

A neuro-fuzzy modeling for the hydrological time series of floods of river indus of Pakistan

Abstract

Many forecasting applications of science and technology are used to predict future values. Today, the flood peak flow discharges are extensively used on annual basis in risk assessment by collecting quantitative data from several sources. The famous rivers of Pakistan i.e. River Jhelum, River Kabul, River Chenab, and the very famous, considering upper area parts and the lower area parts of the river, the Indus River are the prime sources of flooding. These aforesaid rivers are the prime tributaries of the Indus River System which is the one, from all, most notable rivers of the world and for Pakistan, it is a supreme river. River Indus is Pakistan's longest river with seven (7) different gauge stations, Dams and various barrages, and plays a significant role in irrigation and maximum generation of power in Pakistan. In the present research the flood risk in the Indus River has been calculated by utilizing the historical peak streamflow discharges recorded data on the daily basis. Nowadays, Adaptive Neuro-Fuzzy Inference System (ANFIS) model which is widely used to analyze these hydrological time series data sets obtained from different gauge stations. Adaptive Neuro-Fuzzy Inference Systems (ANFIS) merges the potentiality of Fuzzy Inference Systems (FIS) and Artificial Neural Networks (ANN) to work out problems of different kinds. For this, the data covered eleven years from 2002 to 2012 daily (6-months each year) streamflow period. From our analysis, the root mean square error (RMSE) shows that the ANFIS model generated more satisfactory results than other models with minimum prediction errors. The ANFIS model is very pliable and has feasibilities of integrating the essence of a fuzzy system for real world.

Keywords: neuro-fuzzy network, fuzzy logic, fuzzy inference system, hydrological modeling, river indus, root mean square error, adaptive neuro-fuzzy inference systems

Volume 3 Issue 2 - 2019

Salman Bin Sami,¹ Tanveer Ahmad Siddiqi,²
 Muhammad Jawed Iqbal³

¹Management Science Department, SZABIST (Karachi Campus), Pakistan

²Department of Mathematics, University of Karachi, Pakistan

³Department of Mathematics and Institute of Space and Planetary Astrophysics, University of Karachi, Pakistan

Correspondence: Salman Bin Sami, Management Science Department, SZABIST (Karachi Campus), Karachi-Pakistan, Tel +92-3333175611, Email salmankhan909@live.com

Received: February 20, 2019 | **Published:** April 04, 2019

Introduction

Today floods are very commonly occurring natural disaster in the world. Flood damages all kind of lives i.e. humans, animals and vegetation, and destroys environments. Very heavy losses are sustained to the economy during floods. Looking at the flash flood data for the 2 or 3 decades, the speed of flood has increased due to global warming.¹ Therefore, engineers and planners want to get accurate time of these events. Technology has extensively been used to study the severity of flood damages, i.e. constructively or non-constructively. Constructive measures require huge amounts of money and time. This is due to the fact that things to consider include; making of dams, reservoirs and change to river flow. As far as the non-structural measures are concerned they include flood- forecasting to plan for floods and to provide relief if floods occur. Time series analysis predicts future values by analysis of data taken from the past. Pakistan lies in the Western region of the Asian Subcontinent where it stands 23 degrees to 37 degrees north and from 60 degrees to 77 degrees east. Pakistan has five provinces and a tribal area; Punjab, Sindh, Khyber Pakhtunkhwa, Balochistan & Gilgit-Baltistan. All of these places face different types of weather conditions. There are temperature variations in the four seasons throughout the year. Different conditions including heavy rain, bring floods. The snow melts on the mountains canal so cause floods. Pakistan has long history of floods in River Indus and its nearby areas. Pakistan has coped with many floods in 1950, 1956, 1973, 1976, 1978, 1988, 1992 and 2010. Graphs or sketches of floods between the years 1922 and 2010 vary. The 2010, flood proved to be the most

unfavorable, devastating and calamitous in the history of Pakistan. This research is an attempt to scrutinize the consequences of floods on habitat along the marked off flood areas of Indus River of Pakistan.

Figure 1 shows the neighborhood elevation of the flooded areas in Pakistan. The most significant reason of floods in Pakistan are the monsoon rains in which there is heavy rainfall which is intensified by the snow melt flow in river drainage. Land sliding also results in floods. These floods cause heavy losses including loss of human life and animals, structural damages including bridges, sewerage systems, roadways and canals, raising waterborne diseases and disintegration of farmland resulting in shortage of food. Nevertheless, in the annual report for the year 2010, FFC (Federal Flood Commission) described that highest damage is observed in the lower part of Indus River specially in Sindh where flow is at a peak in the neighboring areas.¹ River Indus is one of the longest rivers of the world. It has seven (7) dams or barrages mainly and has a length of 1800 (miles). Its total discharged area is around 450000 (square miles) out of which 275000 (square miles) lies in semi-arid deserts of Pakistan while the rest comprises of mountains of Pakistan. The River Indus runs through the region of Ladakh which is in Jammu and Kashmir, Gilgit-Baltistan and flows through Pakistan in southerly direction along the whole length of Pakistan to converge into the Arabian Sea near the port city of Karachi in Sindh. Indus River is largely observed by its seven gauge stations which are Tarbela Dam, Chashma Barrage, Kalabagh or Jinnah Barrage, Taunsa Barrage, Guddu Barrage, Sukkur Barrage and Kotri Barrage as shown in Figure 2.

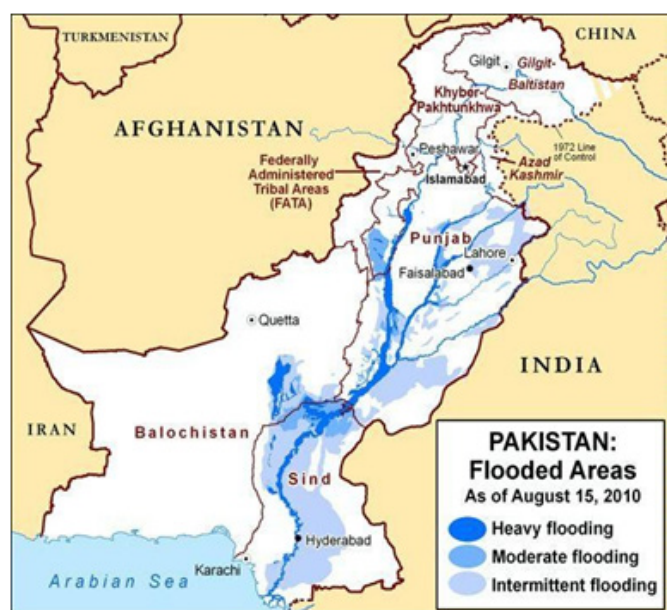


Figure 1 The map of the Pakistan for flooded areas.



Figure 2 The Map of Pakistan for the locus plans of River Indus.

Different flood risk levels; medium, high or very high floods have been observed to occur at these seven gauge stations. Very high-level flood risk is observed at Sukkur Barrage, while Kotri and Guddu Barrages are seen to be fluctuating between high and very high flood risks. Chashma and Taunsa Barrages have high flood risks but Tarbela Dam and Kalabagh have medium flood risks. In this paper, we are proposing to use ANFIS technique for constructing time data series model of river flow for the Basin of River Indus in Pakistan using annual flood peak discharges from gauge sites. In 2005, Ologunorisa and Abawa described that many approaches are used to estimate flood risk like barometrical conditions and equipment, hydrological equipment, socioeconomic elements and a blend of hydro-barometrical and socioeconomic elements in addition to geological data network.² Smith³ states that to estimate flood risk we need to observe the probability of events that took place and its outcomes. Uncertainty in resources of water result in natural dissimilarity of geophysical developments and changes in complex socioeconomic

characteristics. Similarly, Khan et. al. quantified the probabilities of incidents of floods to evaluate River Indus flood risks using archived information about maximum peak discharges in Pakistan.⁴

According to many authors various countries like USA, Korea, Australia, Pakistan, etc., are using barometrical parameters to assess flood risk. A technique for graphing risk of flood based on provincial or local weather information and area has been presented by Kalma and Laughlin.⁵ Khan used a more effective technique called GIS technique to study flood risks in the affected areas near River Indus in Pakistan. This technique included remote sensing, geological data system and digital image processing.⁶ Khan used satellite information yielding the importance of making dams to minimize flood risks. Nawaz and Shafique did the same work on river Jhelum in 2003.⁷ For Dams and Rivers, several forecasting techniques strived successfully to use the Linear and non-linear regression's techniques used by Burn and McBean in 1985,⁸ El-Fandy, Karunanithi and Awwadin 1994.^{9–11} These studies yielded good dams river flow forecasts. Selas and Smith generated synthetic stream flows using hydrological time series modeling.¹² Stedinger and Taylor also assumed stream flow image by developing five different models.¹³ Time series forecasting is a highly demanding and rapidly growing area for researchers in various fields of interest; in engineering, finance, physics and also in medicine. Traditionally, researchers have used several modeling and forecasting techniques for time series, such as, AR (Auto-regressive) method discovered in 1970,¹⁴ ARMA (Auto-regressive moving average) and ARIMA (Auto-regressive integrated moving average) disaggregation models formulated by Valencia and Schaake in 1973,¹⁵ etc. The process to predict weather from time series data is called time series forecasting. These approaches are used only for time series data to generate forecasts. The time series data used is the numeric data type for months of heavy stream flows such as a day or several hours.

In the year 2010, Hassan and Ansari performed some nonlinear methods to predict the continuing behavior of Indus River.¹⁶ Sudheer presented an ANN model for the same purpose but realized that further research on ANN model is required to shape up the peak data flows efficiently.¹⁷ In this study the potential of the ANFIS (adaptive neuro-fuzzy inference system) model has been investigated for the Indus basin. Moreover, Nayaka and Sudheer, in India, investigated a hydrological model of time series for Baitarani River's basin stream flow in Orissa state by ANFIS model.¹⁸ In recent years, the models which are widely used for prediction; adaptive neuro-fuzzy inference system (ANFIS), artificial neural network (ANN) and genetic algorithm are highly systematic tools for complex hydrologic modeling systems. Fuzzy logic procedure to describe the human-like expertise system and to make the decisions was first developed in 1965. The ANFIS model keeps the complete capacity of the ANN approach, and simplicity of the procedure of model building. Adaptive neuro fuzzy inference system (ANFIS) is the fuzzy-rule based or mapping of fuzzy algorithm that was developed by Tagaki-Sugeno-Kang (TSK) and Yasukawain in 1993.¹⁹ For water resources predictions ANFIS has proved to be used on a higher scale during the last decade by various scientists. Nowadays it is used for many applications such as, database management, the forecasting of resources of water and planning.

Neuro-fuzzy model

Neuro-fuzzy model introduces the process of utilizing various learning techniques developed in the literature of neural system to fuzzy modeling or FIS (fuzzy inference system). The fuzzy inference system or FIS structure basically consists of three elements:

- A rulebase, in which we have fuzzy rules selection
- A database element, which designates the functions called membership functions (mfs) for the rules of fuzzy; and
- A reasoning mechanism element, in which we execute the inference strategy based on the defined rules to yield an output results as can see in Figure 3.

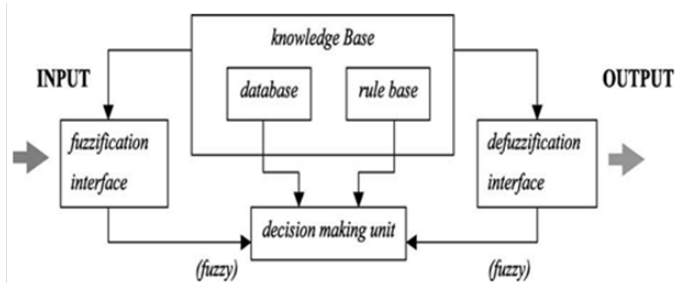


Figure 3 Brisky output by Fuzzy inference system.

From the input area to the output area, fuzzy inference system executes nonlinear plotting. This plotting is carried out by a number of if-then rules of fuzzy, in which each rule reports the behavior of the plotting locally. The parameters of the if-then rules (referred to as antecedents or premises in fuzzy modeling) define a fuzzy region of the input space, and the output parameters (also consequents in fuzzy modeling) specify the corresponding output. Thus, this FIS (fuzzy inference system) relies on the calculated parameters for its proficiency. Nonetheless,²⁰ for the fuzzy set (delineated by antecedents) the choice of pattern leading to an input is not accompanied by any method. But the fuzzy inference rule of system and its structure make it easy to assimilate human proficiency for the system into the process of modeling directly to settle on significant no. of membership functions for every number of input, etc. For parameter calculation there will be leading numerical data. Nowadays, perception of the adaptive network, which is an abstract principle of the ordinary backpropagation neural network, is applied to tackle the parameter identification problem in a FIS. ANN (adaptive neural network) is the multi layered furnish shape structure whose overall output behavior is determined by the value of a collection of modifiable parameters. More specifically, the configuration of an adaptive network is composed of a set of nodes connected through directional links, where each node is a process unit that performs a static node function on its incoming signal to generate a single node output. Every function of node consists of parameterized value function for some malleable parameters. It is noticeable in the adaptive network that the attachments, between nodes with no weights, only show the flow direction of signals affiliated with these links. In 1994, some of the readers mentioned²¹ additional characteristics of adaptive networks. The Jang²² established an understanding strategy and a novel architecture for FIS that uses an understanding algorithm of neural network for building a collection of if-then rules from the determined input–output couples for fuzzy with the convenient MFs.²³ Using this framework for the ANN (adaptive neural networks) for the strategy of flourishing a FIS is known as an adaptive neuro-fuzzy inference system (ANFIS).

Architecture of ANFIS

The common construction for ANFIS is reported in the Figure 3. When designing ANFIS, the main choice for FIS is a model of particular objective system. Many categories of FIS are presented in the literature.^{24–26} Each example is portrayed by part of their consequent parameters. Present research utilizes the fuzzy model of Sugeno.^{26,27} Therefore, the linear equation and parameters are the most

consequent parts of this FIS and can be determined by the method of simple least square error method.

Let's us consider, for FIS, x and y are the two input variables and the one output variable is z . The fuzzy model of Sugeno for the first order has two rules that are if-then rules which are conveyed as,

Rule I: Let x stands for A_1 and y stands for B_1 , then for f_1

$$= p_1 x + q_1 y + r_1 \quad (1)$$

Rule II: Let x stands for A_2 and y stands for B_2 , then for f_2

$$= p_2 x + q_2 y + r_2 \quad (2)$$

Here, for the input variables x & y the MFs are A_1 , A_2 & B_1 , B_2 respectively, and the parameters for the output function are p_1 , q_1 , r_1 and p_2 , q_2 , r_2 . and Figure 4(A) explains fuzzy reasoning technique to obtain the output function i.e. f for this Sugeno model from the provided input variables $[x, y]$. The related equivalent architecture of ANFIS is described in Figure 4(B), where the same layer nodes have corresponding functions. The following are procedures of ANFIS functioning:

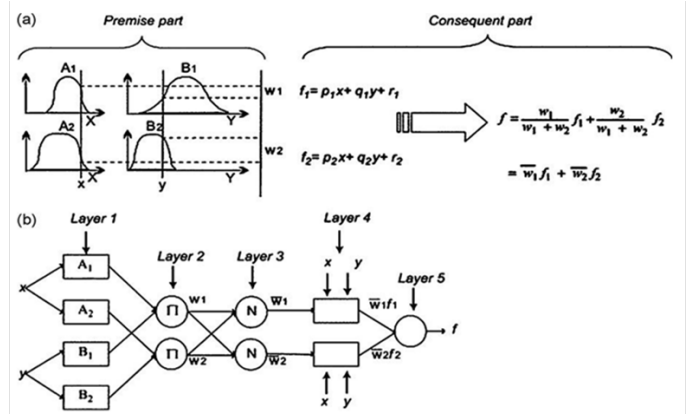


Figure 4(a) Shows the inference system of Fuzzy; (b) Shows an Equivalent architecture of ANFIS.

Layer I: Here, every node of layer 1 creates grades of membership for variable inputs. To define the node output OP_i^I we have:

$$OP_i^I = \mu A_i(x) \text{ where } i = 1, 2 \quad (3)$$

$$OP_i^I = \mu B_{i-2}(y) \text{ where } i = 3, 4 \quad (4)$$

Here, the variable x or y is defined as input node & A_i or B_{i-2} is associated as a fuzzy set, designated by the structure of the membership functions which is any suitable continuous and piecewise differentiable function like Gaussian, generalized bell, triangular & trapezoidal shaped in this node. Considering MF has the shape of generalized bell function, then the output

result OP_i^I will be,

$$OP_i^I = \mu A_i(x) = \frac{1}{(x - c_i)^{b_i}} \quad (5)$$

Here, the shape of MF can be changed by the parameter set $\{a_i, b_i, c_i\}$ from maximum 1 to minimum 0.

Layer II: Here each node multiplies the approaching signals, symbolized by Π , and the rule of firing strength is illustrated by the

output OP_i^2 which can be calculated as:

$$OP_i^2 = w_i = \mu A_i(x) = \mu B_i(y), i = 1, 2 \quad (6)$$

Layer III: For this layer the i^{th} node is symbolized by N_i , which calculates the normalized firing strengths as:

$$OP_i^3 = \bar{w}_i = \frac{w_i}{w_1 + w_2} \quad i = 1, 2 \quad (7)$$

Layer IV: In layer 4 the node i calculates the contribution of the i^{th} rule towards the output of the model, with the following node function:

$$OP_i^4 = \bar{w}_i f_i = \bar{w}_i (p_i x_i + q_i y_i + r_i) \quad (8)$$

Here $\{p_i, q_i, r_i\}$ represents the parameter set and layer 3 output is represented by \bar{w}_i .

Layer V: In layer 5, the total output of ANFIS is calculated by a single node as:

$$OP_i^5 = \text{Total Output} = \frac{\sum_i w_i f_i}{\sum_i w_i} \quad (9)$$

Computation of parameters (Hybrid algorithm)

Accession parameters are the premise parameters $\{a_i, b_i, c_i\}$ in ANFIS, which reports the structure of MFs, and $\{p_i, q_i, r_i\}$ consequent parameters which report the final output of this ANFIS system. In 1986, backpropagation algorithm by Rumelhart, the basic learning rule of an adaptive network based on the gradient descent rule, can be successfully applied to calculate these parameters.²⁸ Nevertheless, Jang insisted that the rule of gradient descent is normally slow and gets trapped in local minima.²² Jang suggested a much faster learning algorithm, which merges the technique of least squares estimate (LSE) and gradient descent to locate parameters as presented below:

The adaptive network has only one output and is supposed to be:

$$\text{output} = F(I, S) \quad (10)$$

Here, the input variables set is I and the parameters set is S . If the existence of function H is there in such a way that the composite function $H \circ F$ in rare elements of S is linear, then the method of least squares is used to recognize these elements. Precisely, parameters of the set S are divided into two sets:

$$S = S1 \oplus S2 \quad (11)$$

Here \oplus sign shows the direct summation in such a way that $S2$ elements possess linear $H \circ F$. Then H is applied to Eq.10, to get:

$$H(\text{output}) = H \circ F(I, S) \quad (12)$$

Which is linear for $S2$ elements; For the provided values of $S1$ elements, the training data P can be plugged into Eq.12 in order to attain the equation of matrix;

$$AX = B \quad (13)$$

Here, the unknown vector X has the parameters of $S2$ elements. Suppose $|S_2| = M$, then $P \times M$, $M \times 1$ and $P \times 1$ are the dimensions of A , X and B . Therefore, for the training data the no. of pairs (P) are generally larger than the no. of linear parameters (M). This is an over-determined problem and usually Eq.13 has no accurate solution.

A LSE for X can be required that reduces the squared error $\|AX - B\|^2$

. Figure 4 shows the architecture of ANFIS and shows that for the provided parametric values of the premise part, the total and final output is conveyed by a linear combination of the parameters of the consequent part. Formally, the final output f can be changed into:

$$f = w_1 f_1 + w_2 f_2 = (w_1 x) p_1 + (w_1 y) q_1 + (w_1) r_1 + (w_2 x) p_2 + (w_2 y) q_2 + (w_2) r_2 \quad (14)$$

Which is linear for parameters (p_1, q_1, r_1, p_2, q_2 and r_2) of the consequent part. In an ANFIS, the overall no. of parameters (S), as a result, can be classified into two sets in such a way that the set of premise part parameters is $S1$ and the set of consequent part parameters is $S2$. Accordingly, the hybrid-learning algorithm combines the methods of least squares and back propagation gradient descent method, which is applied for a fruitful or a powerful search of the excellent parameters of ANFIS. More indicatively, the algorithm of hybrid learning in the forward pass, the node output goes forward until layer 4 and parameters of the consequent part are recognized by method of least squares. The propagation of the error signal backwards in the backward pass and the premise parameters are corrected by the gradient descent technique. As stated earlier, the consequent part parameters thus recognized are excellent under condition that the premise part parameters are confirmed. Consequently, the approach of hybrid algorithm converges much faster since in the original back-propagation method, it minimizes the dimension of the search space. A comprehensive and complete description can be found for this algorithm.²⁹

The defuzzification

Basic ANFIS takes the fuzzy inputs or the crisp inputs, but the total and final outputs are sets of fuzzy. Consequently, a defuzzification approach is required to convert a fuzzy set to a value of crisp. By²¹ the crisp output is usually secured using different defuzzification approaches. The approach of Takagi-Sugeno is applied for the present inspection.²⁶ Xiong, et al.³⁰ do not have a direct defuzzification technique or rather they combine two techniques, the decision of logic and the procedures of defuzzification into the one composite procedure.

Model configuration

The course and broad facts regarding applied models are discussed in this section.

Non-linear time series forecasting

These forecasting are to be made with the help of Neuro-fuzzy. For this purpose, we have observed and evaluated the flow forecast of the three gauge stations (Tarbela Dam in FATA, Chashma Barrage in Punjab and Sukkur Barrage in Sindh). Using Neuro-fuzzy system, we have selected readings of ten years daily stream flow from 2002-2011. Each year contains data of 6-months (from 1st April to 30th September peak months). The daily stream flow data of six peak months of the foresaid stations in 2012 is also chosen to test and verify the model. This data is then given to the system network.

Recruited of the neuro-Fuzzy network

The managerial fuzzy interference network was chosen to be executed by Sugeno model accompanied by the product operator. Two categories of Neuro-fuzzy system network have also been implemented that have different membership functions. Flow discharge is taken at

two values, the present time dt and the past time d_{t-1} for one network and flow discharge is taken at three values that are, the present time dt , the past time $dt-1$ and one step before the past time d_{t-2} . Five different ANFIS models are considered as 2, 3, 4, 6 and 8 membership functions. For all model networks the membership functions are predicted as Gaussian and Triangular shaped once. Hence the first system network uses dt and d_{t-1} as an input for two variables x and y and the second network system uses dt , d_{t-1} and d_{t-2} as an input for three variables x , y and z for the executed fuzzy rule base. Then in accordance with the relation the yielding results are generated as the final outcomes in defuzzification. We can see that ten different system networks are applied and the outcomes are determined and discussed in the coming sections.

Work explanation

Various categories of traditional statistic can be referred to as statistical work explanation. But here we used root mean square error (RMSE). This estimation test can also be estimated by using the below equation.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (d_i^o - d_i^p)^2}{n}} \quad (15)$$

Where, d_i^o is the observed flow of stream at any time t , and d_i^p is the predicted flow of stream at any time t .

Result testing and review

For the present study, we have collected the data from Federal Flood Commission (FFC), situated in Islamabad, which is comprised of eleven years recording at three gauge stations. Data investigation

of three previously mentioned gauge stations reveals that the surveyed data is highly varied and disordered. Flow behavior of Tarbela Dam, Chashma and Sukkur Barrages, from 2010 to 2012, are displayed in Figure 5. Parameters which are interrelated to three foregoing stations are provided in Table 1. From the Table 1, it is clear that standard deviation and the difference between the maxima and minima are large, so that gauge station modeling will be complex. In this regard, for fuzzy network, we consider input data for ten years from year 2002 to year 2011 of the daily stream flow to check different models as training data. Moreover, different models were also tested using testing data for only 1 year i.e. 2012, the daily stream flows (6 peak months' data) for each gauge station. Entered the Input data into the neuro-fuzzy network and results are shown in Tables below.

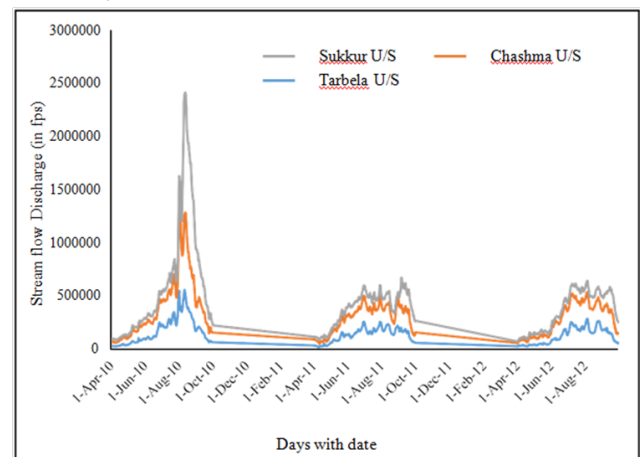


Figure 5 Stream flows of Tarbela, Chashma and Sukkur Barrage through 2010 – 2012.

Table 1 Estimated parameters corresponding to Dam and Barrages for 10 years & 1 year alone (2002–2011 & 2012)

Estimated parameters	Tarbela Dam		Chashma Barrage		Sukkur Barrage	
	2002–2011	2012	2002–2011	2012	2002–2011	2012
Average (fps)	141643.2	123941	177492.7	158845.2	116962.8	81611.91
Standard deviation (fps)	88743.25	76086.45	101397.2	79063.51	124973.6	47645.55
Minimum amount (fps)	18800	26000	23493	26169	16405	15630
Maximum amount (fps)	557100	284000	957309	276745	1130995	214780
Ratio of average to standard deviation (fps)	1.6	1.63	1.75	2.01	0.94	1.71

Outcomes and analysis

For obtaining outcomes, we have processed the network in various approaches and RMSE is used to figure out the outcomes. Tables 2-4 are the calculated results for the three previously mentioned

stations. It is of great importance to note that the mentioned outcomes are informed for Gaussian membership functions only because its calculations and outcomes are better than the Triangular membership functions with error tolerance 0.001.

Table 2A Error evaluation for Tarbela dam in 2012 of flow prediction as testing data

Serial no.	Different input variations	No. of membership functions	RMSE
1	$dt, dt - 1$	2	113.787
2	$dt, dt - 1$	3	108.474
3	$dt, dt - 1$	4	105.902
4	$dt, dt - 1$	6	104.87
5	$dt, dt - 1$	8	97.42

Table Continued....

Serial no.	Different input variations	No. of membership functions	RMSE
6	dt, dt – l, dt - 2	2	101.913
7	dt, dt – l, dt - 2	3	90.465
8	dt, dt – l, dt - 2	4	85.161
9	dt, dt – l, dt - 2	6	68.037
10	dt, dt – l, dt - 2	8	51.114

Table 2B Error evaluation for Tarbela dam in 2002-2011 of flow prediction as training data

Serial no.	Different input variations	No. of membership functions	RMSE
1	dt, dt – l	2	155.639
2	dt, dt – l	3	134.878
3	dt, dt – l	4	129.603
4	dt, dt – l	6	125.749
5	dt, dt – l	8	124.374
6	dt, dt – l, dt - 2	2	129.476
7	dt, dt – l, dt - 2	3	117.745
8	dt, dt – l, dt - 2	4	114.935
9	dt, dt – l, dt - 2	6	110.249
10	dt, dt – l, dt - 2	8	107.461

Table 3A Error evaluation for Chashma Barrage in 2012 of flow prediction as testing data

Serial no.	Different input variations	No. of membership functions	RMSE
1	dt, dt – l	2	117.188
2	dt, dt – l	3	109.666
3	dt, dt – l	4	107.837
4	dt, dt – l	6	101.392
5	dt, dt – l	8	95.455
6	dt, dt – l, dt - 2	2	114.15
7	dt, dt – l, dt - 2	3	100.935
8	dt, dt – l, dt - 2	4	97.08
9	dt, dt – l, dt - 2	6	77.261
10	dt, dt – l, dt - 2	8	69.699

Table 3B Error evaluation for Chashma Barrage in 2002-2011 of flow prediction as training data

Serial No.	Different input variations	No. of membership functions	RMSE
1	dt, dt – l	2	186.708
2	dt, dt – l	3	169.368
3	dt, dt – l	4	162.853
4	dt, dt – l	6	156.526
5	dt, dt – l	8	156.105
6	dt, dt – l, dt - 2	2	158.58
7	dt, dt – l, dt - 2	3	155.167
8	dt, dt – l, dt - 2	4	148.86
9	dt, dt – l, dt - 2	6	141.91
10	dt, dt – l, dt - 2	8	139.753

Table 4A Error evaluation for Sukkur Barrage in 2012 of flow prediction as testing data

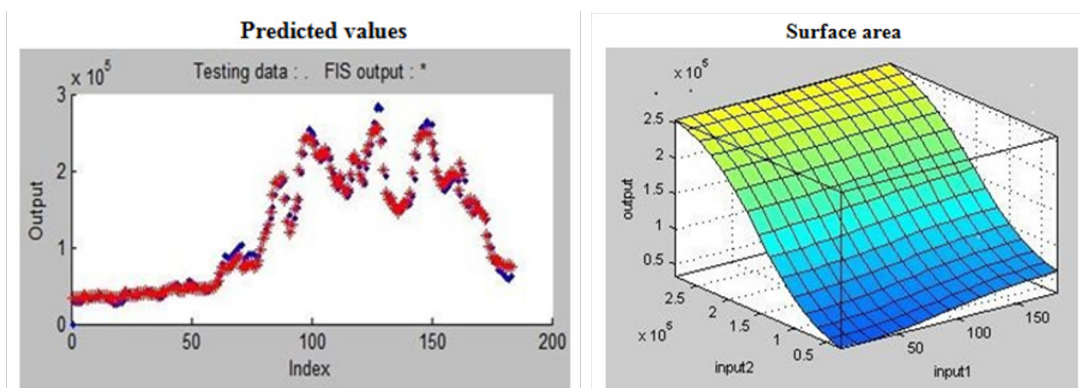
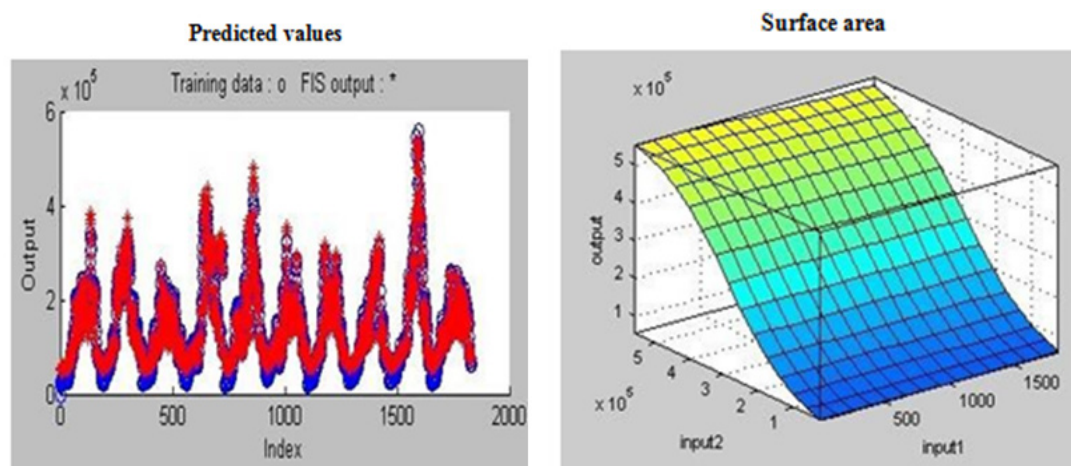
Serial no.	Different input variations	No. of membership functions	RMSE
1	dt, dt – l	2	86.433
2	dt, dt – l	3	80.395
3	dt, dt – l	4	75.492
4	dt, dt – l	6	68.732
5	dt, dt – l	8	63.513
6	dt, dt – l, dt - 2	2	77.205
7	dt, dt – l, dt - 2	3	69.544
8	dt, dt – l, dt - 2	4	60.337
9	dt, dt – l, dt - 2	6	48.093
10	dt, dt – l, dt - 2	8	44.276

Table 4B Error evaluation for Sukkur Barrage in 2002-2011 of flow prediction as training data

Serial No.	Different input variations	No. of membership functions	RMSE
1	dt, dt - 1	2	170.088
2	dt, dt - 1	3	149.33
3	dt, dt - 1	4	138.725
4	dt, dt - 1	6	118.698
5	dt, dt - 1	8	112.94
6	dt, dt - 1, dt - 2	2	135.046
7	dt, dt - 1, dt - 2	3	114.976
8	dt, dt - 1, dt - 2	4	106.599
9	dt, dt - 1, dt - 2	6	93.434
10	dt, dt - 1, dt - 2	8	87.921

Tarbela dam results

Results of Tarbela dam for daily stream flow prediction for the year 2012 and in years 2002-2011 are presented in Table 2A and 2B

**Figure 6A** The predicted values and surface area for Tarbela Dam in year 2012.**Figure 6B** The predicted values and surface area for Tarbela Dam in year 2002 to 2011.

respectively for different inputs and different membership functions. Curves in Figure 6A, indicate the predicted values and surface area for Tarbela dam using observed data and applied neuro-fuzzy technique for the year 2012. In the graph above of predicted values in Figure 6A, the output data on y- axis means stream data values in *fps* and on x-axis index means no. of days i.e. 183 peak days for the year 2012 and also Figure 6B indicates the predicted values and surface area for the years 2002 to 2011. In the graph above of predicted values in Figure 6B, the output data on y- axis means stream data values in *fps* and on x-axis index means no. of days i.e. 1830 peak days for above mentioned years.

Chashma barrage results

Same as previous calculations for Tarbela Dam, we can see the results of Chashma barrage for flow prediction in year 2012 and in years 2002 - 2011 are presented in Table 3A and 3B. Curves in Figure 7A, indicate the predicted values and surface area for Chashma barrage using the observed data and applied neuro-fuzzy technique for the year 2012 and Figure 7B shows the same for the years 2002 to 2011.

Sukkur barrage results

The predicted results of Sukkur barrage of daily flow for different inputs and MFs in years 2012 and in years 2002 - 2011 are presented in Tables 4A&4B respectively. Curves in Figure 8A, indicate the predicted values and surface area for Sukkur barrage using the observed data and applied neuro-fuzzy technique for the year 2012 and Figure 8B is for the years 2002 to 2011.

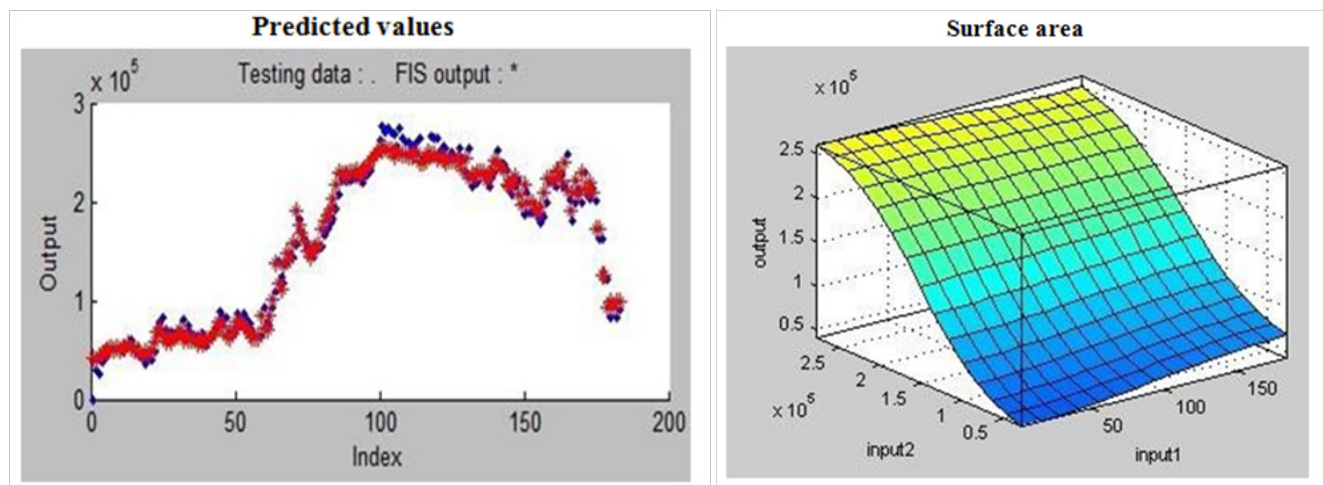


Figure 7A The predicted values and surface area for Chashma Barrage in year 2012.

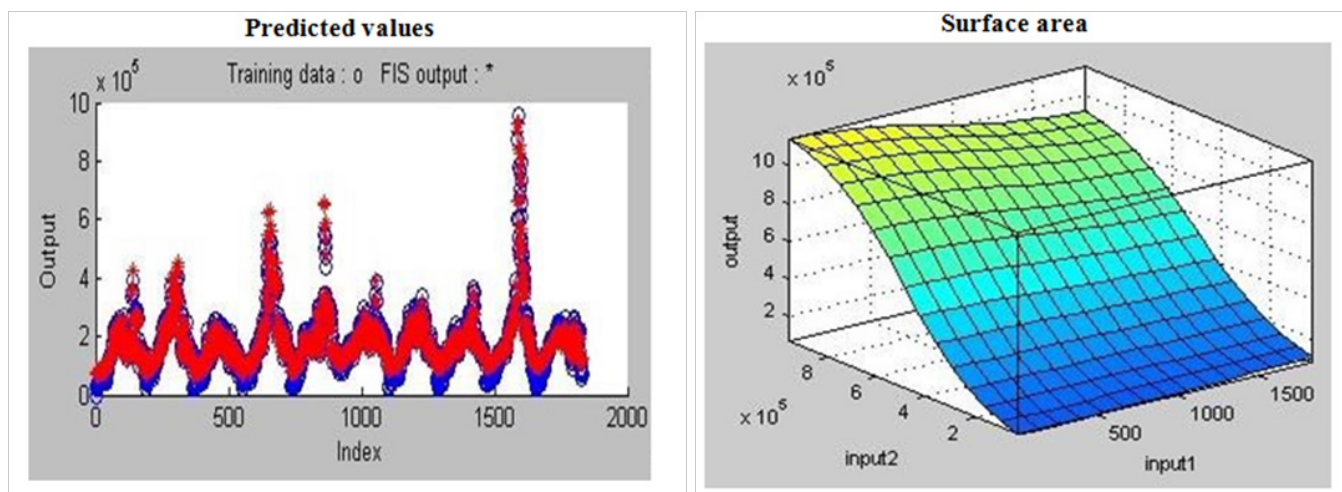


Figure 7B The predicted values and surface area for Chashma Barrage in year 2002 to 2011.

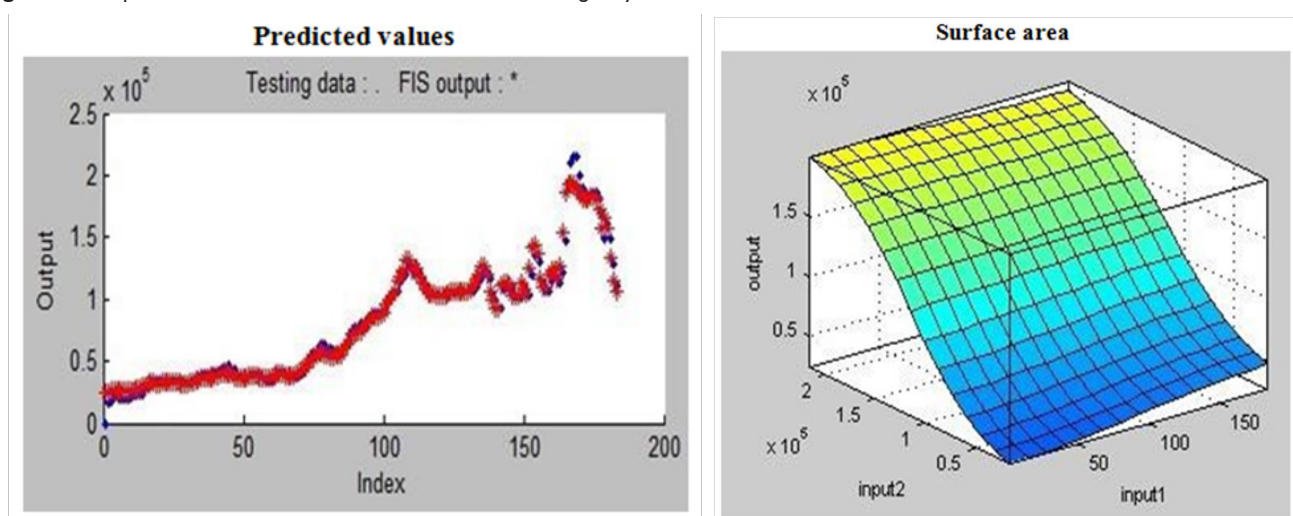


Figure 8A The predicted values and surface area for Sukkur Barrage in year 2012.

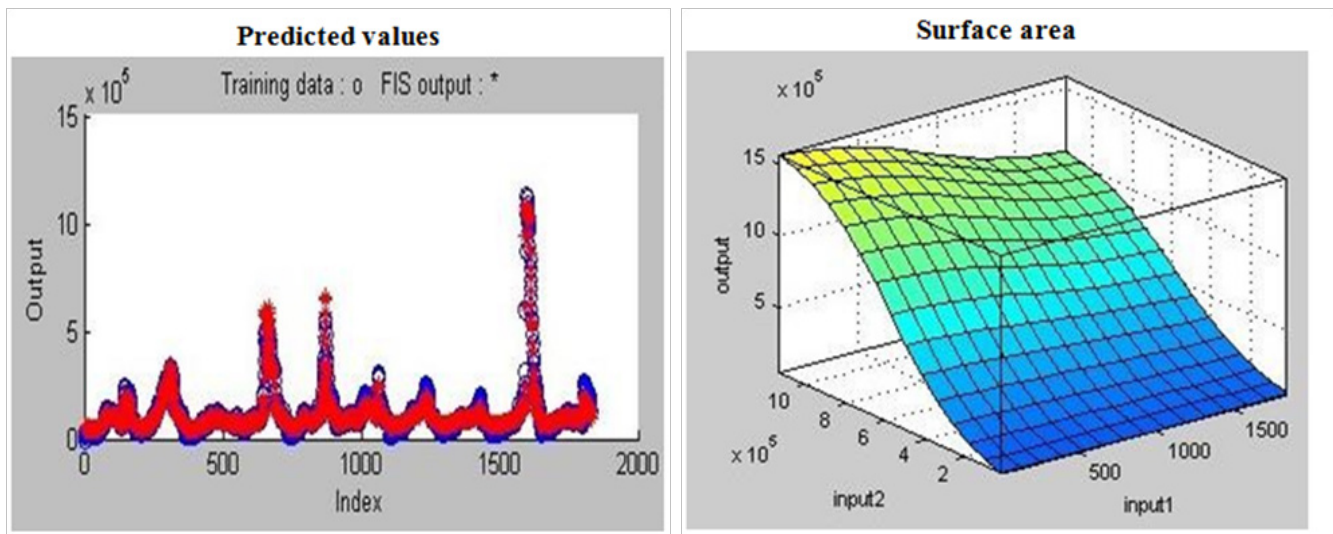


Figure 8B The predicted values and surface area for Sukkur Barrage in year 2002 to 2011.

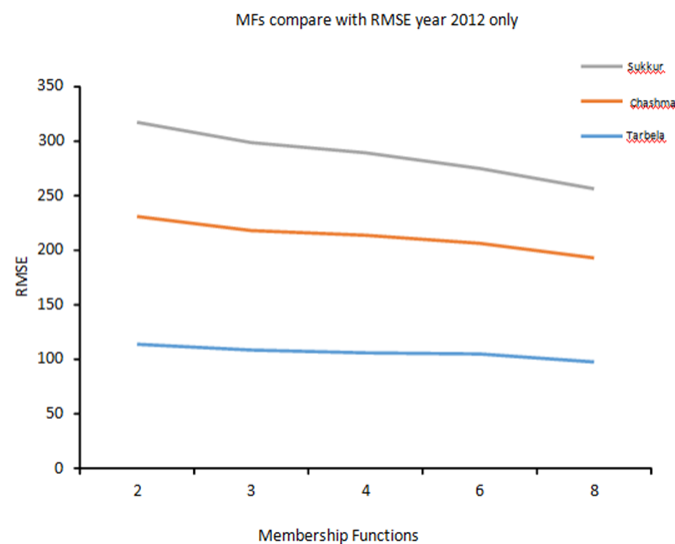


Figure 9A Structure for 2-inputs d_t & d_{t-1} .

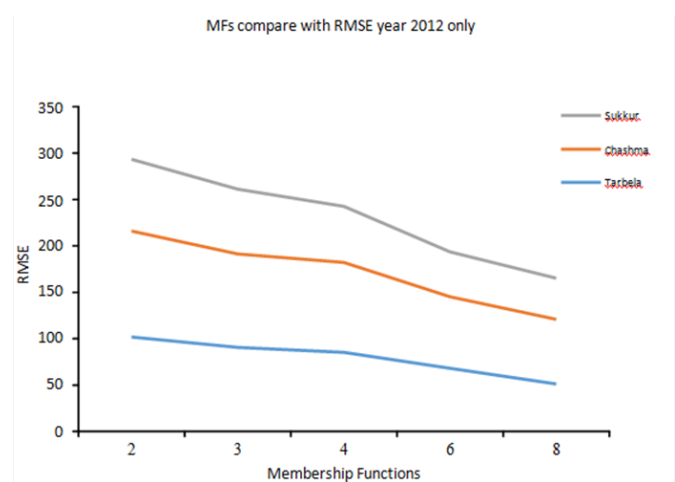


Figure 9B Structure for 3-inputs d_t , d_{t-1} & d_{t-2} .

Discussion on flood analysis results

We can see the modeled results in Table 2A–Table 4B. It is apparent that the results are much better and faster by the application of the ANFIS model. Comparing the results of these three stations shows from the ratio of average to standard deviation and by calculating the error RMSE, a discrete reduction in the errors in all by increasing the inputs than MFs is shown in this Neuro-fuzzy model. It is surprising that the RMSE error is slowly decreasing with an increase in the membership functions, while increasing the number of inputs makes efficiency decrease in the calculated error. So one thing should be apparent that suitable number of membership functions should be interpreted and for this the rise in inputs is a better option for predicted values.

Furthermore, we find that the results of the three stations indicate that the best results can be attained by the structure of neuro-fuzzy model and the best outcomes with minimum errors are for the year 2012, for Tarbela Dam using data inputs d_t , d_{t-1} , d_{t-2} and eight membership functions (MFs). Table 2A shows these results. From Table 3A, it can be seen that for Chashma barrage the structure with minimum error using data inputs d_t , d_{t-1} , d_{t-2} and eight membership functions (MFs) was obtained for the year 2012. But for Sukkur barrage where the flood risk is very high, Table 4b shows an enormous decrease or drop in the error i.e. 87.92 for large amount of data

i.e. 10 years data used on same structure with d_t , d_{t-1} , d_{t-2} inputs and eight membership functions (MFs) while the error in the structure with d_t , d_{t-1} inputs and two membership functions is calculated as 170.087. In Table 1, very low of ratio of average to standard deviation for the 10 years' data is shown for Sukkur barrage.

Figure 9A & 9B shows that the performance evaluation for the three-gauge station relies on the no. of membership functions for two-input structure and three-input structure. It is visible that better results might appear by increasing the input data $\{d_t, d_{t-1}, d_{t-2}, \dots, d_{t-n}\}$ to the network instead of increasing the membership functions.

Conclusion

An Adaptive Neuro-Fuzzy Inference System (ANFIS) has been executed at three different gauge stations to presume cyclic behavior of

river flow. Input variables and membership functions were applied using two different neuro-fuzzy systems performed 5-times with different membership functions; 2, 3, 4, 6, 8 with 2 and 3 inputs data. To do so, we have accumulated ten years' stream flow data of three gauge stations of River Indus flow and have used it as training data. One-year stream flow behavior is executed as testing data. The system is executed in different levels. For this system, it is apparent that better results may be obtained by raising the no. of inputs to the network instead of raising the no. of membership functions. Comparison between the observed and the predicted data values and this analysis of qualitative data achievement of the model stipulates that ANFIS can be utilized fruitfully for time series hydrological data. The ANFIS model is very adaptable and has possibilities of integrating the fuzzy nature of real world systems.

Acknowledgment

None.

Conflicts of interest

The authors declare that there is no conflict of interest.

References

1. Annual Flood report. *From Federal Flood Commission*. Department of Ministry of Water and Power, and Government of Pakistan–Islamabad. 2010.
2. Ologunorisa TE, Abawua MJ. Flood risk assessment: A review. *App Sci Environ Mgt J*. 2005;9(1):57–63.
3. Smith K. *Environmental hazards*. Rutledge, London. 1996.
4. Khan B, Iqbal MJ, Yosufzai MAK. Flood risk assessment of River Indus of Pakistan. *Arab Journal of Geoscience*. 2011;4(1–2):115–122.
5. Kalma JD, Laughlin GP. Flood risk mapping for landscape planning: a methodology. *International Journal of Theoretical and Applied Climatology*. 1990;42:41–51.
6. Khan G. Flood hazard assessment and mitigation along River Indus from Chashma Barrage to Sukkur Barrage using satellite image. *M.Phil. thesis, Institute of Space and Planetary Astrophysics, University of Karachi*. 2007.
7. Nawaz F, Shafique M. Data integration for flood risk analysis by using GIS/RS as tools. In *proceedings of MAP Asia, Second Annual Asian Conference and Exhibition in the field of GIS, GPS, Aerial Photography and Remote Sensing*. 2003.
8. Burn D, McBean E. River flow forecasting model for the Sturgeon River. *J Hydraulic Eng*. 1985;111(2):316–333.
9. Awwad H, Valdes J, Restrepo P. Streamflow forecasting for Han River Basin, Korea. *J Water Resour Planning Manage*. 1994;120(5):651–673.
10. El-Fandy M, Ashour Z, Tael S. Time series models adoptable for forecasting Nile floods and Ethiopian rainfalls. *Bull Am Meteorol Soc*. 1994;75(1):1–12.
11. Karunanithi N, Grenney WJ, Whitley D, et al. Neural networks for river flow prediction. *J Comput Civ Eng, ASCE*. 1994;8(2):10–220.
12. Salas JD, Smith RA. Physical basics of stochastic models of annual flows. *Water Resour Res*. 1981;41:428–430.
13. Stedinger JR, Taylor MR. Synthetic stream–flow generation model verification and validation. *Water Resource Res*. 1982;18(4):909–918.
14. Box GEP, Jenkins GM. *Time Series Analysis: Forecasting and Control*. Holden–Day, Oakland, CA. 1970.
15. Valencia RD, Schaake JC. Disaggregation processes in stochastic hydrology. *Water Resour Res*. 1973;9(3):58–585.
16. Daniyal Hassan, Rakhshinda Bano, Kamran Ansari, et al. *Hydrology of Upper Indus Basin*. 2010.
17. Sudheer KP, Nayak PC, Rangan DM. Rainfall runoff modeling using artificial neural network technique, Report No. CS/AR–16/1999–2000. *National Institute of Hydrology, Roorkee, India*. 2000.
18. Nayak PC, Sudheer DM, Rangan, et al. A neuro–fuzzy computing technique for modeling hydrological time series. *J Hydrol*. 2004;291(1–2):52–66.
19. Sugeno M, Yasukawa T. A fuzzy–logic based approach to qualitative modeling. *IEEE Transactions on Fuzzy Systems*. 1993;1(1):7–31.
20. Ojala T. *Neuro–Fuzzy systems in control*. MSc. Thesis, Tampere University of Technology, Tampere, Finland. 1995.
21. Brown M, Harris C. *Neuro–fuzzy Adaptive Modeling and Control*. Prentice Hall. 1994.
22. Jang JSR. Rule extraction using generalized neural networks. In *Proceedings of the fourth IFSA World Congress. Volume for Artificial Intelligence*. 1991;4:82–86.
23. Jang JSR. ANFIS: Adaptive network based fuzzy inference systems. *IEEE Trans. Syst., Man Cybernet*. 1993;23(3):665–685.
24. Mamdani EH, Assilian S. An experiment in linguistic synthesis with a fuzzy logic controller. *Int J Man Mach Stud*. 1975;7(1):1–13.
25. Tsukamoto Y. An Approach to Fuzzy Reasoning Method. In: Gupta M, et al. *Advances in Fuzzy Set Theory*, North–Holland, and Amsterdam. 1979.
26. Takagi T, Sugeno M. Fuzzy identification of systems and its application to modeling and control. *IEEE Trans Syst Man Cybernet*. 1985;15(1):116–132.
27. Sugeno M, Kang GT. Structure identification of fuzzy model. *Fuzzy Sets Syst*. 1988;28(1):15–33.
28. Rumelhart DE, Hinton and GE, Williams RJ. Learning representations by back–propagating errors. *Nature*. 1986;323:533–536.
29. Jang JSR, Sun CT. Neuro–fuzzy modeling and control. *Proceedings IEEE*. 1995;83(3):378–406.
30. Xiong LH, Shamseldin AY, O'Connor KM. A nonlinear combination of the forecasts of rainfall–runoff models by the first order Takagi–Sugeno fuzzy system. *Journal of Hydrology*. 2001;245(1–4):196–217.