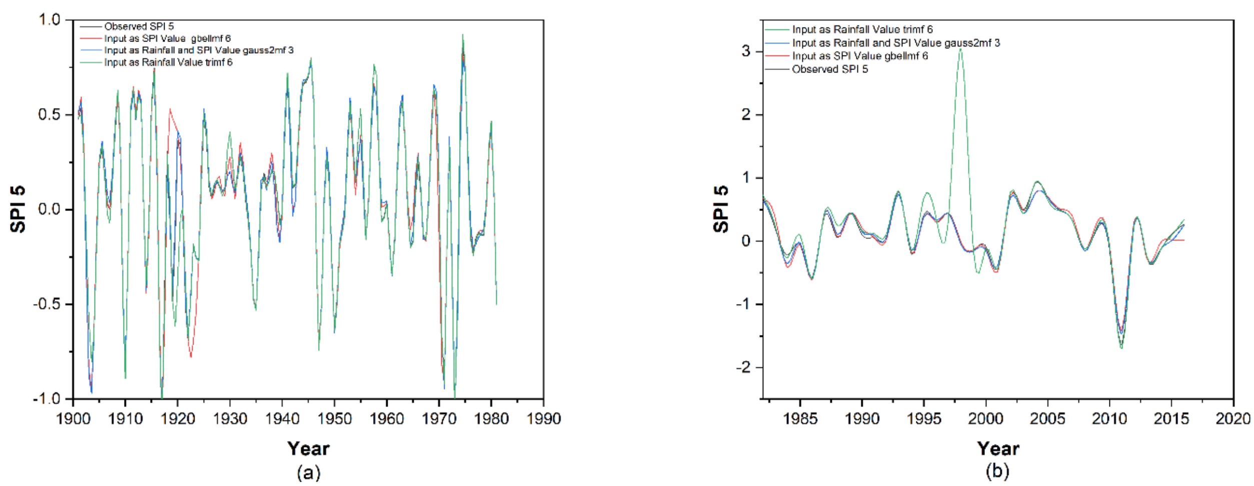

Fig. 6. Comparison of the observed and predicted value of SPI 5 (a) in the training dataset (b) in the testing dataset.

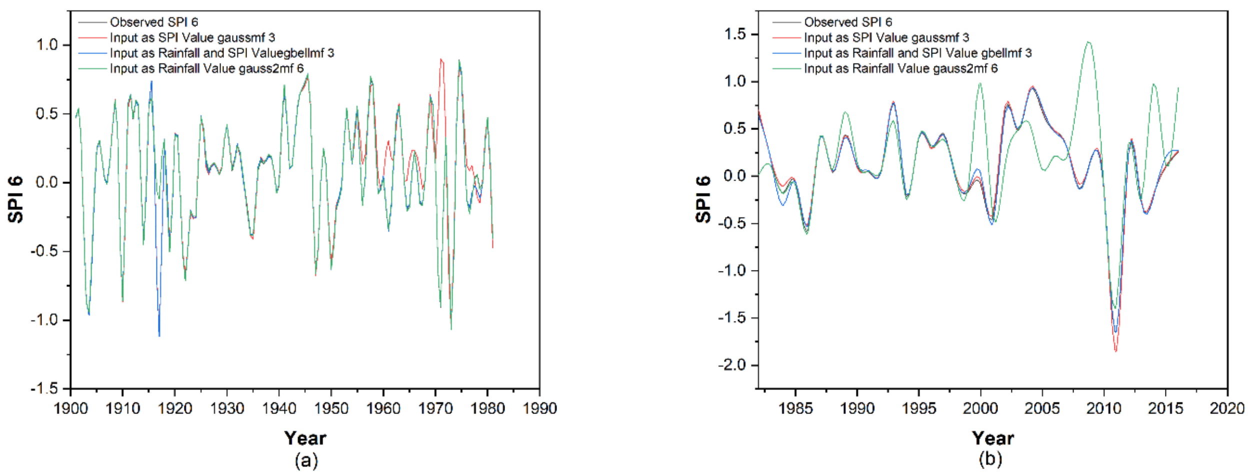

Fig. 7. Comparison of the observed and predicted value of SPI 6 (a) training dataset (b) testing dataset.

4. Conclusions

In the present study integrated SPI- ANFIS approach is being utilized to forecast the rainfall condition in the Surat region, Gujarat. SPI index is a widely used and proven index to understand the rainfall condition in any region. The SPI values are calculated analysis points out that there was a phenomenon of extreme and severe drought conditions in the study area in the last 117 years. Therefore, it was needed to understand the future trend of rainfall.

Moreover, SPI values are predicted using different models using the ANFIS technique in the research. The authors designed a fuzzy logic model using the Sugeno fuzzy inference system. Different ANFIS forecasting models for SPI-5 and SPI-6 were trained and tested. When the results of the ANFIS models are compared, it is seen that only the performances of models composed of precipitation values belonging to the previous time step are lower than the performances of the other models. Evaluation and performance of the ANFIS model show that the best result for each entity SPI 5 and SPI 6 is gained by only SPI value and a combination of both SPI and Rainfall Value as an input parameter, respectively. It is better not to use the "only Rainfall" value as an input parameter for forecasting SPI 5 and SPI 6. ANFIS models are simple to construct, and computation time is also less.

All in all, it's worth mentioning that sensible water resource planning and management entail forecasting future events while considering that most predictions are based on previous events. The present research will assist policymakers, government authorities, and engineers in building strong policies on water conservative measures and solving different hydrological problems.

Funding

This research received no external funding.

Conflicts of interest

The authors declare no conflict of interest.

References

[1] Gowri L, Manjula KR, Sasireka K, Deepa D. Assessment of Statistical Models for Rainfall Forecasting Using Machine Learning Technique 2022;2:51–67.

[2] Poornima S, Pushpalatha M. Prediction of Rainfall Using Intensified LSTM Based Recurrent Neural Network with Weighted Linear Units. Atmosphere (Basel) 2019;10:668. https://doi.org/10.3390/atmos10110668.

[3] Bacanli UG, Firat M, Dikbas F. Adaptive Neuro-Fuzzy Inference System for drought forecasting. Stoch Environ Res Risk Assess 2009;23:1143–54. https://doi.org/10.1007/s00477-008-0288-5.

[4] Palanichamy J, Palani S, Anita Hebsiba G, Viola J, Tungsrimvong A, Babu B. Simulation and Prediction of Groundwater Quality of a Semi-Arid Region Using Fuzzy Inference System and Neural Network Techniques. J Soft Comput Civ Eng 2022;6:110–26. https://doi.org/10.22115/SCCE.2022.285106.1314.

[5] Barua S, Ng AWM, Perera BJC. Artificial Neural Network–Based Drought Forecasting Using a Nonlinear Aggregated Drought Index. J Hydrol Eng 2012;17:1408–13. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000574.

[6] Belayneh A, Adamowski J, Khalil B, Ozga-Zielinski B. Long-term SPI drought forecasting in the Awash River Basin in Ethiopia using wavelet neural network and wavelet support vector regression models. J Hydrol 2014;508:418–29. https://doi.org/10.1016/j.jhydrol.2013.10.052.

[7] Hosseini-Moghari S-M, Araghinejad S, Azarnivand A. Drought forecasting using data-driven methods and an evolutionary algorithm. Model Earth Syst Environ 2017;3:1675–89. https://doi.org/10.1007/s40808-017-0385-x.

[8] Choubin B, Malekian A, Golshan M. Application of several data-driven techniques to predict a standardized precipitation index. Atmosfera 2016;29:121–8. https://doi.org/10.20937/ATM.2016.29.02.02.

[9] Deo RC, Kisi O, Singh VP. Drought forecasting in eastern Australia using multivariate adaptive regression spline, least square support vector machine and M5Tree model. Atmos Res 2017;184:149–75. https://doi.org/10.1016/j.atmosres.2016.10.004.

[10] Mokhtarzad M, Eskandari F, Jamshidi Vanjani N, Arabasadi A. Drought forecasting by ANN, ANFIS, and SVM and comparison of the models. Environ Earth Sci 2017;76:729. https://doi.org/10.1007/s12665-017-7064-0.

[11] Zhang R, Chen Z-Y, Xu L-J, Ou C-Q. Meteorological drought forecasting based on a statistical model with machine learning techniques in Shaanxi province, China. Sci Total Environ 2019;665:338–46. https://doi.org/10.1016/j.scitotenv.2019.01.431.

[12] Khan MMH, Muhammad NS, El-Shafie A. Wavelet based hybrid ANN-ARIMA models for meteorological drought forecasting. J Hydrol 2020;590:125380. https://doi.org/10.1016/j.jhydrol.2020.125380.

[13] Zhang Y, Yang H, Cui H, Chen Q. Comparison of the Ability of ARIMA, WNN and SVM Models for Drought Forecasting in the Sanjiang Plain, China. Nat Resour Res 2020;29:1447–64. https://doi.org/10.1007/s11053-019-09512-6.

[14] Fung KF, Huang YF, Koo CH, Soh YW. Drought forecasting: A review of modelling approaches 2007–2017. J Water Clim Chang 2020;11:771–99. https://doi.org/10.2166/wcc.2019.236.

[15] Gandhi FR, Patel JN. Groundwater potentiality deciphering and sensitivity study using remote sensing technique and fuzzy approach. Acta Geophys 2022;70:265–82. https://doi.org/10.1007/s11600-021-00711-5.

[16] Patel JN, Gandhi FR. Statistical temporal analysis of trend and variability of rainfall in Surat district, Gujarat, India. Int J Hydrol Sci Technol 2022;1:1. https://doi.org/10.1504/IJHST.2022.10044459.

[17] CGWB. CGWB Surat District 2013;492001.

[18] McKee TB, Doesken NJ, Kleist J. The relationship of drought frequency and duration to time scales. Proc. 8th Conf. Appl. Climatol., vol. 17, Boston; 1993, p. 179–83.

[19] Rezaei E, Karami A, Yousefi T, Mahmoudinezhad S. Modeling the free convection heat transfer in a partitioned cavity using ANFIS. Int Commun Heat Mass Transf 2012;39:470–5. https://doi.org/10.1016/j.icheatmasstransfer.2011.12.006.

[20] Hayati M, Rashidi AM, Rezaei A. Prediction of grain size of nanocrystalline nickel coatings using adaptive neuro-fuzzy inference system. Solid State Sci 2011;13:163–7. https://doi.org/10.1016/j.solidstatesciences.2010.11.007.