

Research Article

Performance Evaluation of Two ANFIS Models for Predicting Water Quality Index of River Satluj (India)

Sharad Tiwari ¹, Richa Babbar ², and Gagandeep Kaur¹

¹Department of Electrical and Instrumentation Engineering, Thapar Institute of Engineering & Technology, Patiala, Punjab 147004, India

²Department of Civil Engineering, Thapar Institute of Engineering & Technology, Patiala, Punjab 147004, India

Correspondence should be addressed to Richa Babbar; richa.babbar@thapar.edu

Received 18 August 2017; Revised 30 January 2018; Accepted 7 February 2018; Published 20 March 2018

Academic Editor: Giuseppe Oliveto

Copyright © 2018 Sharad Tiwari et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Water quality index is the most convenient way of communicating water quality status of water bodies, but its evaluation requires subjectivity in terms of user involvement and dealing with uncertainty. Recently, artificial intelligence algorithms that are appropriate for nonlinear forecasting and also dealing with uncertainties have been applied to various domains of water quality forecasting. This paper focuses on development of a data-driven adaptive neurofuzzy system for the water quality index using a real data set obtained from eight different monitoring stations across River Satluj in northern India. Novelty in the paper lies in the estimation of water quality index using two different clustering techniques: fuzzy C-means and subtractive clustering-based ANFIS and assessing their predictive accuracy. Each model was used to train, validate, and test the index that was obtained from seven water quality parameters including pH, conductivity, chlorides, nitrates, ammonia, and fecal coliforms. The models were evaluated on the basis of statistical performance criteria. Based on the evaluations, it was found that the SC-ANFIS method gave more accurate result as compared to the FCM-ANFIS. The tested model, SC-ANFIS model, was further used to identify those sensitive parameters across various monitoring stations that were capable of causing change in the existing water quality index value.

1. Introduction

Surface water quality evaluation is an issue that draws the attention of regulatory agencies time and again for the purpose of safeguarding various intended uses. In this regard, continuous water quality monitoring is undertaken so as to assess the water quality and thereof propose adequate measures for its management. One of the several ways in which the large amount of data generated from the monitoring stations is assimilated for easy communication to the stakeholders is water quality index (WQI). The water quality index is a single number that expresses water quality by aggregating the measurements of several water quality parameters. Its output ranges from 0 to 100. A value of 100 represents excellent water quality condition, while zero indicates water not suitable for the intended use.

Several advancements are seen in evaluating WQI, ever since Horton [1] proposed the first index in the year 1965. These advancements have mainly come in the form of using

soft computing tools such as data mining techniques, artificial intelligence, and fuzzy modeling system. The techniques have the capability to take care of uncertainties that often accumulate in a traditional way of evaluating WQI [2, 3]. Babbar and Babbar [3] employed different data mining techniques such as support vector machines, k nearest neighbor, decision trees, and naïve bayes to develop the predictive environment to classify water quality into understandable terms based on the overall index of pollution. Various researchers have successfully employed AI techniques to evaluate the water quality status [4–13]. However the choice of methods depends on the quantity and quantity of the data available. Fuzzy logic methods on the other hand are heuristic system description that uses if-then rules to establish quantitative relationships among the input and output variables [14]. However, the main problem with the fuzzy model is that there is no systematic procedure to define the membership function parameters, which must be predetermined by expert knowledge about the modeled system. At the same time, artificial neural

network (ANN) has the ability to learn from an input and output pair and adapt to it in an interactive way. Based on these understanding, a hybrid technique adaptive neuro-fuzzy inference system (ANFIS) is gaining popularity in dealing with ill-defined and uncertain domains such as water quality predictions. ANFIS is a technique that embeds the fuzzy inference system into the framework of adaptive networks. ANFIS thus draws the benefits of both ANN and fuzzy techniques in a single framework. One of the major advantages of the ANFIS method over fuzzy systems is that it eliminates the basic problem of defining the membership function parameters and obtaining a set of fuzzy if-then rules. The learning capability of ANN is used for automatic fuzzy if-then rule generation and parameter optimization in ANFIS [15]. Jang [16] introduced the concept of ANFIS and since then it has been successfully implemented in various water quality problems. Yan et al. [17] developed the ANFIS model for classifying the water quality status of river and compared its performance with ANN. In this study, different types of membership functions such as generalized bell, Gaussian, trapezoidal, and triangular were compared and tested for training and testing data. The best-fit model was obtained using the Gaussian membership function. Sahu et al. [18] employed ANFIS to predict WQI of groundwater.

Rahimzadeh et al. [19] suggested ANFIS approach for the prediction of oily wastewater microfiltration permeate volume. Similarly, Talebizadeh and Moridnejad [20] compared ANN with ANFIS in forecasting lake level fluctuations, where ANFIS turned out to be superior to ANN in terms of efficiency. For water quality prediction, ANFIS has been applied for biochemical oxygen demand estimation based upon other water quality parameters as inputs [21]. Despite prediction model improvements, increase in model accuracy while avoiding overfitting is still a challenge for most researchers. According to recent researches, model performance can be significantly improved if an appropriate hybrid of multiple models is used for forecasting and prediction than using a single model in this regard [22].

As is apparent from the above studies, ANFIS has its popularity in terms of its applications and is also well established in the literature, but the choice of algorithms in automatically generating membership function is not given due considerations. Therefore, this study presents the automatic generation of membership function for WQI using fuzzy C-means (FCM) and subtractive clustering- (SC-) based ANFIS model which yields the homogeneous clusters (or classes) of WQI. The clustering method is one of the important methods for data analysis and decision making that allows retrieval of the useful information by grouping or categorizing multidimensional data in clusters. Since water quality data are a multidimensional data, it is expected that by performing two different clustering techniques a model performance can be evaluated and compared.

Thus the present study aims to develop the ANFIS model based on two different clustering methods and statistically identify the best amongst the two. The second objective is to employ the identified best model to find out the most sensitive water quality parameter that can cause a change in the predicted water quality.

2. Materials and Methods

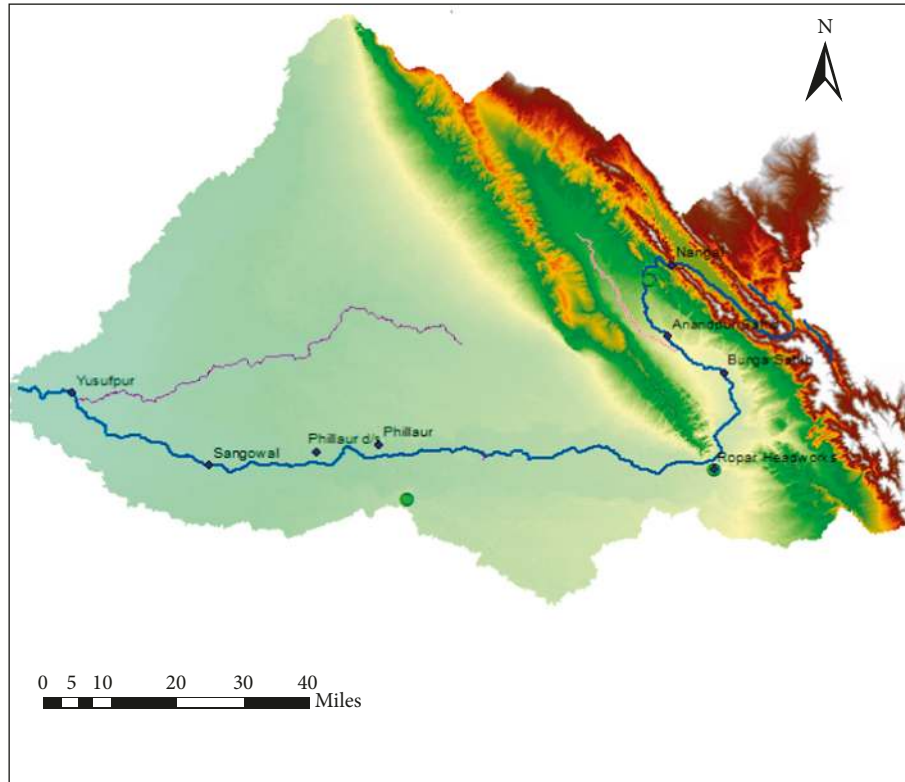
2.1. Study Area. River Satluj is one of the five rivers of Indus River system that joins Indus River on its eastern side. It originates from the *Manasarovar-Rakas* Lakes in western Tibet at a height of 4,570 m within 80 km of the source of the Indus. In its course of travel, it flows through the Himalayan range in the Indian state of Himachal Pradesh and enters in the plains of Punjab from the Shivalik hills near Nangal, India. The river carries historical importance because of the fact that its water allocation is a part of famous Indus water treaty between India and Pakistan and also world's highest gravity dam, Bhakra Nangal Dam, is built across the river at a point where it enters Punjab (India). Satluj River is extensively used for irrigation as well as drinking purposes. The hydrology of the river is controlled by snowmelt from Himalayas and South Asian monsoons.

This study focuses 238 km stretch of Satluj River flowing in the Punjab state of India. The watershed area of the river in this stretch is about 10,880 km² and is geographically bounded between 31° 45' N, 74° 57' E and 30° 45' N, 76° 50' E. The river is regularly monitored at 8 monitoring stations by Thapar University, Patiala (India), under the river monitoring program initiated by the Government of India in 1996. These monitoring stations are strategically located for comprehensive study on water quality of the river in Punjab (India). The locations of the monitoring stations along the main river are shown in Figure 1.

Studies have shown that the water quality U/S of Headwork Nangal is of "A" class or of pristine quality. But further downstream, the river quality deteriorates in the study stretch area. This deterioration in the river quality is because of release of toxic effluents from industries and untreated sewage discharges, making river water unfit for any beneficial use.

2.2. Water Quality Parameters. The water quality parameters that are routinely monitored at eight monitoring sites are pH, conductivity, chlorides, dissolved oxygen (DO), 5-day biochemical oxygen demand (BOD₅), total dissolved solids (TDS), suspended solids (SS), ammonical-N, nitrates, total phosphorous (TP), and fecal coliform (FC). All these parameters are monitored on monthly basis and analyzed according to the APHA standard methods [23]. For the purpose of this study, the historic water quality data from the year 1996–2012 are taken into consideration. For each station and for each parameter, the month wise 16-year average concentration was computed as given in Table 1. From Table 1, it is visible that water quality deteriorates in terms of DO (% sat.), BOD, and fecal coliform from station 6. This is because confluence of small rivulets: *Budha Nallah* and *Chitti Bein* happens after station 6. Both these rivulets discharge a large quantity of industrial and domestic effluent in River Satluj.

2.3. Water Quality Index. The formulation for the water quality index for River Satluj is given in the study of Sharma and Reddy [24]. In it, the calculations for water quality index



- ◆ Monitoring stations
- Cities

FIGURE 1: Satluj River and its monitoring stations.

TABLE 1: Average values of water quality parameters along River Satluj.

	Station 1	Station 2	Station 3	Station 4	Station 5	Station 6	Station 7	Station 8
pH	7.70	8	7.88	7.91	7.91	7.88	7.57	7.70
Conductivity (μ S/cm)	3.22	0.23	0.21	0.21	0.23	0.23	0.38	0.45
Chlorides (mg/l)	7.98	10.52	9.16	10.51	12.75	12.99	35.98	41.79
DO (% sat.)	81.41	87.51	89.50	88.66	82.63	81.72	49.26	52.72
BOD ₅ (mg/l)	1.24	1.59	1.31	1.67	3.62	4.05	19.19	18.22
SS (mg/l)	52.67	64.42	60.49	100.58	136.24	138.43	125.44	153.47
TDS (mg/l)	131	147.38	139.88	140.72	161.69	167.89	276.72	276.32
Ammonical-N (mg/l)	0.31	0.35	0.35	0.43	0.41	0.61	4.09	0.84
Nitrates (mg/l)	2.39	5.51	1.15	3.52	2.98	1.07	0.98	1.70
TP (mg/l)	0.33	0.26	0.23	0.32	0.39	0.40	0.59	0.74
FC (MPN/100 ml)	720.41	519.51	321.48	591.02	469.17	1566.95	1198058.28	237681.12

were performed taking into consideration three different water uses: ecological, irrigation, and municipal and domestic use. The steps involved included (a) selection of parameters, (b) assignment of weights to the selected parameter, (c) transformation of the monitored parameter values into common environmental scale units through use of parameter rating curves/equations (PRC/E) as given in Table 2, and (d) aggregation of the parameter values into the final score.

On the basis of the final score obtained, the water quality is classified into six classes with highest quality having maximum score as very poor (<45), poor (45–60), fair (61–69), good (70–79), very good (80–90), and excellent (91–100).

Mathematically, the WQI is expressed as

$$WQI = \sum_{i=1}^n W_i Q_i, \tag{1}$$

where Q_i is the subindex for i th water quality parameter, W_i is the weighting factor of the weight associated with each water quality parameter based, and n is the number of water quality parameters.

In this study, WQI computed was meant for municipal and domestic water use. The parameters considered in this particular category were pH, conductivity, chlorides, TDS, ammonia-N, nitrates, and fecal coliform. The selection of the

TABLE 2: Transformation equations for selected water quality parameters (source: Sharma and Reddy [24]).

Parameters	Concentration, x (range)	Transformation equation
pH	2.7–5.0	$y = 6.889x^2 - 40.00x + 57.79$
	5.0–7.5	$y = -10.66x^2 + 161.9x - 513.8$
	7.5–9.5	$y = -22.33x^2 + 346.3x - 1242.8$
	9.5–12.0	$y = 2.666x^2 - 69.33x + 448$
Conductivity ($\mu\text{S/cm}$)	0–4500	$y = 8E - 07x^2 - 0.025x + 104.1$
TDS (mg/L)	50–3000	$y = 6E - 06x^2 - 0.046x + 96.90$
Ammonical-N (mg/L)	0–2	$y = -7x^2 - x + 100$
	2–40	$y = -20.5 \ln(x) + 84.61$
Nitrate-N (mg/L)	0–23	$y = 0.161x^2 - 7.953x + 104.5$
Chloride (mg/L)	0–1200	$y = 7E - 05x^2 - 0.157x + 99.60$
Fecal coliform (MPN/100 ml)	0–1000	$y = 80.702e^{-0.001x}$
	1000–100000	$y = -4.7772 \ln(x) + 63$

parameters was based on the fact that these parameters are considered useful for characterizing municipal and industrial waste, are routinely monitored for any polluted river, and therefore, form a small set of parameters that sufficiently convey information of the overall water quality for designated uses. More parameters can be used for estimation of WQI; however, more the number of parameters, more is the logistic concern for their monitoring and analysis. Since every parameter has its own importance with regard to an expression for water quality, there comes the need of ranking them in terms of their relative importance over the other parameter by assigning them a weighting factor which while formulating the index was obtained from expert opinion analysis given in detail in the study of Sharma and Reddy [24]. The transformation equations that were used to develop WQI for 16-year data in this study are given in Table 2. These transformation equations basically represent parameter rating curves that meant to bring different units of different parameters on a single scale unit for the purpose of direct aggregation in (1). These equations were developed on the basis of water quality standards and Indian effluent standards. In the equations given in Table 2, x represents the concentration range and y represents the scaled value of the parameter under consideration.

3. Description of ANFIS

ANFIS is a multilayer feed-forward network that uses neural network learning algorithms and fuzzy logic to map an input space to an output space [25]. Jang [16] suggested adaptive neuro-fuzzy inference system (ANFIS) to construct an input-output mapping based on the initial given fuzzy system and available input-output data pairs by using learning procedures. This system can achieve a highly nonlinear mapping and is superior to common linear methods in producing nonlinear time series [26]. In the process of mapping input space to an output space, two commonly employed fuzzy inference systems (FIS) are used in various

applications. These are Mamdani inference system and Sugeno inference system, which are described in the literature [27, 28]. The consequences of the fuzzy rules for these two inference models are different, and thus their aggregation and defuzzification procedures also differ accordingly. The Sugeno system is, however, considered more compact and computationally efficient than Mamdani's system [27]. The consequence parameter in Sugeno FIS is either a linear equation, called *first-order Sugeno FIS*, or constant coefficient, called *zero-order Sugeno FIS* [26].

If it is assumed that the system includes two inputs, x_1 and x_2 , and the output y and the rule base contains two fuzzy if-then rules; then the representation of rules for the first-order Sugeno FIS can be expressed as

$$\text{Rule 1: If } (x_1 = A_1) \text{ and } (x_2 = B_1) \text{ then} \\ f_1 = p_1 * x_1 + q_1 * x_2 + b_1, \quad (2)$$

$$\text{Rule 2: If } (x_1 = A_2) \text{ and } (x_2 = B_2) \text{ then} \\ f_2 = p_2 * x_1 + q_2 * x_2 + b_2,$$

where p_i , q_i , and r_i are the linear parameters in the consequent part of the Sugeno fuzzy inference system.

The architecture of ANFIS consists of five layers. Each layer contains several nodes described by the node function. Adaptive nodes, denoted by squares, represent the parameter sets that are adjustable in these nodes, whereas fixed nodes, denoted by circles, represent the parameter sets that are fixed in the system. The output data from the nodes in the previous layers are the input in the present layer. The description of each layer in the ANFIS architecture is given below.

Layer 1 is the fuzzification layer in which each node represents membership grade of the crisp inputs and each nodes output θ_i^1 is computed by

$$\theta_i^1 = \mu_{A_i}(x_1), \quad i = 1, 2, \\ \theta_i^1 = \mu_{B_{i-2}}(x_2), \quad i = 3, 4, \quad (3)$$

where x_1 and x_2 are the crisp input to the node i , A_i and B_i are the linguistic labels characterized by the proper membership functions μ_A and μ_B , respectively. The Gaussian membership function is given by

$$\mu_{A_i}(x_1) = e^{-((x_1 - b_i)^2 / 2a_i^2)}, \\ \mu_{B_i}(x_2) = e^{-((x_2 - b_i)^2 / 2a_i^2)}, \quad (4)$$

where $\{a_i, b_i\}$ are the parameter set of the membership function in the premise part of fuzzy if-then rules that modify the shapes of membership functions. The parameter in the input layer is called the premise parameter.

In layer 2, each node provides the strength of rules by means of multiplication operator given in (5). The output of this layer is firing strength θ_i^2 as the products of the corresponding degree obtained from layer 1.

$$\theta_i^2 = \omega_i = \mu_{A_i}(x_1)\mu_{B_i}(x_2), \quad i = 1, 2. \quad (5)$$

The membership values represented by $\mu_{A_i}(x_1)$ and $\mu_{B_i}(x_2)$ are multiplied in order to find the strength of the rule

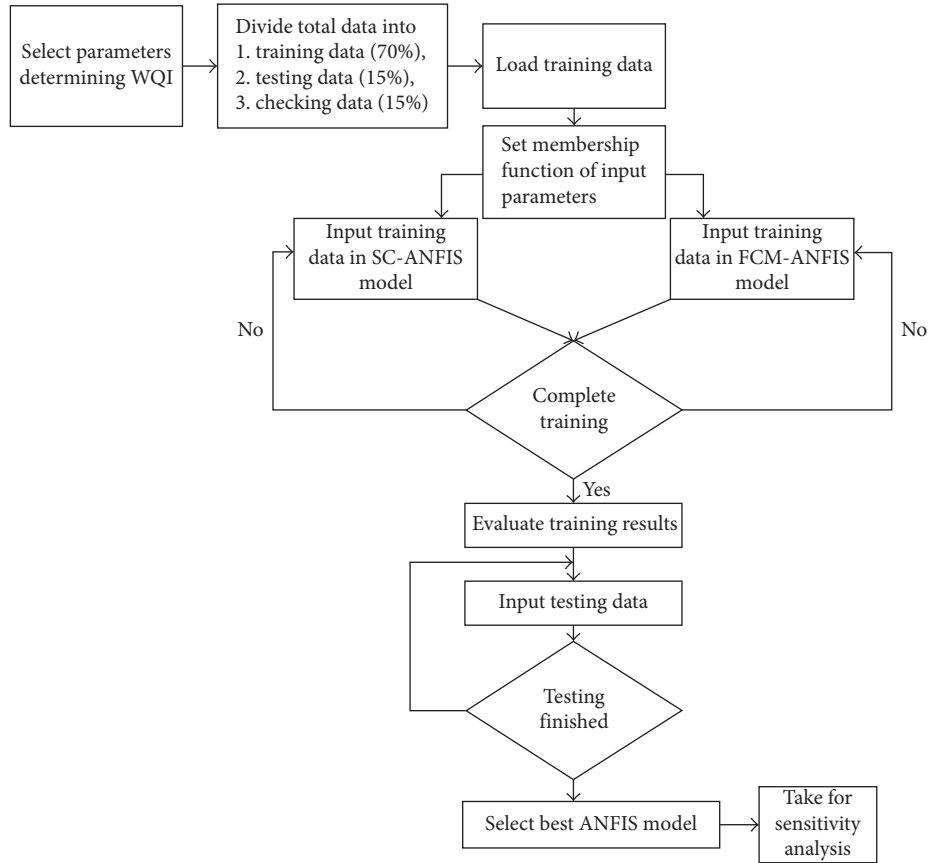


FIGURE 2: Flow chart depicting the proposed methodology.

where the variable x_1 has linguistic value A and x_2 has linguistic value B .

Layer 3 is the normalization node, which normalizes the strength of all rules according to the equation given below:

$$\theta_i^3 = \bar{\omega}_i = \frac{\omega_i}{\sum_i \omega_i}, \quad i = 1, 2. \quad (6)$$

Layer 4 is a layer of adaptive node, and every node in this layer computes the contribution of each i th rule towards the overall output and the function defined as

$$\theta_i^4 = \bar{\omega}_i f_i = \bar{\omega}_i (p_i * x_1 + q_i * x_2 + b_i), \quad i = 1, 2, \quad (7)$$

where $\bar{\omega}_i$ is the output of layer 3 and $\{p_i, q_i, b_i\}$ are the parameter set. Parameters in this layer are referred to as consequent parameters.

Layer 5 is the output layer in which the single node computes the overall output by summing all the rules from the previous layer. Accordingly, the defuzzification process transforms each rule's fuzzy results into a crisp output in this layer. The output, θ_i^5 , is computed as in (7).

$$\theta_i^5 = y = \sum_i \bar{\omega}_i f_i = \frac{\sum_i \omega_i f_i}{\sum_i \omega_i}. \quad (8)$$

ANFIS applies the hybrid-learning algorithm, which consists of the combination of "gradient descent" and "least-squares" methods to update the model parameters. Each epoch of this hybrid learning procedure is composed of

TABLE 3: Parameter values for different clustering-based ANFIS algorithms used.

Parameters for ANFIS by FCM algorithm	Parameters for ANFIS by SC algorithm
Membership function = Gaussian	Membership function = Gaussian
Number of clusters = 15	Number of clusters = 12
Number of rules = 15	Number of rules = 12
Number of exponent for the partition matrix = 3	Cluster radius = 0.3
	Squash factor = 1.7
	Accepted ratio = 0.5
	Rejected ratio = 0.15

a forward pass and a backward pass. In the forward pass of the hybrid learning procedure, the node output goes forward until layer 4 and the consequent parameters are identified by the least squares method. In the backward pass, the error signal propagates backwards and the premise parameters are updated by gradient descent. The detailed description of this algorithm is given in Jang and Sun [25].

4. Modeling with ANFIS

The proposed methodology for ANFIS application to water quality evaluation is shown in the form of a flow chart in Figure 2. In the current study, the input parameters of the

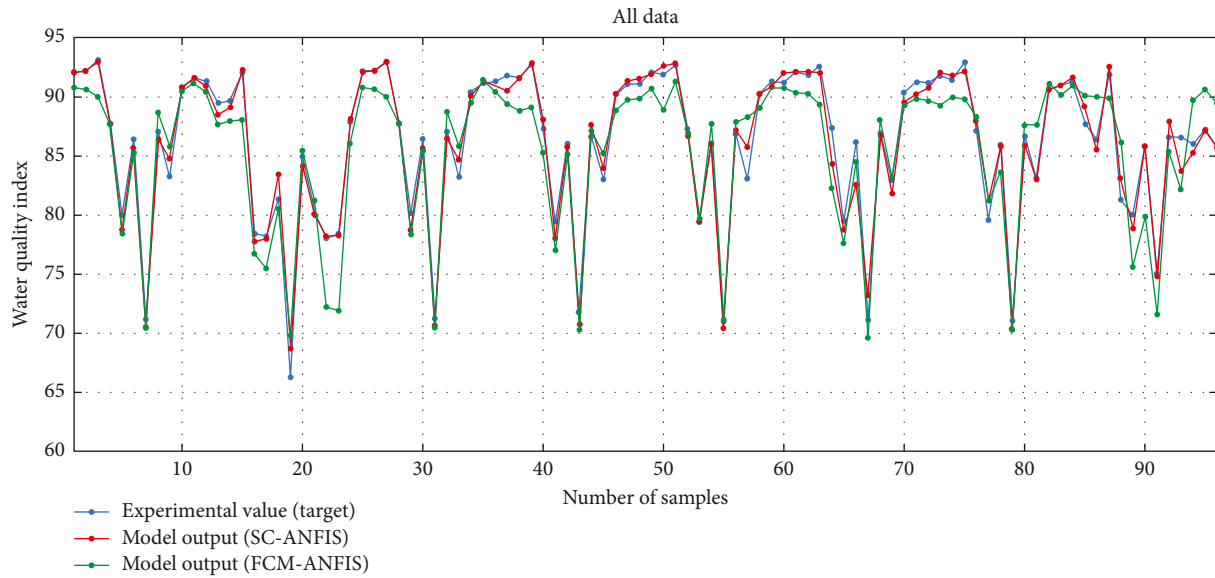


FIGURE 3: Comparison of measured and predicted WQI by FCM-ANFIS and SC-ANFIS.

ANFIS under consideration are pH, conductivity, TDS, chlorides, nitrates, ammonia and fecal coliforms and the output is the WQI. The data with selected water quality parameters were first converted into the number of principal components and then loaded as input into the ANFIS taking the Gaussian input parameter membership function.

The relationship between water quality index and input variables using ANFIS to generalize the relationship was of the form

$$\text{WQI} = f(\text{pH, conductivity, TDS, chlorides, nitrates, ammonia-N, fecal coliform}). \quad (9)$$

For the purpose of modeling, the water quality data obtained from eight monitoring stations equaling a total of 204 observations were divided into three sets: training data, checking data, and testing data. The training data were used for the training of ANFIS, while the checking data were used for verifying the identified ANFIS. The testing data were used to evaluate the model performance. The first data set, containing 70% of the records, was used as the training data; the second data set containing 15% of the records was used as the checking data, while the remaining 15% data were applied as the testing data. The target values represented the WQI computed using (1).

In this study, MATLAB Fuzzy Logic Toolbox ANFIS GUI was used as a modeling tool. Two separate clustering algorithms were used to automatically generate Gaussian-shaped membership functions. These algorithms are fuzzy *c*-means (FCM) and subtractive clustering (SC). Both these techniques generate fuzzy if-then rules but are different in terms of implementation. Fuzzy *c*-means (FCM) is a data clustering technique wherein each data point belongs to a cluster to some degree that is specified by a membership grade. Originally introduced by Bezdek [29], it provides a method that shows how to group data points that populate some multidimensional space into 12 different clusters.

TABLE 4: Water quality performance based on FCM and SC methods.

	SC-ANFIS	FCM-ANFIS
RMSE (test data)	1.6092	2.7743
RMSE (train data)	0.8701	1.1590
MSE (test data)	2.5896	7.6966
MSE (train data)	0.7570	1.3434
R^2 (train data)	0.9919	0.9852
R^2 (test data)	0.9827	0.9015
R^2 (all data)	0.9911	0.9821

On the other hand, if there is no clear idea about how many clusters there should be for a given set of data then subtractive clustering is useful. Subtractive clustering is a fast, one-pass algorithm for estimating the number of clusters and the cluster centers in a set of data [30]. The cluster estimates obtained can be used to initialize iterative optimization-based clustering methods and model identification methods like ANFIS. In this study, 15 cluster centers were determined for the given 204 data sets. The number of fuzzy rule set would be equal to the number of cluster centers, each representing the characteristic of the cluster. The details on various parameters, and their values taken for modeling with the two above clustering methods are given in Table 3.

For each algorithm defined above, ANFIS works on the model and tunes it by means of a hybrid technique combining gradient descent back propagation and mean least squares optimization algorithms. At each epoch, an error measure, which is the sum of the squared difference between actual and desired output, is reduced. When the values of the premise parameters were learned, the overall WQI was obtained as a linear combination of these parameters.

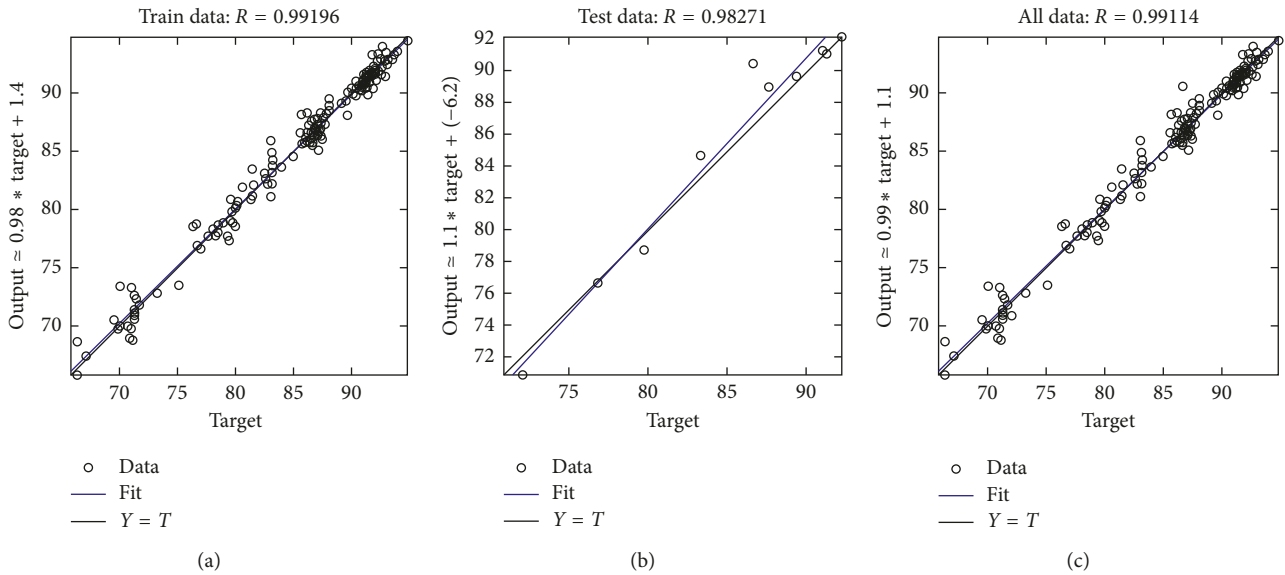


FIGURE 4: Fitness of the SC-ANFIS-WQI prediction model.

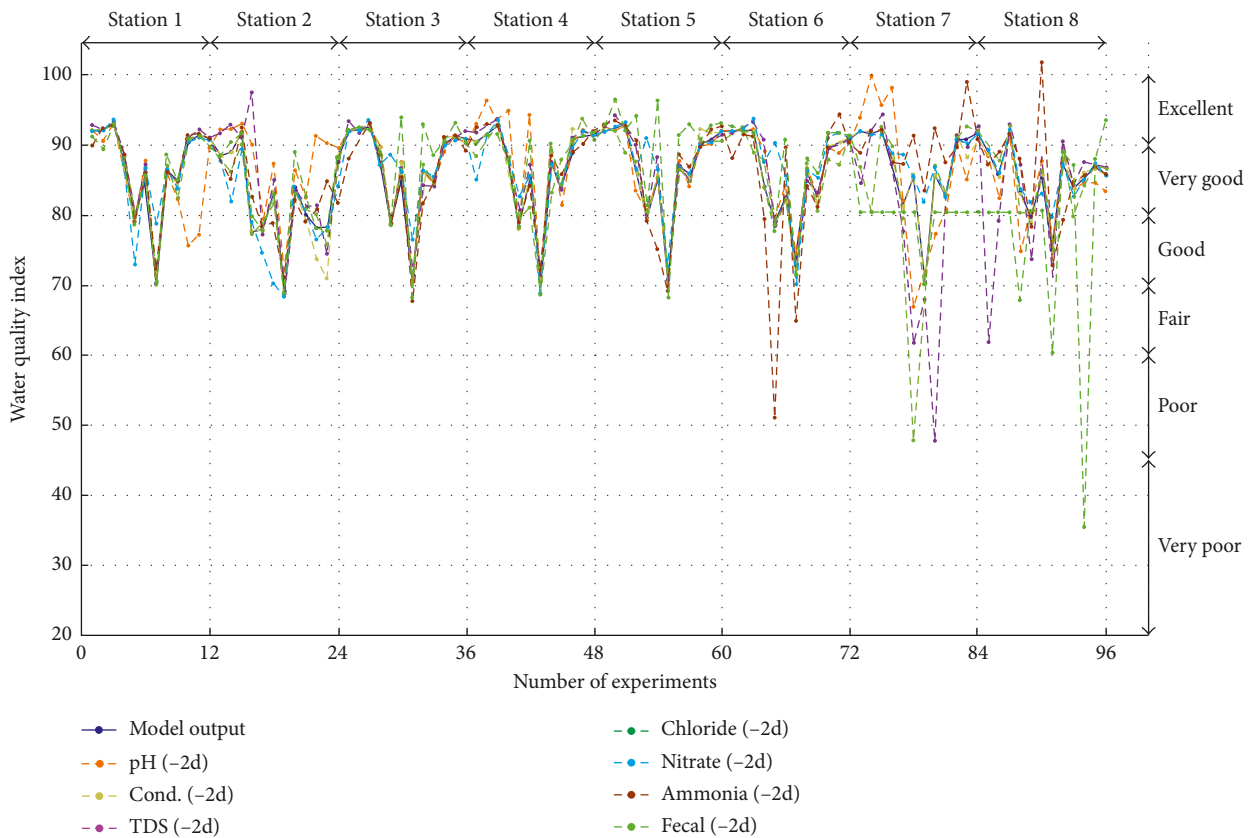


FIGURE 5: Sensitivity effect of -2d variation of each input on WQI.

The measured and the predicted WQI values over a range of records are shown in Figure 3. It is evident from this figure that the measured and predicted WQI SC-ANFIS values are in close agreement as compared to the FCM-ANFIS method.

In order to further verify the result, the performances of the two ANFIS models were evaluated according to different statistical criteria. The correlation coefficient (R^2), root mean square error (RMSE), and MSE were obtained using (10). The evaluated values are given in Table 4.

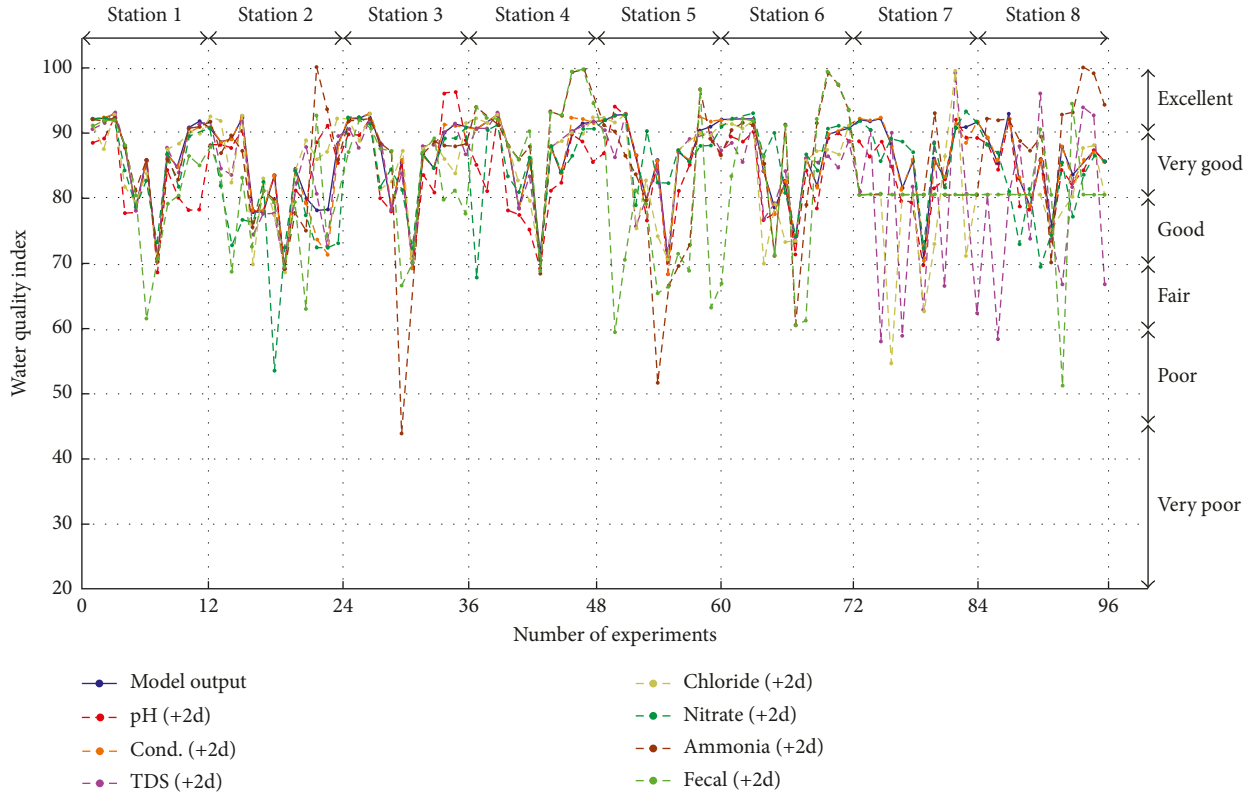


FIGURE 6: Sensitivity effect of +2d variation of each input on WQI.

$$R^2 = 1 - \frac{\sum_{i=1}^n (X_{\text{model}} - X_{\text{obs}})^2}{\sum_{i=1}^n (X_{\text{model}} - \overline{X_{\text{model}}})^2},$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=0}^n (X_{\text{obs}} - X_{\text{model}})^2}{n}}, \quad (10)$$

$$\text{MSE} = \sqrt{(X_{\text{obs}} - X_{\text{model}})^2},$$

where X_{obs} , X_{model} , and $\overline{X_{\text{model}}}$ are the observed value, modeled values, and average value, respectively.

From Table 4, it is found that the R^2 values for both training and testing data are 0.9919 and 0.98271, respectively. The RMSE and MSE values from Table 4 also indicate that the SC-ANFIS model predictions are very close to the experimental value as compared to the FCM-ANFIS model.

Figure 4 shows the overall fitness of the SC-ANFIS method, showing that the predicted WQI values were plotted against the measured ones. The WQI values lie around a straight line passing through the origin, which implies a very close agreement between the two.

5. Sensitivity Analysis

Sensitivity analysis is a process that helps us to find out how model output values are affected by changes in model input values. The purpose of performing sensitivity analysis was to

determine those parameters that can change the output value (WQI), to an extent that water quality class shifts from its existing class. For this, each input water quality parameter was perturbed by ± 2 times its standard deviation, and the changes in WQI were noted using the SC-ANFIS predictive model. The perturbation of ± 2 times the standard deviation was taken so as to account the variability of parameter and the associated influence on WQI. Finally, the resulting WQI was compared to the reference values. Figures 5 and 6 show the sensitivity effect of $-2d$ and $+2d$ variation of each input on WQI, respectively, with respect to eight monitoring stations.

The WQI changes its existing class when the perturbed parameter value shifts the WQI trend in a particular group of observation data shown on x -axis. In Figure 5, stations 6, 7, and 8 are visibly sensitive to different parameters. Station 6 is sensitive to ammonia, station 7 to TDS and chloride, and station 8 is sensitive to ammonia and fecal coliforms. The change in WQI is significantly high such that the existing water quality class shifts from good to poor; good to poor and good to very poor at stations 6, 7, and 8, respectively.

Similarly by studying the pattern of change in Figure 6, it is observed that fecal coliform is the most sensitive parameters and any perturbation in this parameter is switching the WQI of all stations. At station 8, the WQI changes from good to poor because of $+2d$ change in fecal coliform as compared to other stations where maximum change in water quality class is from good to fair. Chlorides and TDS are most sensitive parameters at station 7, and ammonia is the most sensitive parameter at station 3 and station 5.

The water quality parameters such as ammonia, chlorides, TDS, and fecal coliforms are the indicator parameters that are strongly linked to municipal or domestic waste. The sensitivity of Satluj River to these four parameters across its monitoring stations indicates that the river is under the influence of strong municipal waste, and fluctuations in these parameters should be given due considerations when communicating the water quality class.

6. Conclusions

In this study, two different clustering algorithms were used to develop the ANFIS model for water quality index prediction of River Satluj in northern India. The formulation for WQI of River Satluj is established in literature, and the index was computed by considering seven water quality parameters, crucial for municipal use. The data set for 16 years across eight monitoring stations on the river was used.

The two ANFIS models that were based on subtractive clustering and fuzzy c-means methods were trained, validated, and tested for modeling WQI. Based on the statistical evaluations, it was found that SC-ANFIS model predictions at training and testing stages were very close to the experimental value when compared to the FCM-ANFIS model.

The SC-ANFIS model, because of its good predictive capability over the FCM-ANFIS model, was further used to perform sensitivity analysis. The effect of perturbation in each water quality parameter was modeled and analyzed. The analyses showed that ammonia, chlorides, and fecal coliform were the most sensitive parameters that were capable of switching the existing water quality class to poorer class and hence warrant more attention in terms of their analysis for WQI computation.

The study reveals that ANFIS modeling with SC-ANFIS can be a useful approach to characterization of water quality in the form of water quality index. Since the approach obviates the otherwise lengthy computations of WQI, the present study holds its importance in developing a model and employing it for faster dissemination of information as well identifying the critical water quality parameters affecting WQI. The future scope of the work lies in the usage of a combination of hybrid SC-FCM and ANN (neurodynamic fuzzy expert system) to evaluate water quality.

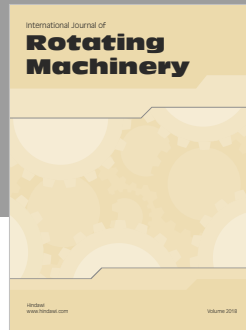
Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] R. K. Horton, "An index-number system for rating water quality," *Journal of the Water Pollution Control Federation*, vol. 37, no. 3, pp. 300–306, 1965.
- [2] K. Chau, "A review on integration of artificial intelligence into water quality modeling," *Marine Pollution Bulletin*, vol. 52, no. 7, pp. 726–733, 2006.
- [3] R. Babbar and S. Babbar, "Predicting river water quality index using data mining techniques," *Environmental Earth Sciences*, vol. 76, no. 14, p. 504, 2017.
- [4] N.-B. Chang, H. W. Chen, and S. K. Ning, "Identification of river water quality using the Fuzzy Synthetic Evaluation approach," *Journal of Environmental Management*, vol. 63, no. 3, pp. 293–305, 2001.
- [5] J. L. Jin, L. Liu, J. Ding, and Q. Fu, "Logistic curve model of groundwater quality evaluation," *Environmental Pollution Control*, vol. 25, no. 1, pp. 46–48, 2003.
- [6] T. Sun, S. B. Pan, and Y. J. Li, "Application of artificial neural network model to groundwater quality assessment and classification," *Hydrogeology and Engineering Geology*, vol. 3, pp. 58–61, 2004.
- [7] L. K. Wang, B. H. Men, Q. Fu, and T. G. Lu, "Matter element model for evaluating the water quality pollution," *System Sciences and Comprehensive Studies in Agriculture*, vol. 20, no. 2, pp. 96–98, 2004.
- [8] L. Zhang, G. X. Wang, and M. R. Luo, "Application of an improved grey association analysis method in water environmental quality assessment," *Acta Scientiarum Naturalium Universitatis Sunyatseni*, vol. 43, no. 21, pp. 234–236, 2004.
- [9] Z. H. Zou, Y. Yun, and J. H. Sun, "Entropy method for determination of weight of evaluating indicators in fuzzy synthetic evaluation for water quality assessment," *Journal of Environmental Sciences*, vol. 18, no. 5, pp. 1020–1023, 2006.
- [10] F. Nasiri, I. Maqsood, G. Huang, and N. Fuller, "Water quality index: a fuzzy river-pollution decision support expert system," *Journal of Water Resources Planning and Management*, vol. 133, no. 2, pp. 95–105, 2007.
- [11] Z. H. Zou and X. L. Wang, "Application of BP modeling based on random samples to assessment on natural water quality," *Environmental Engineering*, vol. 25, no. 1, pp. 69–71, 2007.
- [12] L. J. Wang and Z. H. Zou, "Application of improved attributes recognition method in water quality assessment," *Chinese Journal of Environmental Engineering*, vol. 2, no. 4, pp. 553–556, 2008.
- [13] L. Liu, J. Z. Zhou, X. L. An, Y. C. Zhang, and L. Yang, "Using fuzzy theory and information entropy for water quality assessment," *Expert Systems with Applications*, vol. 37, no. 3, pp. 2517–2521, 2009.
- [14] H. Vernieuwe, O. Georgieva, B. De Baets, V. R. Pauwels, N. E. Verhoest, and F. P. De Troch, "Comparison of data-driven Takagi-Sugeno models of rainfall-discharge dynamics," *Journal of Hydrology*, vol. 302, no. 1–4, pp. 173–186, 2005.
- [15] P. C. Nayak, K. P. Sudheer, D. M. Rangan, and K. S. Ramasastri, "A neuro fuzzy computing technique for modeling hydrological time series," *Journal of Hydrology*, vol. 291, no. 1–2, pp. 52–66, 2004.
- [16] J. S. R. Jang, "ANFIS: adaptive network-based fuzzy inference system," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 23, no. 3, pp. 665–685, 1993.
- [17] H. Yan, Z. Zou, and H. Wang, "Adaptive neuro fuzzy inference system for classification of water quality status," *Journal of Environmental Sciences*, vol. 22, no. 12, pp. 1891–1896, 2010.
- [18] M. Sahu, S. S. Mahapatra, H. B. Sahu, and R. K. Patel, "Prediction of water quality index using neuro fuzzy inference system," *Water Quality, Exposure and Health*, vol. 3, no. 3–4, pp. 175–191, 2011.
- [19] A. Rahimzadeh, F. Z. Ashtiani, and A. Okhovat, "Application of adaptive neuro-fuzzy inference system as a reliable approach for prediction of oily wastewater microfiltration permeate volume," *Journal of Environmental Chemical Engineering*, vol. 4, no. 1, pp. 576–584, 2016.
- [20] M. Talebizadeh and A. Moridnejad, "Uncertainty analysis for the forecast of lake level fluctuations using ensembles of ANN

- and ANFIS models,” *Expert Systems with Applications*, vol. 38, no. 4, pp. 4126–4135, 2011.
- [21] A. M. Ahmed and S. M. Shah, “Application of adaptive neuro-fuzzy inference system (ANFIS) to estimate the biochemical oxygen demand (BOD) of Surma River,” *Journal of King Saud University–Engineering Sciences*, vol. 29, no. 3, pp. 237–243, 2017.
- [22] T. Taskaya-Temizel and M. C. Casey, “A comparative study of autoregressive neural network hybrids,” *Neural Networks*, vol. 18, no. 5-6, pp. 781–789, 2005.
- [23] APHA (American Public Health Association), *Standard Methods for the Examination of Water and Waste-Water*, American Public Health Association, Washington, DC, USA, 20th edition, 1998.
- [24] S. Sharma and A. S. Reddy, “Development of water quality indices for designated best uses of surface water bodies,” *International Journal of Engineering Science and Innovative*, vol. 2, no. 5, pp. 82–93, 2013.
- [25] J. S. R. Jang and C. T. Sun, “Neuro-fuzzy modeling and control,” *Proceedings of the IEEE*, vol. 83, no. 3, pp. 378–406, 1995.
- [26] J.-S. R. Jang, C.-T. Sun, and E. Mizutani, *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*, Prentice Hall, Upper Saddle River, NJ, USA, 1997.
- [27] T. Takagi and M. Sugeno, “Fuzzy identification of systems and its applications to modeling and control,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 15, no. 1, pp. 116–132, 1985.
- [28] A. El-Shafie, M. R. Taha, and A. Noureldin, “A neuro-fuzzy model for inflow forecasting of the Nile River at Aswan high dam,” *Water Resources Management*, vol. 21, no. 3, pp. 533–556, 2007.
- [29] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press, New York, NY, USA, 1981.
- [30] S. Chiu, “Fuzzy model identification based on cluster estimation,” *Journal of Intelligent & Fuzzy Systems*, vol. 2, no. 3, pp. 267–278, 1994.



Hindawi

Submit your manuscripts at
www.hindawi.com

