

Three Different Adaptive Neuro Fuzzy Computing Techniques for Forecasting Long-Period Daily Streamflows



Ozgur Kisi, Jalal Shiri, Sepideh Karimi and Rana Muhammad Adnan

Abstract A modeling study was presented here using three different adaptive neuro-fuzzy (ANFIS) approach algorithms comprising grid partitioning (ANFIS-GP), subtractive clustering (ANFIS-SC) and fuzzy C-Means clustering (ANFIS-FCM) for forecasting long period daily streamflow magnitudes. Long-period data (between 1936 and 2016) from two hydrometric stations in USA were used for training, evaluating and testing the approaches. Five different input combinations were applied based on the autoregressive analysis of the recorded streamflow data. A sensitivity analysis was also carried out to investigate the effect of different model architectures on the obtained outcomes. When using ANFIS-GP, the double-input model gives the best results for different model architectures, while the triple-input models produce the most accurate results using both ANFIS-SC and ANFIS-FCM models, which is due to increasing the model complexity for ANFIS-GP by using more input parameters. Comparing the all three algorithms it was observed that the ANFIS-FCM generally gave the most accurate results among others.

O. Kisi (✉)

Faculty of Natural Sciences and Engineering, Ilia State University, Tbilisi, Georgia
e-mail: ozgur.kisi@iliauni.edu.ge

J. Shiri · S. Karimi

Faculty of Agriculture, Water Engineering Department, University of Tabriz, Tabriz, Iran
e-mail: j_shiri2005@yahoo.com

S. Karimi

e-mail: karimi_sepide@yahoo.com

R. M. Adnan

Faculty of Agricultural and Biosystems Engineering and Technology,
Muhammad Nawaz Sharif University of Agriculture, Multan, Pakistan
e-mail: adnan.ikram@mnsuam.edu.pk

© Springer Nature Singapore Pte Ltd. 2018

S. S. Roy et al. (eds.), *Big Data in Engineering Applications*,
Studies in Big Data 44, https://doi.org/10.1007/978-981-10-8476-8_15

303

1 Introduction

Accurate streamflow forecast is very important in water resources system planning, design, operation and management as well as identifying hydrologic drought spells [6], controlling flood events [28], optimizing hydrologic system [17], determining environmental flow portions [33], modeling surface water-groundwater interactions [10], and modeling suspended sediment load in rivers [18]. Traditionally, conceptual simple models have been developed by numerous researchers to describe the rainfall-runoff process for computing the total amount of surface water flows. Although such models do not require more detailed information on the physical parameters, they can produce acceptable results in some cases [34]. In the contrary, physically-based models of river flow forecast are generally time consuming and complex which need lots of input variables for simulating river flow magnitudes [2]. So, using autoregressive moving average (ARMA) approaches for forecasting the streamflow magnitudes using the previously recorded flow magnitudes have been proposed as alternatives for physically-based models [23].

As a substitute, application of heuristic models e.g. adaptive neuro-fuzzy inference system (ANFIS) in streamflow forecasting has become viable. For instance, Wang et al. [35], Kagoda et al. [16] and Humphrey et al. [14] applied artificial neural networks (ANNs) models for streamflow forecasting. Shiri and Kisi [30] introduced a wavelet-neuro-fuzzy model of streamflow forecasting. Sharma et al. [29] compared neuro-fuzzy model with a physically based watershed model for streamflow forecasting and concluded that the neuro-fuzzy model was equally comparable to physical model especially when rain gauge stations were not adequate. Ballini et al. [3] applied ANFIS for seasonal river flow forecast. Nayak et al. [25] ANFIS modeling approach to model the long-term daily river flow magnitudes in India and reported that ANFIS gave promising results in this case. Vernieuwe et al. [34] compared data-driven Takagi–Sugeno models for rainfall–discharge dynamics modeling. Zounemat Kermani and Teshnelab [39] introduced ANFIS approach as a strong method of daily streamflow prediction when compared to ANN and traditional regression models. El-Shafie et al. [9] proposed ANFIS technique to forecast the inflow for the Nile River at Aswan High Dam on monthly basis. Wang et al. [36] examined different heuristic models for forecasting monthly discharge time series and introduced the ANFIS models as the most accurate technique in this field. Rath et al. [26] applied hierarchical neuro-fuzzy model for real-time flood forecasting. He et al. [11] compared ANN, ANFIS and SVM for forecasting riverflow in a semiarid mountain region and found that the performance of the applied models in river flow forecasting was satisfactory. Yazar [37] introduced a hybrid wavelet-ANFIS model for forecasting monthly streamflow time series. Yilmaz and Muttil [38] utilized different machine learning techniques including ANFIS for runoff estimation. Talei et al. [32] applied Takagi–Sugeno neuro-fuzzy model with online learning for runoff forecasting. Anusree and Varghese [1] compared ANFIS, ANN and MNL models for daily streamflow forecasting and found the ANFIS as the superior model.

The main aim of this study is to forecast long period daily streamflows using three different adaptive neuro fuzzy techniques, i.e. ANFIS with Grid Partition (ANFIS-GP), ANFIS with subtractive clustering (ANFIS-SC) and ANFIS with fuzzy C-Means clustering (ANFIS-FCM). Two different ANFIS-GP methods were considered in the present study: ANFIS-GP with constant output and ANFIS-GP with linear output. The models were also compared according to their complexity and training durations.

2 Methods

2.1 Adaptive Neuro-fuzzy Inference System (ANFIS)

ANFIS is a merger of an adaptive neural network (ANN) and a fuzzy inference system (FIS), where the parameters of FIS are identified by the ANN learning algorithms. ANFIS is able to estimate real continuous functions on a compact set of parameters with any degree of accuracy [15]. There are two approaches for FIS, namely, Mamdani and Assilian [24] and Takagi and Sugeno [31]. The differences between the two approaches corresponds to the consequent part where Mamdani’s method uses fuzzy membership functions, while linear or constant functions are utilized in Sugeno’s method.

2.1.1 ANFIS Architecture

Let’s assume a FIS having two input variables of x and y and one output variable f . The first-order Sugeno fuzzy model, a typical rule set with two fuzzy If-Then rules would read:

$$\text{Rule 1: If } x \text{ is } A_1 \text{ and } y \text{ is } B_1, \text{ then } f_1 = p_1x + q_1y + r_1 \tag{1}$$

$$\text{Rule 2: If } x \text{ is } A_2 \text{ and } y \text{ is } B_2, \text{ then } f_2 = p_2x + q_2y + r_2 \tag{2}$$

where A_1, A_2 and B_1, B_2 are the membership functions (MFs) of the inputs x and y , respectively and p_1, q_1, r_1 and p_2, q_2, r_2 are the parameters of the output function. Here, the output f is the weighted mean of the single rule outputs.

The output of the i th node in layer l is shown as $O_{l,i}$. Every node i in Layer 1 is an adaptive node with node $O_{l,i} = \varphi A_i(x)$, for $i = 1, 2$, or $O_{l,i} = \varphi B_{i-2}(y)$, for $i = 3, 4$, where x (or y) is the input to the i th node and A_i (or B_{i-2}) is a linguistic label (such as ‘low’ or ‘high’) associated with this node. The MFs for A and B are generally described by generalized bell functions as:

$$\varphi A_i(x) = \frac{1}{1 + [(x - c_i)/a_i]^{2b_i}} \quad (3)$$

where $\{a_i, b_i, c_i\}$ is the parameter set. Parameters in this layer are called as premise parameters. The outputs of this layer are the membership values of the premise part. Layer 2 includes the nodes labeled Π which multiply incoming signals and sending the product out. For instance,

$$O_{2,i} = w_i = \varphi A_i(x)\varphi B_i(y), \quad i = 1, 2. \quad (4)$$

Each node output shows the firing strength of a rule. The nodes labeled N computes the ratio of the i th rule's firing strength to the sum of all rules' firing strengths in Layer 3,

$$O_{3,i} = \bar{w}_i = \frac{w_i}{w_1 + w_2}, \quad i = 1, 2. \quad (5)$$

The outputs of this layer are referred to as normalized firing strengths. The nodes of the Layer 4 are adaptive with node functions

$$O_{4,i} = \bar{w}_i f_i = \bar{w}_i(p_i x + q_i y + r_i) \quad (6)$$

where \bar{w}_i is the output of Layer 3, and $\{p_i, q_i, r_i\}$ are the parameter set. Parameters of this layer are called as consequent parameters. The single fixed node of the Layer 5 labeled Σ computes the final output as the summation of all incoming signals

$$O_{5,i} = \sum_{i=1} \bar{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \quad (7)$$

So, an adaptive network which is functionally equivalent to a Sugeno first-order fuzzy inference system is built.

2.1.2 Grid Partition Method

Grid partition (GP) is one of the commonly used methods for producing initial FIS rules for ANFIS building, where the space including input- output parameters is divided into certain partitions called as grids. Each grid expresses a fuzzy surface, and interference areas between grids make a continuous output surface [13, 22]. There is no fixed rule for defining the number of MFs for each variable, so they are identified through a trial and error process. The learning process is begun from zero output and during the learning process, functions and fuzzy rules are trained gradually [7]. Although they are different membership function types which can be applied in modeling various procedures, the literature review shows that triangular MFs are commonly used and the most optimal MF types for practical applications

[27]. Nevertheless, other studies (e.g. [34]) have confirmed that the type of MFs cannot affect the results of simulation majorly, though other studies have demonstrated the major effect of MFs in modeling accuracy (e.g. [19]).

2.1.3 Subtractive Clustering Method

The subtractive clustering method assumes that each data point is a potential cluster center and calculates a measure of the likelihood that each data point would define the cluster center on the basis of the density of surrounding data points. Considering a set of n data points $\{x_1, x_2, \dots, x_i\}$ in m -dimensional space, it is assumed that all data points within a cubic space have been normalized. In subtractive clustering, each of the data points is considered as a potential cluster center. As a result, the density index D_i corresponding to the data x_i can be expressed as follows:

$$D_i = \sum_{j=1}^n \exp\left(-\frac{\|x_i - x_j\|}{(r_a/2)^2}\right) \tag{8}$$

Here, r_a is a positive quantity called cluster radius. If many data points are adjacent to a data point, hence, that data point has the maximum density. After measuring the density of each data point, data point with the highest density is selected as the first data center clustering [12]. If the effect limited area of the center of the first cluster center is removed, following formula is used to measure the other points density.

$$D_i = D_i - D_{c1} \sum_{j=1}^n \exp\left(-\frac{\|x_i - x_{c1}\|^2}{(r_b/2)^2}\right) \tag{9}$$

Here, x_{c1} and D_{c1} are the selected points and density potential, respectively. r_b is a positive constant. To avoid approaching the cluster centers, the r_b constant value is normally larger than r_a (r_b is considered $1.5r_a$). After measuring the density for each data point, the next cluster center x_{c2} is selected and all the measured density for data points will be recalculated. This process continues until a sufficient number of cluster center produce [2, 20].

2.1.4 Fuzzy C-Means Method

In fuzzy clustering, each pattern might belong to several clusters or segment. One of the most functional clustering algorithms is K-mean algorithm. This unsupervised algorithm in large datasets, exposures with some limitations in the process, may not work properly. To deal with the disadvantages, different clustering algorithms have been proposed. Among them, fuzzy c-means as a proper alternative method is used

[21]. Fuzzy C-means (FCM) was developed by Dunn [8] and Bezdek [4] improved it.

The FCM method blocks a set of N vector x_i , $i = 1, \dots, n$, into c fuzzy clusters, where each pattern is corresponded to a cluster with a degree specified by a membership grade u_{ij} between 0 and 1. The final object by the FCM algorithm is to find c cluster centers so that the cost function of the dissimilarity measure can be minimized. The aim is minimizing the objective function that is defined as below:

$$\text{Min} J_{FCM} = \sum_{c=1}^C \sum_{i=1}^N w_{ic}^p \|w_i - v_c\|^2 \quad \text{s.t.} \quad \sum_{c=1}^C w_{ic} = 1, i = 1, 2, \dots, N \quad (10)$$

which p ($1 < p$) is known as fuzzifier portion and N , is the number of data points; C , the number of clusters; w_{ic} , the number of belongings of the i th data point to the c th cluster; v_c is the clusters center and x is the number of the input. for calculating the amount of w_{ic} the following formula is used [5]:

$$w_{ic} = \frac{1}{\sum_{L=1}^C \left(d_{ic}^2 / d_{ij}^2 \right)^{1/(p-1)}} \quad \text{for } i = 1, 2, \dots, N \text{ and } c = 1, 2, \dots, C \quad (11)$$

For beginning of the center vectors, centers are calculated by:

$$v_c = \frac{\sum_{j=1}^N w_{jc}^p x_j}{\sum_{j=1}^N w_{jc}^p} \quad (12)$$

FCM procession continues until a convergence condition is obtained.

3 Case Study

Daily streamflow data from two stations, Murder Creek near Evergreen (Hydrologic Unit Code 03140304, Latitude $31^{\circ}25'06''$, Longitude $86^{\circ}59'12''$, Drainage area 176.00 square miles, Gage datum 178.29 feet a.s.l.) and Choctawhatchee River Near Newton (Hydrologic Unit Code 03140201, Latitude $31^{\circ}20'34''$, Longitude $85^{\circ}36'38''$, Drainage area 686.00 square miles, Gage datum 138.56 feet a.s.l.), Alabama, USA were used in this research. The location of the study area and stations are illustrated in Fig. 1. The reason of selection of these stations is due to having long data period. Data covering the period of 1936–2016 for the both stations were divided into three parts, training (01/01/1938–12/31/1976, 14245 values), validation (01/01/1977–12/31/1996, 7305 values) and testing (01/01/1997–12/31/2016, 7305).

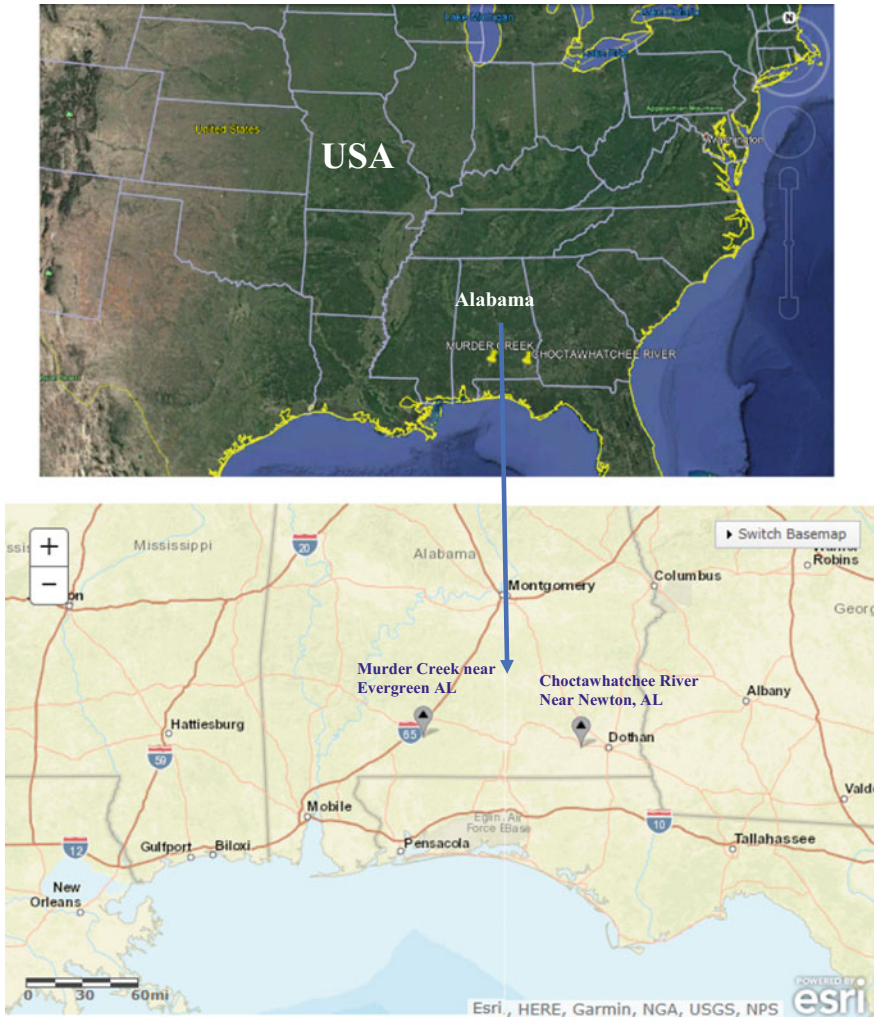


Fig. 1 The location of the studied area

The summary of statistical properties is reported in Table 1 for the used streamflow data. From the table, it is clear that data have highly skewed distributions, skewness coefficient ranging between 8.19 and 20.7. Data of the Choctawhatchee River has more autocorrelations than those of the Murder Creek. This may be due to the high discharge volume of the Choctawhatchee.

Table 1 The statistical properties of the streamflow data sets

Station	Data set	Min	Max	Mean	Sd	Csx	r ₁	r ₂	r ₃
Choctawhatchee River	Training	1.73	584	27.9	34.4	4.52	0.887	0.702	0.564
	Validation	1.23	205	25.8	47.9	20.7	0.789	0.505	0.339
	Testing	0.99	975	22.7	36.7	8.19	0.867	0.650	0.498
Murder Creek	Training	1.11	346	7.93	10.1	10.1	0.668	0.427	0.341
	Validation	1.62	264	8.49	9.68	8.94	0.719	0.399	0.287
	Testing	0.94	241	6.89	9.19	8.45	0.743	0.484	0.357

Min, Max, Mean, Sd, Csx, r₁, r₂ and r₃ show the mean, minimum, maximum, standard deviation, skewness coefficients, lag-1, lag-2 and lag-3 autocorrelations, respectively

4 Application and Results

In this study, the ability of cluster based neuro fuzzy methods, ANFIS-SC and ANFIS-FCM, was investigated in forecasting daily streamflows which have long data period (1936–2016). The results of these methods were compared with the ANFIS-GP which uses all possible rule combinations and generally has higher complexity and computational time when compared to cluster based neuro fuzzy methods. For the ANFIS-GP, two different outputs, constant and linear, were applied to detect the difference with each other. The models were compared according to the four different statistics, root mean square error (RMSE), mean absolute relative error (MARE), determination coefficient (R²) and Nash-Sutcliffe efficiency (NE) which can be expressed as

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Q_{m,i} - Q_{o,i})^2} \tag{13}$$

$$MARE = \frac{1}{n} \sum_{i=1}^n \frac{|Q_{m,i} - Q_{o,i}|}{Q_{o,i}} 100 \tag{14}$$

$$NSE = 1 - \frac{\sum_{i=1}^n (Q_{m,i} - Q_{o,i})}{\sum_{i=1}^n (Q_{o,i} - \bar{Q}_o)} \tag{15}$$

where $Q_{o,i}$ and $Q_{m,i}$ are the observed and estimated streamflows, N is the number of time steps, \bar{Q}_o is the mean of the observed streamflows. First, auto and partial auto correlation analysis were employed and they suggested three previous lags. In the applications, however, five input scenarios comprising five previous lags were used from input(i) to input(iv) comprising Qt-1 to Qt-1, Qt-2, Qt-3, Qt-4 and Qt-5.

The results of the ANFIS-GP models with constant and linear outputs are presented for the Choctawhatchee River in Tables 2 and 3. For the ANFIS-GP models, different number of triangular membership functions were tried and the best one that gave the minimum RMSE in the validation period was selected. It is clear from

Table 2 Results of ANFIS-GP models with constant output—Choctawhatchee River

Data set	Statistics	Input (i)	Input (ii)	Input (iii)	Input (iv)	Input (v)	Mean
Training	RMSE (m ³ /s)	51.78	46.67	16.99	10.99	11.52	25.79
	MARE (%)	48.52	16.36	7.700	6.180	5.950	15.13
	R ²	49.36	15.99	9.200	6.950	7.200	16.11
	NE	45.06	9.160	7.180	4.930	5.610	12.93
	Duration (s)	48.68	22.05	10.27	7.260	7.570	17.49
Validation	RMSE (m ³ /s)	31.37	29.44	29.04	29.21	30.18	29.85
	MARE (%)	29.66	16.54	16.96	17.29	18.60	19.81
	R ²	0.594	0.657	0.664	0.655	0.618	0.638
	NE	0.570	0.622	0.632	0.628	0.602	0.611
Testing	RMSE (m ³ /s)	18.08	16.15	15.70	15.96	16.33	16.44
	MARE (%)	42.28	21.08	21.91	22.48	24.38	26.42
	R ²	0.759	0.811	0.821	0.814	0.804	0.802
	NE	0.758	0.807	0.817	0.811	0.802	0.799

Table 3 Results of ANFIS-GP models with linear output—Choctawhatchee River

Data set	Statistics	Input (i)	Input (ii)	Input (iii)	Input (iv)	Input (v)	Mean
Training	RMSE (m ³ /s)	15.81	13.53	13.31	13.14	12.88	13.73
	MARE (%)	24.60	15.46	16.44	15.60	15.54	17.53
	R ²	0.789	0.845	0.850	0.854	0.860	0.839
	NE	0.789	0.845	0.850	0.854	0.860	0.839
	Duration (s)	3.680	8.710	26.42	116.6	619.0	154.9
Validation	RMSE (m ³ /s)	31.37	10,687	22,450	1711	1648	7306
	MARE (%)	29.66	76.32	233.5	23.33	20.14	76.59
	R ²	0.594	0.016	0.054	0.241	0.257	0.233
	NE	0.570	-49,861	-220,032	-1278	-1185	-54,471
Testing	RMSE (m ³ /s)	18.08	6744	13,108	2042	1930	4768
	MARE (%)	42.28	64.13	203.4	28.39	27.48	73.14
	R ²	0.759	0.010	0.000	0.041	0.034	0.169
	NE	0.758	-33,719	-127,390	-3093	-2762	-33,392

the tables that the ANFIS-GP with constant output has the best accuracy in test period for the 3rd input combination while the 1st input provides the best results for the ANFIS-GP with linear output. First model comprising constant output seems to be superior to the second model. This can also be seen from the mean of the all input combinations. From the mean statistics, it is clear that the training accuracy of ANFIS-GP models with linear output is better compared to the other one. The 2nd model with linear output can approximate better than the 1st model with constant output because it has higher number of parameters and more flexible than the latter one. Assume that we used 2 Gaussian membership functions (each has 2 parameters) and 5 inputs for each model. In this case, the premise parameters of the both

models will be $2 * 2^5 = 64$. The 1st model will have 32 rules, each has 1 constant output parameter and totally it will have 32 output parameters while the 2nd model will have 32 rules, each has 6 output parameters and totally $32 * 6 = 192$ output parameters. For this reason, the training duration of the ANFIS-GP models with linear output is also much higher than those of the ANFIS-GP models with constant output especially when the number of inputs is higher than 2.

Tables 4 and 5 report the results of ANFIS-FCM and ANFIS-SC models with respect to RMSE, MARE, R^2 and NE for the Choctawhatchee River. For the ANFIS-FCM models, different number of cluster numbers (vary between 1 and 8)

Table 4 Results of ANFIS-FCM models—Choctawhatchee River

Data set	Statistics	Input (i)	Input (ii)	Input (iii)	Input (iv)	Input (v)	Mean
Training	RMSE (m ³ /s)	15.58	13.77	13.63	13.70	13.71	14.08
	MARE (%)	16.91	15.17	15.66	15.63	15.84	15.84
	R ²	0.795	0.839	0.843	0.841	0.841	0.832
	NE	0.795	0.839	0.843	0.841	0.841	0.832
	Duration (s)	13.20	20.80	28.23	20.98	25.43	21.73
Validation	RMSE (m ³ /s)	28.92	24.32	24.75	24.60	24.75	25.47
	MARE (%)	17.05	15.09	15.79	15.83	16.18	15.99
	R ²	0.635	0.759	0.756	0.756	0.750	0.731
	NE	0.635	0.742	0.733	0.736	0.733	0.716
Testing	RMSE (m ³ /s)	18.00	15.99	15.69	15.70	15.71	16.22
	MARE (%)	19.39	16.89	18.54	17.99	18.51	18.26
	R ²	0.760	0.811	0.817	0.817	0.817	0.804
	NE	0.760	0.811	0.817	0.817	0.817	0.804

Table 5 Results of ANFIS-SC models—Choctawhatchee River

Data set	Statistics	Input (i)	Input (ii)	Input (iii)	Input (iv)	Input (v)	Mean
Training	RMSE (m ³ /s)	15.64	14.04	13.76	13.77	13.78	14.20
	MARE (%)	19.55	17.74	16.83	16.90	17.04	17.61
	R ²	0.793	0.833	0.840	0.839	0.839	0.829
	NE	0.793	0.833	0.840	0.839	0.839	0.829
	Duration (s)	12.89	16.68	20.64	30.70	36.03	23.39
Validation	RMSE (m ³ /s)	34.17	33.35	33.44	33.66	33.94	33.71
	MARE (%)	22.71	20.74	18.68	18.54	18.74	19.88
	R ²	0.503	0.534	0.528	0.518	0.506	0.518
	NE	0.490	0.515	0.512	0.505	0.497	0.504
Testing	RMSE (m ³ /s)	18.71	17.17	16.40	16.62	16.92	17.16
	MARE (%)	29.84	27.41	23.58	23.08	23.44	25.47
	R ²	0.741	0.783	0.802	0.796	0.788	0.782
	NE	0.740	0.781	0.801	0.795	0.788	0.781

which decides the number of rules were tried and the best one that gave the minimum RMSE in the validation period was selected. For the ANFIS-SC models, different number of radii values (vary between 0.1 and 1) which decides the number of membership functions and rules were tried. It is apparent from the tables, the 3rd input combination has the best accuracy for the both methods and after 3 lags input, the accuracy of the models does not considerably increase. Comparison with the ANFIS-GP models indicates that the ANFIS-FCM model slightly performs superior to the ANFIS-GP with constant output and ANFIS-SC models. The training duration of the ANFIS-GP with constant output is less than those of the cluster

Table 6 Results of ANFIS-GP models with constant output—Murder Creek

Data set	Statistics	Input (i)	Input (ii)	Input (iii)	Input (iv)	Input (v)	Mean
Training	RMSE (m ³ /s)	7.520	7.370	7.350	15.63	13.60	10.30
	MARE (%)	35.33	28.05	25.91	20.75	16.11	25.23
	R ²	0.450	0.471	0.474	0.793	0.843	0.606
	NE	0.450	0.471	0.474	0.793	0.843	0.606
	Duration (s)	3.720	6.950	12.71	26.61	64.08	22.81
Validation	RMSE (m ³ /s)	6.730	6.540	6.450	32.71	30.42	16.57
	MARE (%)	29.03	25.24	24.51	24.98	18.37	24.43
	R ²	0.519	0.548	0.561	0.546	0.623	0.559
	NE	0.516	0.543	0.556	0.533	0.596	0.549
Testing	RMSE (m ³ /s)	6.180	5.990	5.980	18.11	18.17	10.89
	MARE (%)	48.31	35.83	31.66	28.80	20.91	33.10
	R ²	0.554	0.583	0.582	0.757	0.757	0.647
	NE	0.548	0.576	0.577	0.757	0.755	0.642

Table 7 Results of ANFIS-GP models with linear output—Murder Creek

Data set	Statistics	Input (i)	Input (ii)	Input (iii)	Input (iv)	Input (v)	Mean
Training	RMSE (m ³ /s)	7.140	6.880	6.720	15.31	13.15	9.840
	MARE (%)	19.15	18.55	19.19	16.97	16.09	17.99
	R ²	0.503	0.539	0.560	0.802	0.854	0.651
	NE	0.503	0.539	0.560	0.802	0.854	0.651
	Duration (s)	4.100	8.760	26.62	116.2	941.1	219.4
Validation	RMSE (m ³ /s)	6.640	6.280	13.93	7442	7013	2896
	MARE (%)	20.26	18.89	18.82	29.37	48.98	27.26
	R ²	0.529	0.578	0.133	0.214	0.018	0.295
	NE	0.529	0.578	-1.075	-24,179	-21,472	-9130
Testing	RMSE (m ³ /s)	5.830	5.480	7.320	4193.5	3420	4768
	MARE (%)	19.83	19.83	22.05	35.92	38.87	73.14
	R ²	0.598	0.646	0.471	0.044	0.027	0.169
	NE	0.598	0.644	0.365	-13,036	-8674	-33,392

based ANFIS-FCM and ANFIS-SC models. The reason of this might be fact that the ANFIS-FCM and ANFIS-SC have linear output comprising more consequent parameters. However, the main advantage of the cluster based neuro fuzzy methods is that their rules are automatically determined based on the selected cluster number or radii value. For example, in case of 8 clusters, we will have only 8 rules for the whole fuzzy model while the ANFIS-GP has 32 rules when the input number is 5.

The statistics of the ANFIS-GP models with constant and linear outputs are compared in Tables 6 and 7 for the Murder Creek. As seen from the tables, the both methods have the best accuracy in the test period for the 2nd input combination. After 2nd input, the accuracy of ANFI-GP model comprising linear output is

Table 8 Results of ANFIS-FCM models—Murder Creek

Data set	Statistics	Input (i)	Input (ii)	Input (iii)	Input (iv)	Input (v)	Mean
Training	RMSE (m ³ /s)	7.100	6.870	7.000	15.58	14.10	10.13
	MARE (%)	18.73	16.94	18.15	22.58	25.42	20.36
	R ²	0.509	0.540	0.523	0.795	0.832	0.640
	NE	0.509	0.540	0.523	0.795	0.832	0.640
	Duration (s)	13.31	19.90	15.26	13.90	16.03	15.68
Validation	RMSE (m ³ /s)	6.670	6.290	6.250	27.21	25.41	14.37
	MARE (%)	19.78	17.71	18.84	27.60	32.46	23.28
	R ²	0.525	0.577	0.583	0.683	0.749	0.623
	NE	0.525	0.577	0.582	0.677	0.718	0.616
Testing	RMSE (m ³ /s)	5.860	5.570	5.620	18.09	16.18	10.26
	MARE (%)	19.63	18.12	19.59	35.66	43.00	27.20
	R ²	0.594	0.634	0.628	0.757	0.806	0.684
	NE	0.594	0.634	0.627	0.757	0.806	0.683

Table 9 Results of ANFIS-SC models—Murder Creek

Data set	Statistics	Input (i)	Input (ii)	Input (iii)	Input (iv)	Input (v)	Mean
Training	RMSE (m ³ /s)	92.24	126.8	8.690	15.59	14.01	51.46
	MARE (%)	33.02	40.23	8.460	18.70	17.41	23.56
	R ²	0.030	0.453	0.327	0.794	0.834	0.488
	NE	0.030	0.453	0.327	0.794	0.834	0.488
	Duration (s)	0.266	0.543	0.790	30.71	36.03	13.67
Validation	RMSE (m ³ /s)	122.5	272.2	10.96	33.98	33.37	94.60
	MARE (%)	40.22	73.53	11.81	21.83	20.41	33.56
	R ²	0.000	0.034	0.157	0.510	0.531	0.247
	NE	-0.047	-1.347	0.093	0.496	0.514	-0.058
Testing	RMSE (m ³ /s)	114.2	165.1	9.620	18.36	16.67	64.80
	MARE (%)	41.93	57.57	9.910	26.28	25.62	32.26
	R ²	0.002	0.178	0.244	0.751	0.796	0.394
	NE	-0.067	0.061	0.215	0.750	0.794	0.351

worsening. It can be said that increasing input number increases the complexity of the model and this results in worse streamflow forecasts. Similar to the Choctawhatchee River, the training durations of the ANFIS-GP with linear output is higher

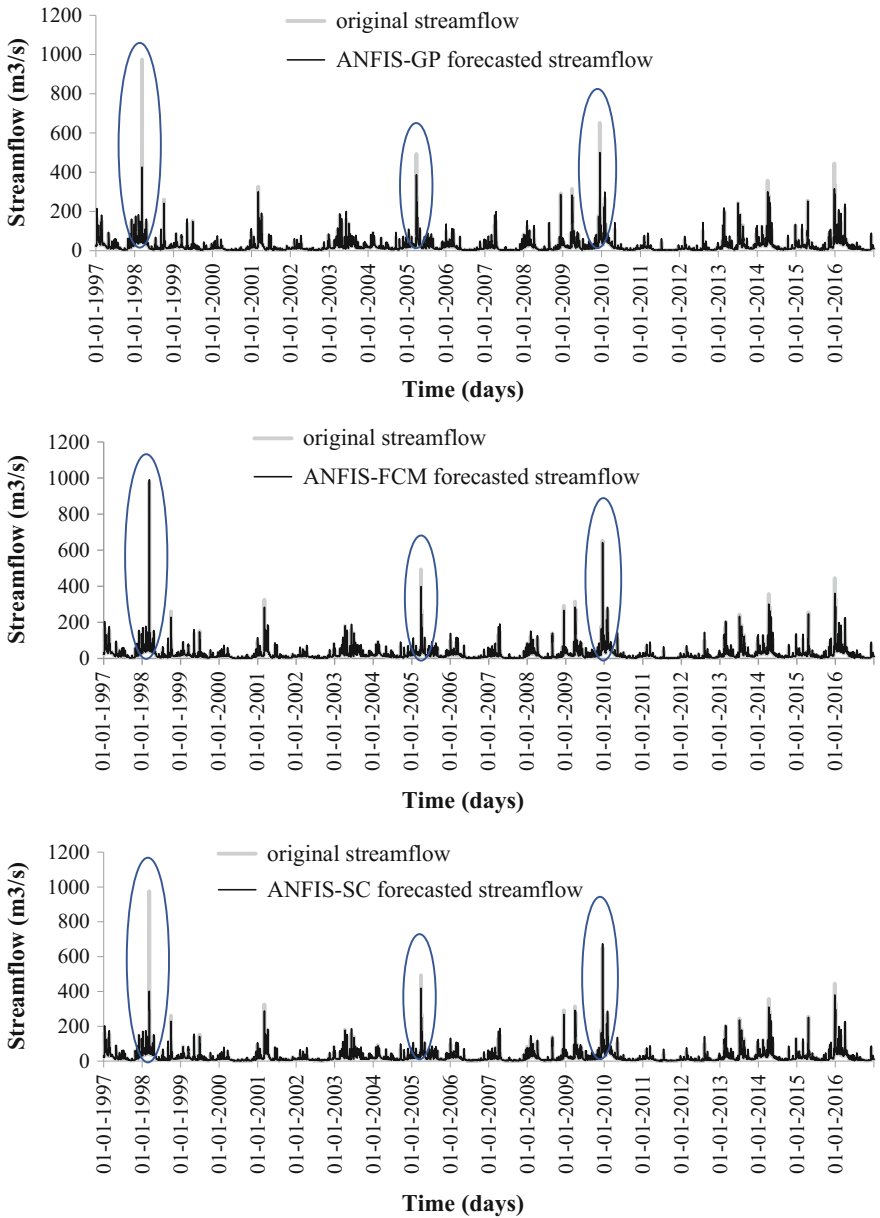


Fig. 2 The time variation of the observed and forecasted streamflows by using the optimal ANFIS-GP-constant, ANFIS-FCM and ANFIS-SC models—Choctawhatchee River

than those of the models with constant output especially for the inputs higher than 2. Tables 8 and 9 present the training, validation and testing results of ANFIS-FCM and ANFIS-SC models for the Murder Creek. ANFIS-FCM model has the best accuracy in 2nd input combination while the 3rd input combination provides the best accuracy for ANFIS-SC model. Comparison of the Tables 6, 7 and 8 clearly shows that the ANFIS-GP model with linear output slightly performs superior to the ANFIS-GP with constant output and ANFIS-FCM models. ANFIS-SC model has the worst accuracy even though it has the least training duration. Comparison of two stations obviously indicates that the accuracy of the applied models is better for the Choctawhatchee River compared to Murder Creek. The main reason of this might be the fact that the data of first station has lower autocorrelations than the latter one.

The time variation of observed and forecasted streamflows by using the optimal ANFIS-GP-constant, ANFIS-FCM and ANFIS-SC models can be seen from Fig. 2 for the Choctawhatchee River. From the figures, it is clear that the ANFIS-FCM model catches the high streamflow values better than the other models. The

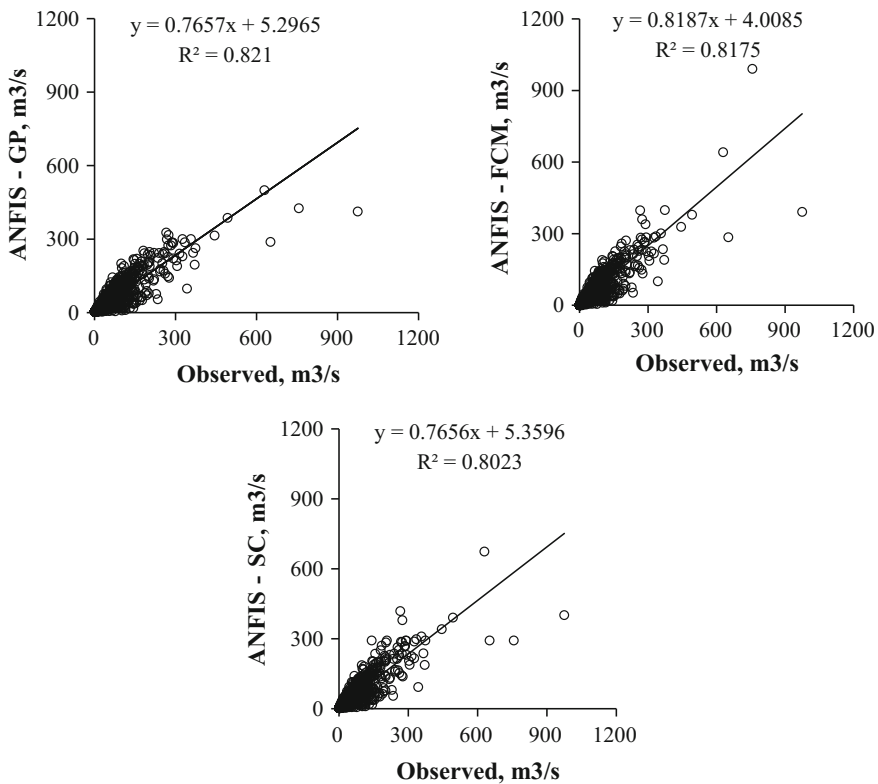


Fig. 3 The scatterplots of the observed and forecasted streamflows by using the optimal ANFIS-GP-constant, ANFIS-FCM and ANFIS-SC models—Choctawhatchee River

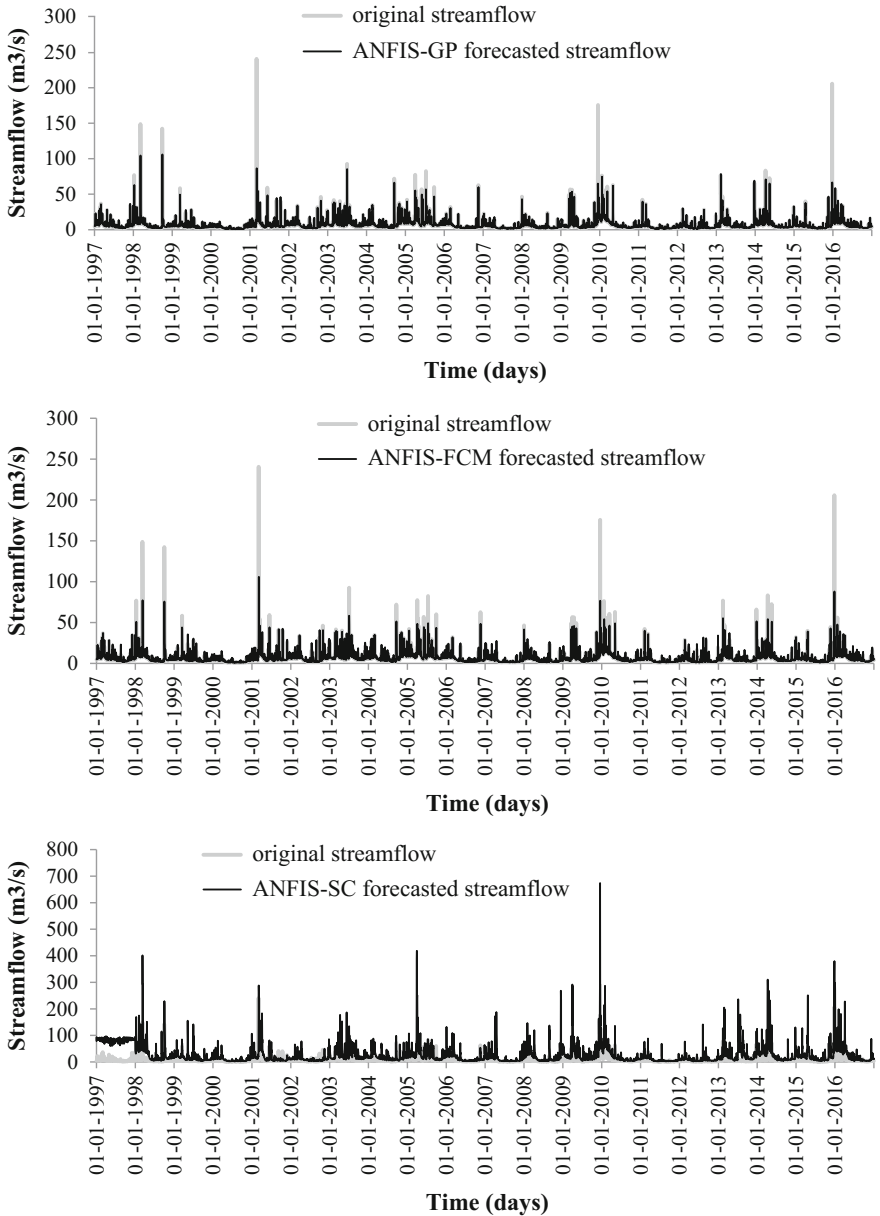


Fig. 4 The time variation of the observed and forecasted streamflows by using the optimal ANFIS-GP-linear, ANFIS-FCM and ANFIS-SC models—Murder Creek

ANFIS-SC also seems to be better than ANFIS-GP model. Figure 3 makes the scatterplot comparison of the applied models. ANFIS-GP model has slightly higher R^2 than the ANFIS-FCM. However, the a and b coefficients of the fit line equation (assume that the fit line is $y = ax + b$) are respectively closer to the 1 and 0 (exact line is $y = x$) for the ANFIS-FCM compared to the ANFIS-GP model. Figure 4 illustrates the time variation graphs of the observed and forecasted streamflows by using the optimal ANFIS-GP-linear, ANFIS-FCM and ANFIS-SC models for the Murder Creek. The ANFIS-GP-linear and ANFIS-FCM models considerably underestimate peak discharges while the ANFIS-SC overestimates. The scatter diagrams of the three methods are given in Fig. 5. As seen from the figure, the ANFIS-GP model has slightly higher R^2 than the ANFIS-FCM while the slope coefficient of the latter model closer to the 1 compared to the first model. The ANFIS-SC seems to be insufficient in forecasting daily streamflows of Murder Creek.

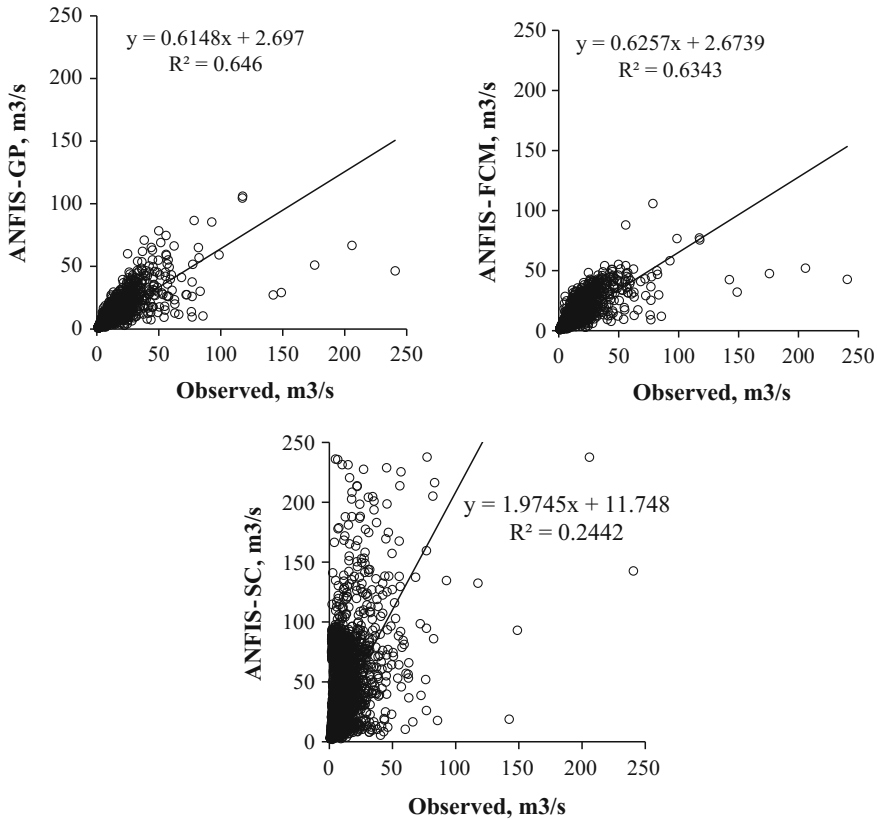


Fig. 5 The scatterplots of the observed and forecasted streamflows by using the optimal ANFIS-GP-linear, ANFIS-FCM and ANFIS-SC models—Murder Creek

5 Conclusion

Long period streamflow data from two hydrometric stations in USA were used in the present research to forecast streamflow magnitudes in daily forecast horizon. Adaptive neuro-fuzzy inference system (ANFIS) with three different running algorithms, namely, ANFIS grid partitioning (ANFIS-GP), ANFIS sub clustering (ANFIS-SC) and ANFIS fuzzy C means (ANFIS-FCM) were then trained, validated and tested using these data. Five input combinations were tried by also considering the auto- and partial-auto-correlation functions of the streamflow records during the study period to see the effect of 5 time lags on the predictions. Using different models and input combinations it was revealed that the best input combination (which can be used to feed the predictive models) is somewhat model-specific, where introducing more input parameters (beyond the double-input combination) has deteriorated the ANFIS-GP accuracy. This might be linked to the model complexity by using more inputs and might dictate a risk of redundancy when using inputs roughly based on linear measures (e.g. auto correlation). Nevertheless, the models architectures had monotonous influence on the predictive models performance that showed the necessity of performing sensitivity analysis on the models architectures. This might be crucially important when using short period data, where the time domain is limited and general trend of data which can affect the predictions are not involved in model training. It was seen that the ANFIS-GP model with linear output produce complex model structure especially in case of high number of inputs compared to ANFIS-GP with constant output.

In this study, high number of membership functions were not tried for that ANFIS-GP model because its parameters exponentially increase when the number of MFs was increased. In future studies, the effect of MFs numbers may be investigated by using high speed computers.

References

1. Anusree, K., & Varghese, K. O. (2016). Streamflow prediction of Karuvannur River Basin using ANFIS, ANN and MNLN models. *Procedia Technology*, 24, 101–108.
2. Aqil, M., Kita, I., Yano, A., et al. (2007). A comparative study of artificial neural networks and neuro-fuzzy in continuous modeling of the daily and hourly behavior of runoff. *Journal of Hydrology*, 337, 22–34.
3. Ballini, R., Soares, S., & Andrade, M. G. (1999). Seasonal streamflow forecasting via a neural fuzzy system. In: 14th Triennial World Congress, Beijing, P.R. China (pp. 5249–5254).
4. Bezdek, J. C. (1981). *Pattern recognition with fuzzy objective function algorithms*. New York: Plenum.
5. Bezdek, J. C., Ehrlich, R., & Full, W. (1984). FCM: The fuzzy C-means clustering algorithm. *Computers & Geosciences*, 10(2–3), 191–203.
6. Chemedo Edossa, D., & Singh Babel, M. (2011). Application of ANN-based streamflow forecasting model for agricultural water management in the Awash River Basin, Ethiopia. *Water Resources Management*, 25, 1759–1773.

7. Cobaner, M. (2011). Evapotranspiration estimation by two different neuro-fuzzy inference systems. *Journal of Hydrology*, 398(3–4), 299–302.
8. Dunn, J. C. (1973). A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters. *Journal of Cybernetics*, 3(3), 32–57.
9. El-Shafie, A., Taha, M. R., & Noureldin, A. (2007). A neuro-fuzzy model for inflow forecasting of the Nile River at Aswan high dam. *Water Resources Management*, 21, 533–556.
10. Gunduz, O., & Aral, M. M. (2005). River networks and groundwater flow: A simultaneous solution of a coupled system. *Journal of Hydrology*, 301(1–4), 216–234.
11. He, Z., wen, X., Liu, H., & Du, J. (2013). A comparative study of artificial neural network, adaptive neuro fuzzy inference system and support vector machine for forecasting river flow in the semiarid mountain region. *Journal of Hydrology*, 509, 379–386.
12. Hiremath, S. M., Patra, S. K., & Mishra, A. K. (2012). ANFIS with subtractive clustering-based extended data rate prediction for cognitive radio. In *Proceeding of the 5th International Conference on Computers and Devices for Communication (CODEC)*. <https://doi.org/10.1109/codec.2012.6509239>.
13. Hu, Y. C. (2007). Sugeno fuzzy integral for finding fuzzy if–Then classification rules. *Applied Mathematics and Computation*, 185, 72–83.
14. Humphrey, G. B., Gibbs, M. S., Dandy, G. C., & Maier, H. R. (2016). A hybrid approach to monthly streamflow forecasting: Integrating hydrological model outputs into a Bayesian artificial neural network. *Journal of Hydrology*, 540, 623–640.
15. Jang, J. S. R., Sun, C. T., & Mizutani, E. (1997). *Neurofuzzy and soft computing: A computational approach to learning and machine intelligence*. New Jersey: Prentice-Hall.
16. Kagoda, P. A., Ndiritu, J., Ntuli, C., & Mwaka, B. (2010). Application of radial basis function neural networks to short-term streamflow forecasting. *Physics and Chemistry of the Earth*, 35 (13–14), 571–581.
17. Kisi, O. (2008). River flow forecasting and estimation using different artificial neural network techniques. *Hydrology Research*, 39(1), 27–40.
18. Kisi, O., Hossein zadeh Dalir, A., Cimen, M., & Shiri, J. (2012). Suspended sediment modeling using genetic programming and soft computing techniques. *Journal of Hydrology*, 450–451, 48–58.
19. Kisi, O., Shiri, J., & Tombul, M. (2013). Modeling rainfall-runoff process using soft computing techniques. *Computers & Geosciences*, 51, 108–117.
20. Kisi, O., Karimi, S., Shiri, J., Makarynsky, O., & Yoon, H. (2014). Forecasting sea water levels at Mukho Station, South Korea using soft computing techniques. *The International Journal of Ocean and Climate Systems*, 5(4), 175–188.
21. Kisi, O., & Zounemat-Kermani, M. (2016). Suspended sediment modeling using neuro-fuzzy embedded fuzzy c-means clustering technique. *Water Resources Management*, 30(11), 3979–3994.
22. Lin, C. T., Lin, C. J., & Lee, C. S. G. (1995). Fuzzy adaptive learning control network with on-line neural learning. *Fuzzy Sets Systems*, 71, 25–45.
23. Maier, H. R., & Dandy, G. C. (1996). Use of artificial neural networks for prediction of water quality parameters. *Water Resources Research*, 32(4), 1013–1022.
24. Mamdani, E. H., & Assilian, S. (1975). An experiment in linguistic synthesis with a fuzzy logic controller. *International Journal of Man-Machine Studies*, 7(1), 1–13.
25. Nayak, P. C., Sudheer, K. P., Rangan, D. M., & Ramasastri, K. S. (2004). A neuro-fuzzy computing technique for modeling hydrological time series. *Journal of Hydrology*, 291, 52–66.
26. Rath, S., Nayak, P. C., & Chatterjee, C. (2013). Hierarchical neuro-fuzzy model for real-time flood forecasting. *International Journal of River Basin Management*, 11(3), 253–268.
27. Russel, S. O., & Campbell, P. F. (1996). Reservoir operating rules with fuzzy programming. *Journal of Water Resources Planning and Management*, 122(3), 165–170.
28. Sarlak, N. (2008). Annual streamflow modelling with asymmetric distribution function. *Hydrological Processes*, 22, 3403–3409.

29. Sharma, S., Srivastava, P., Fang, X., & Kalin, L. (2015). Performance comparison of Adoptive Neuro Fuzzy Inference System (ANFIS) with Loading Simulation Program C++ (LSPC) model for streamflow simulation in El Niño Southern Oscillation (ENSO)-affected watershed. *Expert Systems with Applications*, 42(4), 2213–2223.
30. Shiri, J., & Kisi, O. (2010). Short-term and long-term streamflow forecasting using a wavelet and neuro-fuzzy conjunction model. *Journal of Hydrology*, 394, 486–493.
31. Takagi, T., & Sugeno, M. (1985). Fuzzy identification of systems and its application to modeling and control. *IEEE Transactions on Systems, Man, and Cybernetics*, 15(1), 116–132.
32. Talei, A., Chua, L. H., Queck, C., & Jansson, P. E. (2013). Runoff forecasting using a Takagi-Sugeno neuro-fuzzy model with online learning. *Journal of Hydrology*, 488, 17–32.
33. Tennant, D. L. (1976). Instream flow regimes for fish, wildlife, recreation and related environmental resources. *Fisheries*, 1, 6–10.
34. Vernieuwe, H., Georgieva, O., De Baets, B., Pauwels, V. R. N., Verhoest, N. E. C., & De Troch, F. P. (2005). Comparison of data-driven Takagi-Sugeno models of rainfall-discharge dynamics. *Journal of Hydrology*, 302(1–4), 173–186.
35. Wang, W., Van Gelder, P., Vrijling, J. K., & Ma, J. (2006). Forecasting daily streamflow using hybrid ANN models. *Journal of Hydrology*, 324, 383–399.
36. Wang, W., Chau, K. W., Cheng, C. T., & Qiu, L. (2009). A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series. *Journal of Hydrology*, 374, 294–306.
37. Yarar, A. (2014). A hybrid wavelet and neuro-fuzzy model for forecasting the monthly streamflow data. *Water Resources Management*, 28, 553–565.
38. Yilmaz, A. G., & Muttil, N. (2014). Runoff estimation by machine learning methods and application to the Euphrates Basin in Turkey. *Journal of Hydrologic Engineering*, 19(5), 1015–1025.
39. Zounemat Kermani, M., & Teshnelab, M. (2008). Using adaptive neuro-fuzzy inference system for hydrological time series prediction. *Applied Soft Computing*, 8, 928–936.