

Rainfall-Runoff Modeling of Sutlej River Basin (India) Using Soft Computing Techniques

Athar Hussain, CBP Government Engineering College, Delhi, India

Jatin Kumar Singh, School of Engineering, Gautam Buddha University, Greater Noida, India

A. R. Senthil Kumar, National Institute of Hydrology, Roorkee, India

Harne K R, CBP Government Engineering College, Delhi, India

ABSTRACT

The prediction of the runoff generated within a watershed is an important input in the design and management of water resources projects. Due to the tremendous spatial and temporal variability in precipitation, rainfall-runoff relationship becomes one of the most complex hydrologic phenomena. Under such circumstances, using soft computing approaches have proven to be an efficient tool in modeling of runoff. These models are capable of predicting river runoff values that can be used for hydrologic and hydraulic engineering design and water management purposes. It has been observed that the artificial neural networks (ANN) model performed well compared to other soft computing techniques such as fuzzy logic and radial basis function investigated in this study. In addition, comparison of scatter plots indicates that the values of runoff predicted by the ANN model are more precise than those found by RBF or Fuzzy Logic model.

KEYWORDS

Artificial Neural Networks, Fuzzy Logic, Multilayer Perception, Radial Basis Function, Rainfall-Runoff Modelling

1. INTRODUCTION

The demand for water has increased due to population growth, urbanization and industrialization as a result of which watersheds and river systems have been altered. This will cause greater damage to property and result in loss of life if flooding occurs. Therefore, it's critically important to successfully plan, design and manage these water resources systems. Determining the relationship of transformation of precipitation to runoff is an important issue in surface hydrology. A rainfall runoff model is required to obtain the relationship between rainfall and runoff. This model helps in forecasting river runoff values that can be used in hydrology and hydraulic engineering design and water management purposes.

Generally, three types of models, including deterministic (physical) models, conceptual models and empirical/systems theoretic/black-box models, are being used by hydrologists in order to model this relationship. The deterministic (physical) models describe the relationship using physical law of mass and energy transfer (Dawson and Wilby, 2001). In contrast, in conceptual models, instead of using physical laws of mass and energy transfer, a simplified, but a plausible or reliable conceptual representation of the underlying physics is adopted (Jain and Srinivasulu, 2006).

DOI: 10.4018/IJAEIS.2019040101

This article, originally published under IGI Global's copyright on April 1, 2019 will proceed with publication as an Open Access article starting on February 4, 2021 in the gold Open Access journal, International Journal of Agricultural and Environmental Information Systems (converted to gold Open Access January 1, 2021), and will be distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

An alternative modeling approach for hydrological processes such as rainfall-runoff process is the empirical/systems theoretic/black-box models, which tries to find a relationship between historical inputs and outputs (ASCE Task Committee, 2000a) without detailed understanding of the physics involved in the process under investigation, such as artificial neural networks (ANNs).

1.1. Artificial Neural Networks

In the recent past, Artificial Neural Networks (ANN) modeling has gained significant attention because of its ability to provide better solutions when applied to complex systems that have been poorly described or understood and where input is incomplete or uncertain by nature over other traditional modeling techniques such as empirical models, statistical models (autoregressive, autoregressive moving average models) and physical based models. The advantages of ANN models over physically based models have been described in detail by French et al. (1992).

One of the most groundbreaking rediscoveries was that of back propagation techniques (which were conceived by Rosenblatt) by Rumelhart et al. (1986). In many previous studies, Multilayer Feed Forward Back Propagation Neural Network (MLFBPN) was commonly adopted and it proved to be most powerful tool to 80 percent of practical application in all field of hydrologic engineering and sciences (Hsu et al., 1995; Smith and Sli, 1995). In the present study, a multilayered feed forward back propagation neural network model is developed with rainfall, evaporation and lagged runoff as input to predict runoff. The number of hidden layers and the number of nodes in each hidden layer are usually determined by a trial-and-error procedure. The nodes within neighboring layers of the network are fully connected by links.

1.2. Radial Basis Function Network

The RBF network can be considered as a two-layer feed forward artificial neural network in which the hidden layer performs a fixed non-linear transformation with no adjustable internal parameters. The output layer, which contains the only adjustable weights in the network, then linearly combines the outputs of the hidden neurons (Chen, et al., 1991).

The hidden layer function $h(x)$ is the Gaussian activation function with the parameters r (the radius or standard deviation) and c (the center or average taken from the input space) defined separately at each RBF unit as shown in equation below (Equation 1):

$$h(x) = \exp\left(-\frac{(x-c)^2}{r^2}\right) \quad (1)$$

and the output layer function is of the form of equation (Equation 2):

$$f(x) = \sum_{j=1}^m w_j h_j(x) \quad (2)$$

The function 'newrb' iteratively creates a radial basis network with one neuron at a time. Neurons are added to the network until the sum-squared error falls beneath an error goal or a maximum number of neurons has been reached. The call for this function is:

net = newrb(P, T, GOAL, SPREAD)

The function 'newrb' takes matrices of input and target vectors P and T, and design parameters GOAL and SPREAD, and returns the desired network. At each iteration the input vector that results in lowering the network error the most is used to create a radbas neuron. The error of the new network is checked, and if low enough newrb is finished. Otherwise the next neuron is added. This procedure is repeated until the error goal is met or the maximum number of neurons is reached. It is important that the spread parameter be large enough that the radbas neurons respond to overlapping regions of the input space, but not so large that all the neurons respond in essentially the same manner.

1.3. Fuzzy Logic

A Takagi-Sugeno Fuzzy Inference System(FIS) has been developed by using the subtractive clustering (Chiu 1994) algorithm integrated with a linear least squares estimate algorithm for the modeling of runoff. Clustering is a method where data points can belong in more than one group ("cluster"). Clustering divides data points into groups based in similarity between items and looks to find patterns or similarity between items in a set. Items in clusters should be as similar as possible to each other and as dissimilar as possible to items in other groups. Cluster Estimation Method can be used to estimate sums and products when the numbers you are adding or multiplying cluster near or is close in value to a single number.

The FIS model has been developed on the basis of the assumption that the cluster estimation method, when applied to a collection of input and output data, produces cluster centers in which each cluster center is in essence a prototypical data point that represents a characteristic behavior of the system. Hence, each cluster center can be used as the basis of a rule that illustrates the system behavior (Chiu 1994). The Fuzzy Logic Toolbox (MATLAB) (The Mathworks 2001) is a library of functions implementing a framework for creating, editing, and executing fuzzy inference systems. This toolbox was used to develop the Fuzzy Logic model for modeling the runoff.

1.4. Application of Soft Computing Techniques in the Rainfall-Runoff Process

Rajurkar, et al. (2002) studied the application of artificial neural network (ANN) methodology for modeling daily flows during monsoon flood events for a large size catchment of the Narmada River in Madhya Pradesh, India. They found that a linear multiple-input single-output (MISO) model coupled with the ANN provided a better representation of the rainfall-runoff relationship in such large size catchments compared with linear and nonlinear MISO models. Campolo, et al. (2003) forecasted flood in Arno River by using feed forward neural network approach with standard back propagation training algorithm. They used the information of rainfall, hydrometric data and dam operation at the basin, Italy, to predict the hourly water level variations. They used two years data with some special treatment with as inputs. In the last decade, ANNs have been successfully employed in modeling a wide range of hydrologic processes, including rainfall-runoff processes. (Rajurkar et al., 2004; Ahmad and Simonovic, 2005; De Vos and Rientjes, 2005; Kisi, 2007) found acceptable performance of Artificial Neural Networks (ANNs) in rainfall-runoff modeling. Solaimani (2009) developed Artificial Neural Network (ANN) to modeling the rainfall-runoff relationship in a catchment area located in a semiarid region of Iran. The applications of the feed-forward back propagation for the rainfall forecasting with various algorithms with performance of multilayer perceptions has been illustrated in this study.

Kumar, et al. (2004) have studied the performance of MLP and RBF type neural network models developed for rainfall-runoff modeling of two Indian River basins. The performance of both the MLP and RBF network models were comprehensively evaluated in terms of their generalization properties, predicted hydrograph characteristics, and predictive uncertainty. Merits and limitations of networks of both models were discussed. Suhaimi et al. (2009) studied on Rainfall-runoff modeling using Radial Basis Function Neural Network for Sungai Tinjar Catchment, Miri, and Sarawak. The RBF network developed in this study has successfully modeled rainfall runoff relationship in Tinjar Catchment, Miri, and Sarawak with an accuracy of about 98.3%. Vivekanandan (2014) studied prediction of Rainfall using MLP and RBF networks. This paper illustrates the use of ANN for prediction of rainfall at

Atner, Multai and Dharni stations. Multi-Layer Perceptron (MLP) and Radial Basis Function (RBF) networks are applied to train the network data. Model performance indicators such as correlation coefficient, model efficiency and root mean square error are used to evaluate the performance of the MLP and RBF networks. The paper presents the MLP network is better suited for prediction of rainfall for Atner and Multai whereas RBF network for Dharni.

Deka and Chandramouli (2005) proposed a new approach combining FL and ANNs, which is referred to as fuzzy neural networks (FNN), for river flow prediction. Takagi and Sugeno first introduced their method of fuzzy inference in 1985. In recent years the Takagi–Sugeno fuzzy method (TS) has been used extensively in hydrology and has given many satisfactory results (Vernieuwe et al., 2005, Jacquin and Shamseldin, 2006; Lohani, et al., 2006). Alvisi, et al. (2006) stated that the FL models have a limited capacity for dealing with too detailed information compared to ANN models and this result is in line with other hydrological studies based on the fuzzy rules system that are generally characterized by a low number (from 2 to 5) of input variables. Recently, FL has been used in the field of hydrology and water resources. Bárdossy (2006) proposed a fuzzy unit hydrograph to account for the number of uncertainties raised from both model assumptions and data acquisition in representing the rainfall–runoff transformation. Tayfur and Singh (2006) used ANN and fuzzy logic models for simulating event-based rainfall–streamflow. Firat and Güngör (2007a; 2007b) used stream-flow records of two stations in the Great Menderes basin for adaptive estimation by using the Takagi-Sugeno (TS) approach (Takagi and Sugeno, 1985). They concluded that the TS-type fuzzy model has better skill in estimating the current value from two antecedent values, compared to ANNs. Mukarji, et al. (2009) apply the ANN, adaptive neuro-fuzzy inference system (ANFIS) and ANFGI mode to forecast stream-flow for Ajay River Basin in Jharkhand, India and results observed that ANFIS model predicts better than the ANN model in most of the cases. Gowda and Mayya (2014) apply fuzzy logic model for predicting stream-flow for Nethravathi River basin is located in Dakshina Kannad applying different membership functions and results found that, fuzzy inference system using triangular membership function show a good performance compared to other models developed.

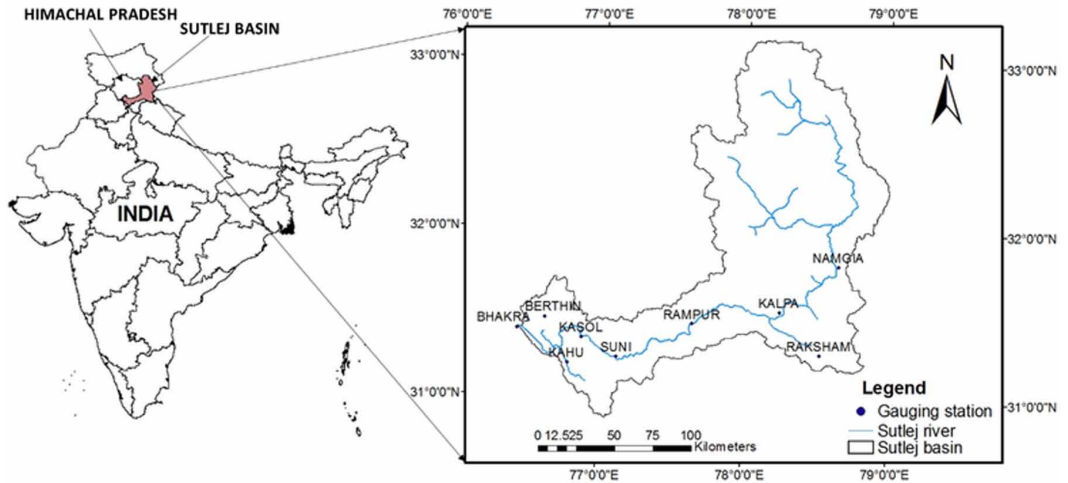
2. MATERIALS AND METHODS

2.1. Study Area

The Sutlej River is one of the main tributaries of the Indus River System and is located in the western Himalayan Region. The Sutlej River originates from Mansarowar Lake in Tibet at an elevation of about 4572 m and is a major tributary of the River Indus. Sutlej plays a key role in the economy of northern India where two out of three persons depend upon agriculture and allied activities for their livelihood. The entire Sutlej basin lies between latitudes 30°N and 33°N and longitudes 76°E and 83°E. The Sutlej River enters India near Shipkila at an elevation of about 6,608 meters and continues to flow in Himachal Pradesh through Wangtoo and Kian before reaching Bhakra reservoir, where the India's highest gravity dam has been constructed. The total length of the river is 1,448 km. The total drainage area of the Sutlej River up to Bhakra Reservoir is about 56,000 km². The Indian part of Sutlej basin upstream of Bhakra dam is about 20,275 km². The present study has been carried out in a part of Sutlej River basin that is confined in the hilly State of Himachal Pradesh, India.

This basin lies between 30°51'23"N and 33°6'30"N latitudes and 76°26'11"E and 78°59'32"E longitudes as shown in figure 1. The elevation of the upper basin varies from about 500 m at Bhakra dam to 7000 m. However, only very small area exists above 6000 m. The mean elevation of the basin is about 3600 m. Although, the basin covers outer, middle and greater Himalayan ranges, the major part of basin lies in the greater Himalayan ranges. Owing to large differences in the relief, the basin is characterized by the diversified climatic patterns. Westerly weather disturbances produce most of the precipitation during winter in the middle and upper parts of the upper basin.

Figure 1. Index map of Sutlej basin area used in present study



The mean annual rainfall (excluding snow) in the outer, middle and outer Himalayan ranges of the basin is about 1300, 700 and 200 mm, respectively (Singh and Kumar, 1997b). The distribution of rainfall indicates that the rainfall is mostly concentrated in the lower part of the basin and has little influence in the greater Himalayan range. The snowline is highly variable, descending to an elevation of about 2000 m during winter and retreats to above 4500 m after the ablation period.

2.1.1. Data Collection and Analysis

The basic data required for the rainfall-runoff modeling are rainfall information, time series of discharge data and evaporation data of the study area. The long-term rainfall, discharge and evaporation data used in the present study has been collected and supplied by Bhakra Beas Management Board (BBMB) for eight gauging sites; Berthin, Bhakra, Kahu, Kasol, Namgia, Raksham, Rampur, Suni. Daily rainfall data for sixteen years starting from 01 January 1988 to 31 December 2004 has been collected from eight rain gauge stations whereas the discharge values for the same period was collected from four discharge measuring stations and evaporation was collected for one measuring station. These data sets were analyzed and transformed for proper use as input to the models. The locations of given stations are shown through Figure 1 earlier while the details of the stations has been summarized in Table 1.

2.2. Methodology

2.2.1. Selection of Input Vector for Model Development

The ANN model for the prediction of runoff is normally developed by using the antecedent rainfall, evaporation, and discharge values of upstream stations as input vector (Cobaner et al. 2009). Determining the number of antecedent rainfall, evaporation, and discharge values involves finding the lags of rainfall, evaporation, and discharge values that have significant influence on the predicted runoff. These influencing values corresponding to the different lags can be very well-established through the statistical analysis of the data series.

The input vector is selected generally by the trial-and-error method. The simple correlation between the dependent and independent variables helps in selecting the significant input vector to the model. To identify the input vector, detailed correlation analysis of the following variables was done: (i) Discharge at Bhakra and rainfall values at Bhakra, Berthin, Kahu, Kasol, Namgia, Raksham,

Table 1. Location details of the stations considered for the study in Sutlej basin

S.No.	Station	Height above M.S.L. (feet)	Latitude	Longitude	Duration of Data Availability Year
1.	Bhakra	1700	31°24'32"N	76°26'28"E	1988 – 2004
2.	Berthin	2155	31°28'15"N	76°37'20"E	1988 – 2004
3.	Kasol	2170	31°21'25"N	76°52'42"E	1988 – 2004
4.	Kahu	2130	31°12'13"N	76°47'15"E	1988 – 2004
5.	Suni	2150	31°14'15"N	77°06'30"E	1988 – 2004
6.	Rampur	3200	31°27'15"N	77°38'40"E	1988 – 2004
7.	Raksham	10269	31°17'48"N	78°32'10"E	1988 – 2004
8.	Namgia	9547	31°48'36"N	78°39'23"E	1988 – 2004

Note: M.S.L. = Mean Sea Level

Rampur and Suni (ii) Discharge at Bhakra with evaporation at Bhakra. (iii) Discharge at Bhakra with discharge at Kasol, Rampur, Suni.

The correlation analysis helps to find out the possible input variable for the modeling, but it does not give the exact lag values. Generally, the trial-and-error method is adapted to find out the significant lag values of the input variable. Sudheer et al. (2002) presented a statistical procedure that avoids the trial-and error procedure. The author reported that the statistical parameters, such as autocorrelation function (ACF), partial autocorrelation function (PACF), and cross-correlation function (CCF), could be used to find out the significant lag values of input variables. These statistical parameters are used in this study to find out lag values of input variables.

The auto-correlation coefficient (Salas et al., 1980) is defined as in equation 3:

$$r_k = \frac{\sum_{t=1}^{N-k} (x_t - \bar{x}_t)(x_{t+k} - \bar{x}_{t+k})}{\left[\sum_{t=1}^{N-k} (x_t - \bar{x}_t)^2 \sum_{t=1}^{N-k} (x_{t+k} - \bar{x}_{t+k})^2 \right]^{1/2}} \quad (3)$$

Where $r_k = \text{lag} - k$ correlation coefficient, the serial correlation coefficient or the autocorrelation function (ACF); x_t = time series for $t = 1, \dots, N$; x_{t+k} = lagged time series for $t = 1, \dots, N-k$; \bar{x}_t = sample mean for $t = 1, \dots, N$; \bar{x}_{t+k} = sample mean for $t = 1, \dots, N-k$; N = sample size.

The partial autocorrelation coefficient (Salas, et al., 1980) may be obtained recursively by the Durbins relations as given below through equation 4, 5 and 6:

$$\phi_1(1) = \rho_1, \phi_1(2) = \frac{\rho_1(1 - \rho_2)}{(1 - \rho_1^2)}, \phi_2(2) = \frac{\rho_2 - \rho_1^2}{(1 - \rho_1^2)} \quad (4)$$

$$\phi_k(k) = \frac{\rho_k - \sum_{j=1}^{k-1} \phi_j(k-1)\rho_{k-j}}{1 - \sum_{j=1}^{k-1} \phi_j(k-1)\rho_j} \quad (5)$$

$$\varnothing_j(k) = \varnothing_j(k-1) - \varnothing_k(k) \varnothing_{k-j}(k-1) \quad (6)$$

To determine the partial autocorrelation function from a sample series x_1, \dots, x_N , the sample autocorrelation r_k should be computed from Equation (3) and the ρ_s are replaced by rs . ρ_s are auto-regression coefficients.

The cross-correlation coefficient (Salas et al., 1980) is defined through equation 7 as:

$$r_k^{ij} = \frac{\sum_{t=1}^{N-k} (x_t^{(i)} - \bar{x}_t^{(i)}) (x_{t+k}^{(j)} - \bar{x}_{t+k}^{(j)})}{\left[\sum_{t=1}^{N-k} (x_t^{(i)} - \bar{x}_t^{(i)})^2 \sum_{t=1}^{N-k} (x_{t+k}^{(j)} - \bar{x}_{t+k}^{(j)})^2 \right]^{1/2}} \quad (7)$$

Where r_k^{ij} = lag-k cross-correlation coefficient; $x_t^{(i)}$ = time series values of series i ; $x_t^{(j)}$ = time series values of series j ; $\bar{x}_t^{(i)}$ = mean of the first N-k values of series i ; and $\bar{x}_{t+k}^{(j)}$ = mean of the last N-k values of series j .

2.2.2. Model Training

The ANN models have been trained by using back propagation (BP) algorithm. The best-known algorithm for training ANNs is the back propagation algorithm. It essentially searches for minima on the error surface by applying a steepest-descent gradient technique. The algorithm is linearly convergent. The back propagation architecture described here and in the accompanying appendices is the basic, classical version, but many variants of this basic form exist. Basically, each input pattern of the training data set is passed through a feedforward network from the input units to the output layer. The network output is compared with the desired target output, and an error is computed based on an error function. This error is propagated backward through the network to each neuron, and correspondingly the connection weights are adjusted.

Backpropagation is a first-order method based on the steepest gradient descent, with the direction vector being set equal to the negative of the gradient vector. Consequently, the solution often follows a zigzag path while trying to reach a minimum error position, which may slow down the training process. It is also possible for the training process to be trapped in a local minimum.

A possible way of preventing overtraining is called regularization. This method involves modifying the error function of performance learning algorithms. For example, if the MSE is used as error function, generalization can be improved by adding a term that consists of the mean of the sum of squares of the network weights and biases:

$$MSEREG = \gamma.MSE + (1 - \gamma).MSW \quad (8)$$

Where

$$MSW = \frac{1}{n} \sum_{j=1}^n w_j^2 \quad (9)$$

Using this performance function will cause the network to have smaller weights and biases, and this will force the network response to be smoother and less likely to overtrain. Bayesian regularization algorithm is used in this study in order to train the given network more efficiently.

The whole data has been divided into two data sets for the training and validation of the models. The daily data from 1, January 1988 to 25, November 1999 (i.e. 70% of the total available data) were considered for the calibration of the model because it contained the extreme values of discharge. The daily data from 26, November 1999 to 31, December 2004 (i.e. 30% of the total available data) were considered for the validation of the model. The ANN models were trained by using the functions of MATLAB (The Mathworks, 2001).

2.2.3. Performance Evaluation

The results obtained from calibration and validation is evaluated in order to determine the difference between observed and predicted values. In most of studies in this thesis, root mean square error (RMSE), model efficiency (EFF) (Nash and Sutcliffe, 1970), and coefficient of correlation (CORR) were used as performance criteria to evaluate the various ANN models. If CORR and EFF value for a model is found to be higher among all other models and RMSE is also lower among all other models, then this model is considered to give the best performance than the other model. These models can be defined by equation 10, 11 and 12 respectively:

$$\text{Root Mean Squared Error (RMSE)} = \sqrt{\frac{\sum_{k=1}^K (t - y)^2}{K}} \quad (10)$$

$$\text{Efficiency (EFF)} = 1 - \frac{\sum (t - y)^2}{\sum (t - \bar{t})^2} \quad (11)$$

$$\text{Coefficient of Correlation (CORR)} = \frac{\sum TY}{\sqrt{\sum T^2 \sum Y^2}} \quad (12)$$

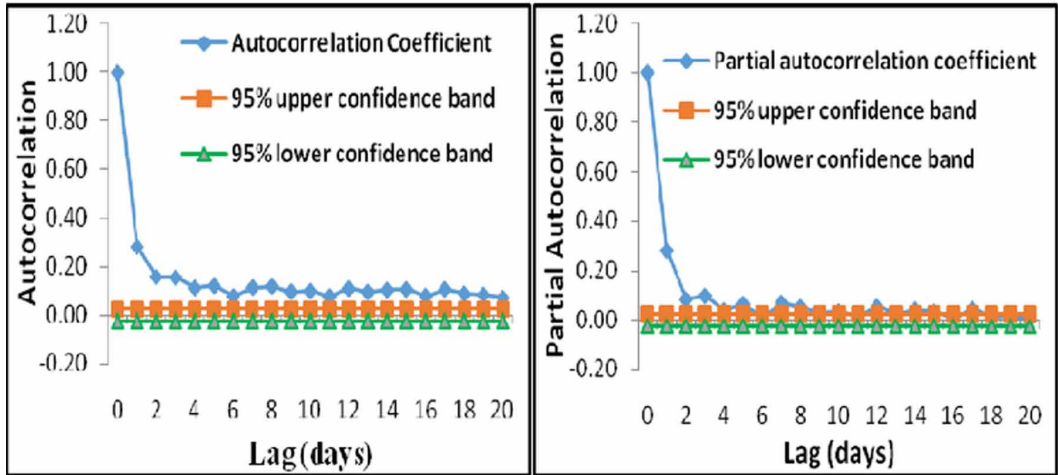
3. RESULTS AND DISCUSSIONS

3.1. Input Vector Selection

The auto correlation coefficient (ACF) and partial autocorrelation coefficient (PACF) of discharge at Bhakra are presented through Figure 2 (a-b) respectively. The cross-correlation coefficient (CCF) between discharge at Bhakra and rainfall at Bhakra, Berthin, Kahu, Kasol, Namgia, Raksham, Rampur and Suni are presented in Figures from 3 (a-h) respectively. The CCF between discharge and evaporation at Bhakra is presented in Figure 4. The CCF between discharge at Bhakra and discharge at Kasol, Rampur and Suni are presented in Figures 5(a-c), respectively.

The partial autocorrelation coefficient of discharge at Bhakra for lag 0 was 1. The partial autocorrelation coefficient values for other lags were less than 0.3. Therefore, we have considered 0 lag for partial autocorrelation coefficient of discharge at Bhakra. The cross-correlation coefficients of discharge at Bhakra with evaporation at Bhakra for lag 0 and 1 were 0.3 and 0.29, respectively.

Figure 2. (a) Autocorrelation of discharge at Bhakra; (b) Partial autocorrelation of discharge at Bhakra.



Therefore, we have considered 0 lag while cross-correlating discharge at Bhakra with evaporation at Bhakra. The cross-correlation coefficients of discharge at Bhakra with rainfall at Bhakra, Berthin, Kahu, Kasol, Namgia, Raksham, Rampur and Suni for lag 0 were 0.4, 0.4, 0.4, 0.41, 0.09, 0.29, 0.26, 0.32, respectively. However, the cross-correlation coefficients of discharge at Bhakra with discharge at Kasol, Rampur and Suni for lag 0 were 0.95, 0.94, and 0.96, respectively. These values are found to be more than all other lagged cross-correlation coefficient values.

On the basis of the values of PACF and CCF of the data series as per values mentioned, the following input vector was selected for neural network training as shown in Equation 13:

$$\text{bhadis}(t) = f(\text{berrain}(t), \text{bharain}(t), \text{kahrain}(t), \text{kasrain}(t), \text{namrain}(t), \text{rakrain}(t), \text{ramrain}(t), \text{sunrain}(t), \text{bhaevap}(t), \text{kasdis}(t), \text{ramdis}(t), \text{sundis}(t), \text{bhadis}(t-1)) \quad (13)$$

In equation 13, bhadis, kasdis, ramdis, sundis are discharge (m^3/s) at Bhakra, Kasol, Rampur and Suni respectively. Similarly, bharain, berrain, kahrain, kasrain, namrain, rakrain, ramrain, sunrain are rainfall values (mm) at Berthin, Bhakra, Kahu, Kasol, Namgia, Raksham, Rampur, Suni, respectively and Bhaevap is evaporation (mm) at Bhakra.

3.2. Model Performance

The performance of the MLP, RBF and FUZZY models during calibration and validation with the input combination derived from statistical procedure as suggested by Sudheer, et al. (2002) and is shown in Tables 2, 3 and 4, respectively.

3.2.1. Artificial Neural Network

The number of the neurons in the hidden layer is established by trial and error procedure starting with one hidden neuron initially and then increasing it up to 10 hidden neurons, based on the performance criteria of the model. The transfer function for the hidden and output layer are log sigmoid and pure linear respectively in the training of the models.

Table 2 indicates that ANN8 performed better than other models during calibration (CORR = 0.99, RMSE = 0.54, EFF = 98.56%) and validation (CORR = 0.98, RMSE = 0.56, EFF = 97.69%) because the coefficients of correlation and Efficiency (%) of ANN8 is found to be higher among all

Figure 3. (a) Cross-correlation of discharge at Bhakra with rainfall at Bhakra (b) Cross-correlation of discharge at Bhakra with rainfall at Berthin (c) Cross-correlation of discharge at Bhakra with rainfall at Kasol (d) Cross-correlation of discharge at Bhakra with rainfall at Kahu (e) Cross-correlation of discharge at Bhakra with rainfall at Namgia (f) Cross-correlation of discharge at Bhakra with rainfall at Raksham (g) Cross-correlation of discharge at Bhakra with rainfall at Rampur (h) Cross-correlation of discharge at Bhakra with rainfall at Suni.

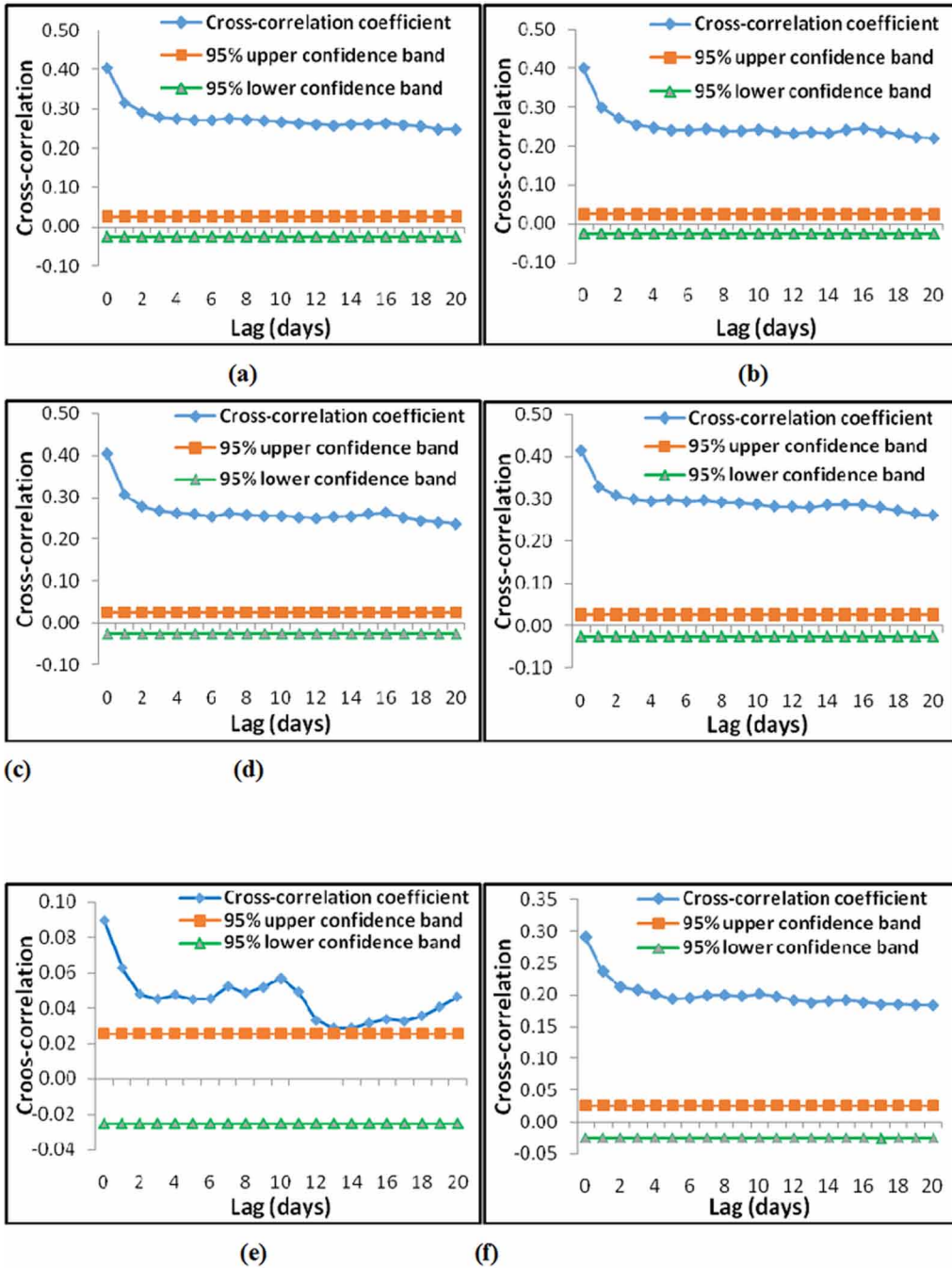
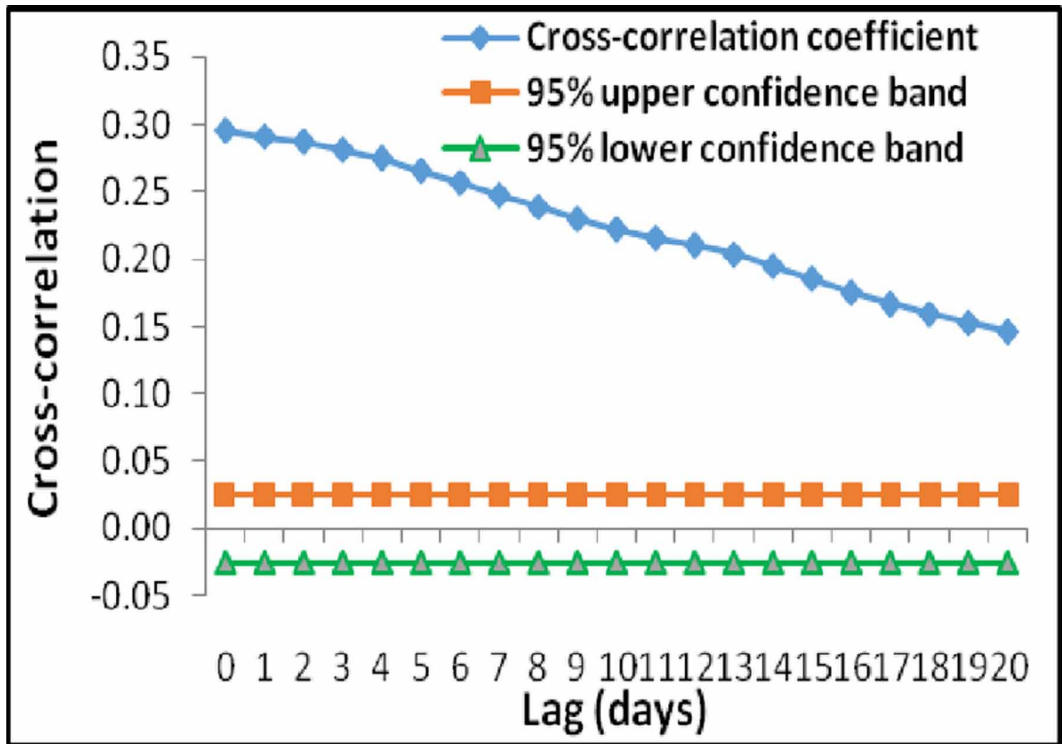


Figure 4. Cross-correlation between evaporation and discharge at Bhakra



other ANN models. Also from table 2 it can be inferred that the RMSE of ANN8 is lower among all other ANN models. The optimum structure of the ANN model has been found to be 8 neurons in the hidden layer.

3.2.2. Radial Basis Function Network

Number of hidden neurons in the hidden layer was investigated by trial and error method. The values investigated are 10, 20, 30, 40, 50, 60, 70, 80, 90 and 100. A smoothing parameter known as spread, has been needed to be set. The spread values as tested are 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9 and 1.

The data summarized in Table 3 indicates that the model RBF9 performed better than other models during calibration (CORR = 0.97, RMSE = 1.11, EFF = 93.61%) and validation (CORR = 0.97, RMSE = 0.87, EFF = 97.24%) because the coefficients of correlation and Efficiency (%) of RBF9 is higher among all other RBF models. Also, RMSE of RBF9 is lower among all other RBF models. The optimum structure of the RBF model was found to be 90 neurons in the hidden layer.

3.2.3. Fuzzy Logic

The Radius in fuzzy logic has been found by trial and error procedure starting with 0.1 and then increasing it up to 1, based on the performance criteria of the model.

The data summarized in table 4 by using fuzzy model indicates that the model FUZZY8 performed better than other models during calibration (CORR = 0.98, RMSE = 0.87, EFF = 96.71%) and validation (CORR = 0.99, RMSE = 0.57, EFF = 97.58%). This is because the coefficients of correlation and Efficiency (%) of FUZZY8 is higher among all other FUZZY models and RMSE of FUZZY8 is lower among all other FUZZY models. The optimum structure of the FUZZY model was found to be 90 neurons in the hidden layer.

Figure 5. (a) Cross-correlation of discharge at Bhakra with discharge at Kasol. (b) Cross-correlation of discharge at Bhakra with discharge at Rampur. (c) Cross-correlation of discharge at Bhakra with discharge at Suni.

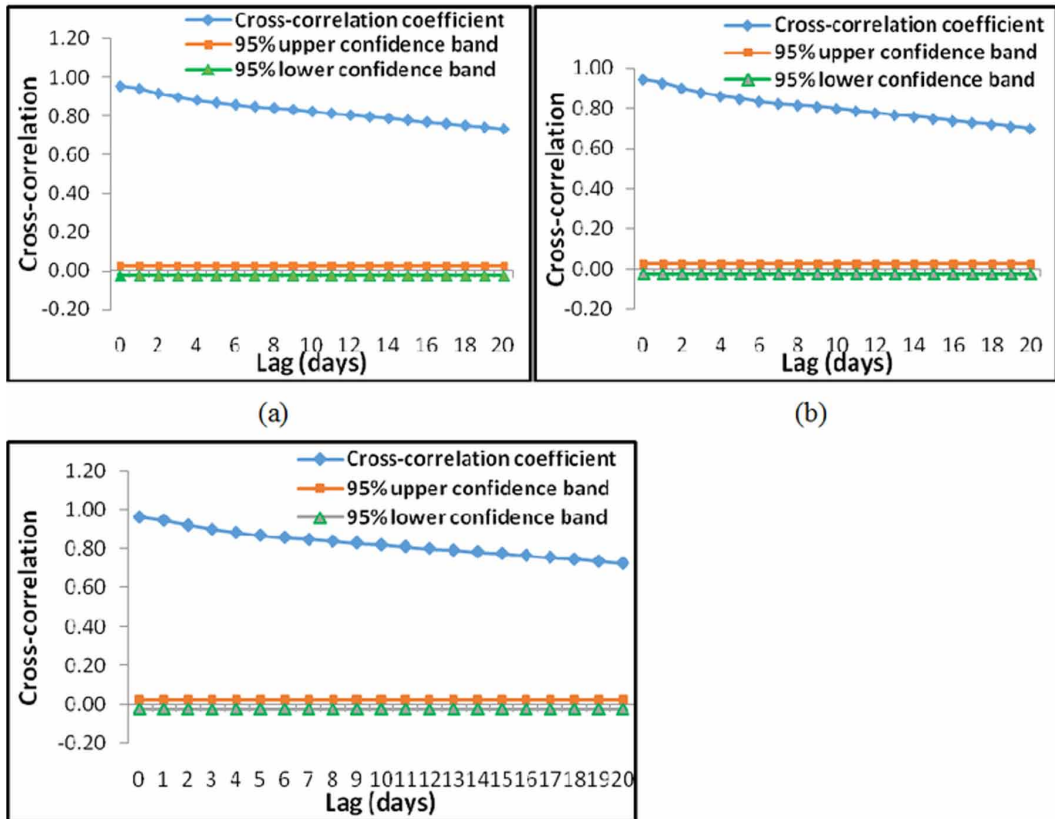


Table 2. Results of ANN model during calibration and validation

Model No.	Input Combinations	ANN Structure	Calibration			Validation		
			CORR	RMSE	EFF (%)	CORR	RMSE	EFF (%)
ANN1	bhaR(t), berR(t), kahR(t), kasR(t), namR(t), rakR(t), ramR(t), sunR(t), bhaE(t), kasD(t), ramD(t), sunD(t), bhaD(t-1)	13-1-1	0.98	0.90	96.47	0.99	0.61	97.23
ANN2		13-2-1	0.98	0.83	96.98	0.99	0.60	97.26
ANN3		13-3-1	0.99	0.74	97.58	0.98	0.63	96.96
ANN4		13-4-1	0.99	0.70	97.85	0.97	0.94	96.76
ANN5		13-5-1	0.99	0.69	97.91	0.98	0.67	96.63
ANN6		13-6-1	0.99	0.64	98.23	0.96	1	92.50
ANN7		13-7-1	0.99	0.58	98.53	0.97	0.88	94.20
ANN8		13-8-1	0.99	0.54	98.56	0.98	0.56	97.69
ANN9		13-9-1	0.99	0.54	98.73	0.97	0.96	93.14
ANN10		13-10-1	0.99	0.57	98.69	0.96	1.04	96.04

Note: bhaR = rainfall at Bhakra; berR = rainfall at Berthin; kah:rainfall at Kahu; kasR = rainfall at Kasol; namR = rainfall at Namgia; rakR = rainfall at raksham; ramR = rainfall at rampur; sunR = rainfall at Suni; bhaE = evaporation at Bhakra; kasD = discharge at Kasol; ramD = discharge at Rampur; sunD = discharge at Suni; bhaD = discharge at Bhakra.

Table 3. Results of RBF model during calibration and validation

Model Number	Input Calibration	ANN Structure	Goal	Spread	Calibration			Validation		
					CORR	RMSE	EFF (%)	CORR	RMSE	EFF (%)
RBF1	bhaR(t), berR(t), kahR(t), kasR(t), namR(t), rakR(t), ramR(t), sunR(t), bhaE(t), kasD(t), ramD(t), sunD(t), bhaD(t-1)	13-10-1	0.1	0.1	0.41	4.40	16.54	0.44	3.41	13.48
RBF2		13-20-1	0.1	0.2	0.71	3.40	50.21	0.76	2.48	54.24
RBF3		13-30-1	0.1	0.3	0.81	3.01	61.04	0.87	1.89	73.20
RBF4		13-40-1	0.1	0.4	0.89	2.16	79.76	0.93	1.40	85.22
RBF5		13-50-1	0.1	0.5	0.92	1.88	84.77	0.95	1.23	88.69
RBF6		13-60-1	0.1	0.6	0.94	1.63	88.56	0.96	1.06	91.56
RBF7		13-70-1	0.1	0.7	0.95	1.47	90.69	0.97	0.92	93.60
RBF8		13-80-1	0.1	0.8	0.97	1.20	93.50	0.98	0.79	95.28
RBF9		13-90-1	0.1	0.9	0.97	1.11	93.61	0.97	0.87	97.24
RBF10		13-100-1	0.1	1	0.97	1.11	93.60	0.98	0.73	96.03

Note: bhaR = rainfall at Bhakra; berR = rainfall at Berthin; kah: rainfall at Kahu; kasR = rainfall at Kasol; namR = rainfall at Namgia; rakR = rainfall at raksham; ramR = rainfall at rampur; sunR = rainfall at Suni; bhaE = evaporation at Bhakra; kasD = discharge at Kasol; ramD = discharge at Rampur; sunD = discharge at Suni; bhaD = discharge at Bhakra.

Table 4. Results of fuzzy logic model during calibration and validation

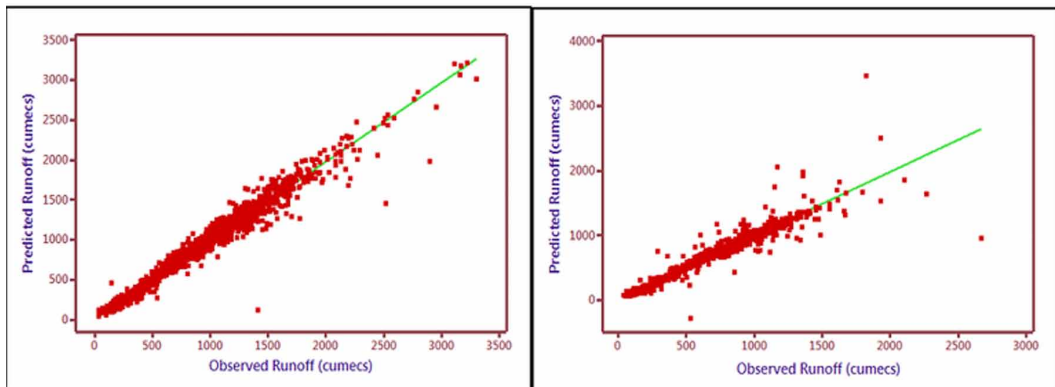
Model No.	Input Combinations	Radius	Calibration			Validation		
			CORR	RMSE	EFF (%)	CORR	RMSE	EFF (%)
FUZZY1	bhaR(t), berR(t), kahR(t), kasR(t), namR(t), rakR(t), ramR(t), sunR(t), bhaE(t), kasD(t), ramD(t), sunD(t), bhaD(t-1)	0.1	0.27	6.38	75.57	0.21	4.93	81.54
FUZZY2		0.2	0.98	0.88	96.63	0.99	0.62	97.10
FUZZY3		0.3	0.98	0.91	96.45	0.99	0.61	97.19
FUZZY4		0.4	0.98	0.91	96.45	0.99	0.61	97.19
FUZZY5		0.5	0.98	0.97	96.67	0.99	0.58	97.46
FUZZY6		0.6	0.98	0.87	96.69	0.99	0.57	97.52
FUZZY7		0.7	0.98	0.87	96.68	0.99	0.57	97.53
FUZZY8		0.8	0.98	0.87	96.71	0.99	0.57	97.58
FUZZY9		0.9	0.98	0.87	96.70	0.99	0.57	97.57
FUZZY10		1.0	0.98	0.90	96.45	0.99	0.61	97.19

Note: bhaR = rainfall at Bhakra; berR = rainfall at Berthin; kah: rainfall at Kahu; kasR = rainfall at Kasol; namR = rainfall at Namgia; rakR = rainfall at raksham; ramR = rainfall at rampur; sunR = rainfall at Suni; bhaE = evaporation at Bhakra; kasD = discharge at Kasol; ramD = discharge at Rampur; sunD = discharge at Suni; bhaD = discharge at Bhakra.

3.3. Analysis of Results of ANN, RBF and Fuzzy Models

The performance of the best ANN, RBF and FUZZY models for the prediction of runoff at Bhakra during calibration and validation is depicted through Figure 6 (a-b) to 8(a-b) respectively. The scatter

Figure 6. (a-b). Scatter plot for the result of best ANN model during (a) Calibration (b) validation.



plots clearly demonstrate the potentiality of the developed ANN, RBF and FUZZY models in the prediction of runoff.

During calibration, the coefficient of correlation for the ANN model (ANN8) is found to be 0.99, whereas, the value of the coefficient of correlation for the RBF model (RBF9) is found to be 0.97. However, for the FUZZY model (FUZZY8) the value comes out to be 0.98. RMSE for ANN model (ANN8) is found to be 0.54 and for RBF model (RBF9) the value comes out to be 1.11. Whereas, for the FUZZY model (FUZZY8) the RMSE value is observed to be 0.87. Therefore, from the present evaluation it can be inferred that in terms of RMSE value, the ANN model performed best (0.54) and RBF model (RBF9) performed worst (1.11) than all other models investigated in this study. However, in terms of efficiency, the ANN model (ANN8) performed the best (98.56%), whereas RBF model (RBF9) performed the worst (93.61%).

Figure 7. (a-b). Scatter plot for the result of best RBF model during (a) calibration (b) validation.

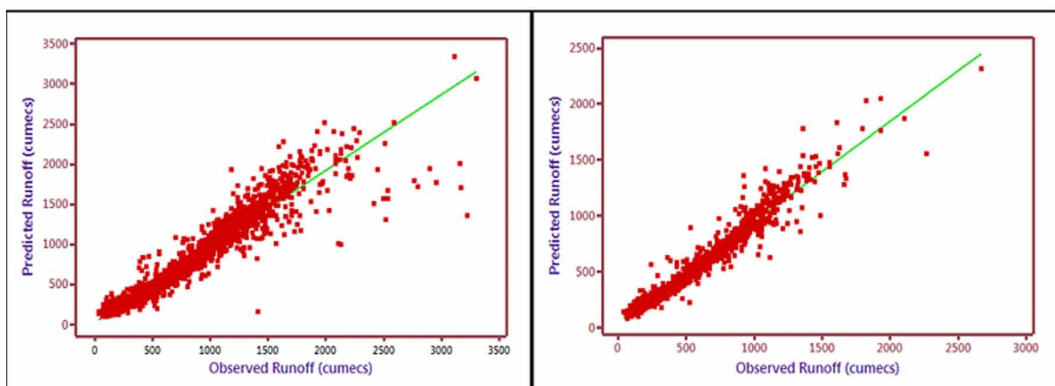
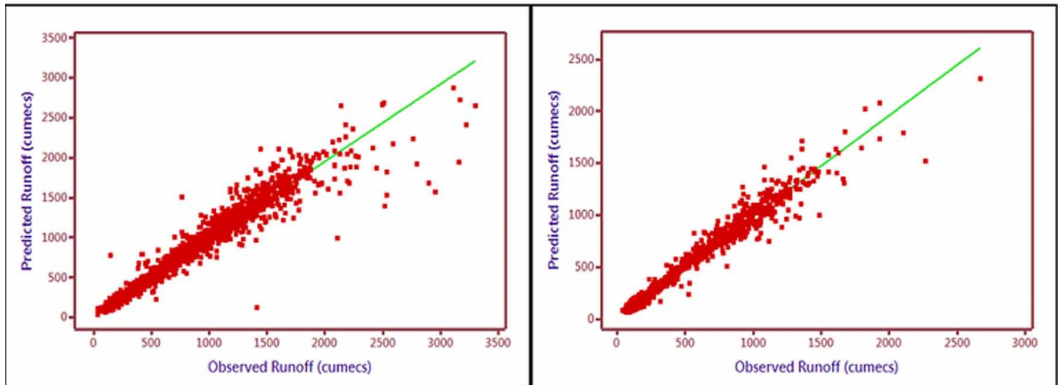


Figure 8. (a-b). Scatter plot for the result of best FUZZY model during (a) calibration (b) validation.



During Validation, the coefficient of correlation for the ANN model (ANN8) is observed to be 0.98. The value of the coefficient of correlation for the RBF model (RBF9) has been found to be 0.97. However, for the FUZZY model (FUZZY8) the value comes out to be 0.99. RMSE for ANN model (ANN8) is observed to be 0.56 and for RBF model (RBF9) the value is 0.87. In the FUZZY model (FUZZY8) the RMSE value is observed to be 0.57. Therefore, from the present evaluation it can be inferred that in terms of RMSE value, the ANN model performed best (0.56) and RBF model (RBF9) performed worst (0.87) than all other models used in this study. However, in terms of efficiency, the ANN model (ANN8) performed the best (97.69%), whereas RBF model (RBF9) performed the best (97.24%).

The results of the calibration and validation of the best ANN, RBF and FUZZY models in terms of various statistical indices are presented in the Table 5. The performance of the best ANN, RBF and FUZZY models in terms of observed and predicted runoff at Bhakra during calibration and validation are presented in Figures 9 and 10 respectively.

From performance indexes as shown in Figures 9 and 10, it has been observed that the ANN model learned the process better than any other model because the efficiency of ANN model (ANN8) is higher than the RBF model (RBF9). However, the FUZZY model and RMSE value is least for the ANN model (ANN8) than any other model during calibration and validation as shown in Table 5. Thus, it can be concluded from the overall performance of the models that the ANN model (ANN8) performed the best and the RBF (RBF9) model performed the worst during both calibration and validation. In general, the performance of all the models used in the modeling process predicted runoff with adequate accuracy but particularly the ANN model performed best with the efficiency of with the efficiency of 98.56%.

Table 5. Comparison of results among the best ANN, RBF and FUZZY Logic models during calibration and validation

Model	Calibration			Validation		
	CORR	RMSE	EFF (%)	CORR	RMSE	EFF (%)
ANN8	0.99	0.54	98.56	0.98	0.56	97.69
RBF9	0.97	1.11	93.61	0.97	0.87	97.24
FUZZY8	0.98	0.87	96.71	0.99	0.57	97.58

Figure 9. Graph showing predicted runoff values by ANN, RBF and Fuzzy logic against observed runoff values during calibration at Bhakra

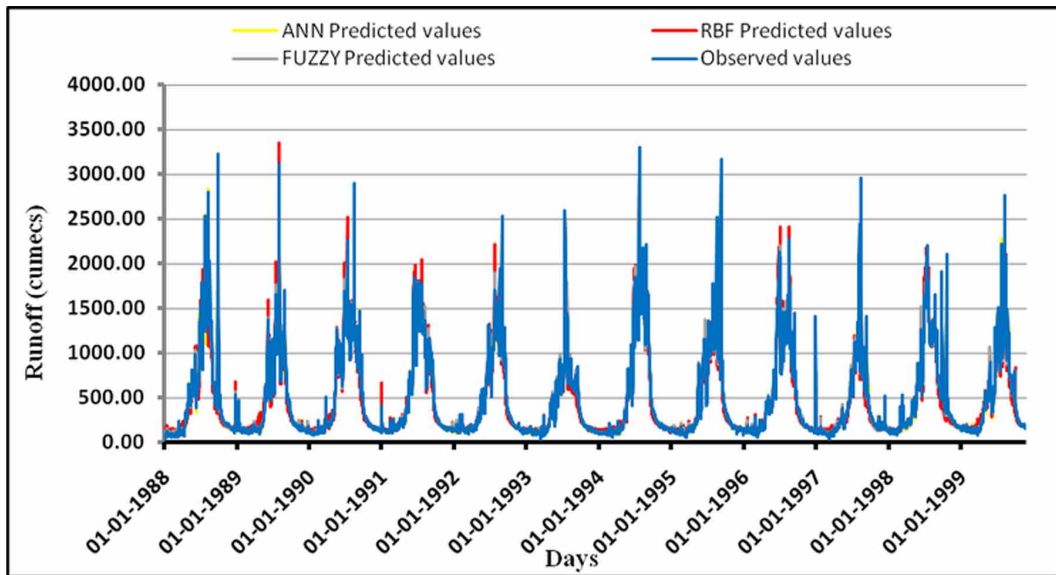
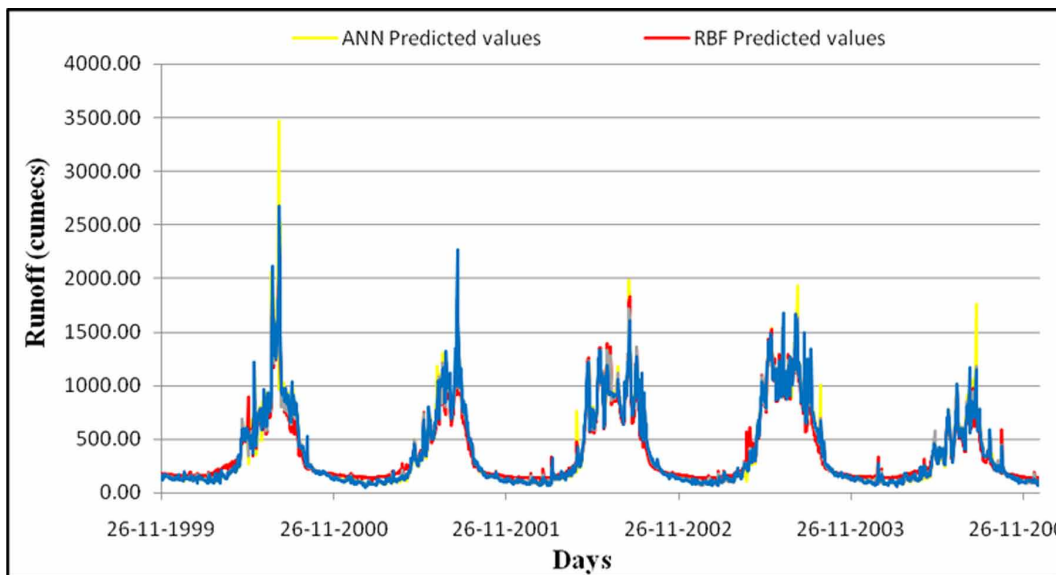


Figure 10. Graph showing predicted runoff values by ANN, RBF and Fuzzy logic against observed runoff values during validation at Bhakra



4. CONCLUSION

This paper presents the findings of a study aimed at a comparative analysis of various techniques available for the purpose of modeling of runoff at the Bhakra gauging site of the Sutlej river basin in northern India. The techniques investigated include Artificial Neural Network (ANN) with back

propagation algorithm, Radial Basis Function and Fuzzy Logic. The statistical parameters ACF, PACF and CCF were used for selection of the input vector to the various models. Daily observed rainfall, evaporation and runoff with 1-day lag has been used as the input. Out of 16 years datasets, 70% of data is used for calibration and 30% is used for validation. The statistical indices such as coefficient of correlation, root mean squared error (RMSE) and model efficiency have been used to evaluate the performance of the model. In general, the performance of all the models used in the modeling process predicted runoff with adequate accuracy but particularly the ANN model performed best with the efficiency of with the efficiency of 98.56%.

REFERENCES

- Ahmad, S., & Simonovic, S. P. (2005). An Artificial Neural Network Model for Generating Hydrograph From Hydro-Meteorological Parameters. *Journal of Hydrology (Amsterdam)*, 315(1), 236–251. doi:10.1016/j.jhydrol.2005.03.032
- Alvisi, S., Mascellani, G., Franchini, M., & Bárdossy, A. (2006). Water level forecasting through fuzzy logic and artificial neural network approaches. *Hydrology and Earth System Sciences*, 10(1), 1–17. doi:10.5194/hess-10-1-2006
- ASCE Task Committee on Application of Artificial Neural Networks in Hydrology. (2000a). Artificial neural networks in hydrology-I: Preliminary concepts. *Journal of Hydrologic Engineering*, 5(2), 115–123. doi:10.1061/(ASCE)1084-0699(2000)5:2(115)
- ASCE Task Committee on Application of Artificial Neural Networks in Hydrology. (2000b). Artificial neural networks in hydrology-II: Hydrologic applications. *Journal of Hydrologic Engineering*, 5(2), 124–137. doi:10.1061/(ASCE)1084-0699(2000)5:2(124)
- Bárdossy, A., Mascellani, G., & Franchini, M. (2006). Fuzzy unit hydrograph. *Water Resources Research*, 42(2), 3–10. doi:10.1029/2004WR003751
- Campolo, M., Soldati, A., & Andreussi, P. (2003). Artificial neural network approach to flood forecasting in the River Arno. *Hydrological Sciences Journal*, 48(3), 381–398. doi:10.1623/hysj.48.3.381.45286
- Chiu, S. L. (1994). Fuzzy model identification based on cluster estimation. *Journal of Intelligent & Fuzzy Systems*, 2, 267–278.
- Cobaner, M., Unal, B., & Kisi, O. (2009). Suspended sediment concentration estimation by an adaptive neuro-fuzzy and neural network approaches using hydro-meteorological data. *Journal of Hydrology (Amsterdam)*, 367(1-2), 52–61. doi:10.1016/j.jhydrol.2008.12.024
- Dawson, C. W., & Wilby, R. L. (2001). Hydrological modelling using artificial neural networks. *Progress in Physical Geography*, 25(1), 80–108. doi:10.1177/030913330102500104
- De Vos, N. J., & Rientjes, T. H. M. (2005). Constraints of artificial neural networks for rainfall-runoff modelling: Trade-offs in hydrological state representation and model evaluation. *Hydrology and Earth System Sciences*, 9(1/2), 111–126. doi:10.5194/hess-9-111-2005
- Deka, P., & Chandramouli, V. (2005). A fuzzy neural network model for deriving the river stage—discharge relationship. *Hydrological Sciences Journal*, 48(2), 197–209. doi:10.1623/hysj.48.2.197.44697
- Firat, M., & Güngör, M. (2007). River flow estimation using adaptive neuro-fuzzy inference system. *Mathematics and Computers in Simulation*, 75(3-4), 87–96. doi:10.1016/j.matcom.2006.09.003
- French, M. N., Krajewski, W. F., & Cuykendall, R. R. (1992). Rainfall forecasting in space and time using neural network. *Journal of Hydrology (Amsterdam)*, 137(1-4), 1–31. doi:10.1016/0022-1694(92)90046-X
- Gowda, C. C., & Moyya, S. D. (2014). Runoff Modeling using different membership functions in Adaptive Neuro fuzzy inference system. *International Journal of Advances in Engineering Sciences*, 4(4), 48–51.
- Hsu, K., Gupta, H. V., & Sorooshian, S. (1995). Artificial neural network modeling of the rainfall-runoff process. *Water Resources Research*, 31(10), 2517–2530. doi:10.1029/95WR01955
- Jacquin, A.P. and Shamseldin, A.Y. (2006). Development of rainfall-runoff models using Takagi-Sugeno fuzzy inference systems. *Journal of Hydrology*, 329, 154–173.
- Jain, A., & Srinivasulu, S. (2004). Development of effective and efficient rainfall-runoff models using integration of deterministic, real-coded genetic algorithms and artificial neural network techniques. *Water Resources Research*, 40(4), 1–12. doi:10.1029/2003WR002355
- Kisi, Ö. (2007). Streamflow forecasting using different artificial neural network algorithms. *Journal of Hydrologic Engineering*, 12(5), 532–539. doi:10.1061/(ASCE)1084-0699(2007)12:5(532)
- Kumar, D. N., Raju, K. S., & Sathish, T. (2004). River flow forecasting using recurrent neural networks. *Water Resources Management*, 18(2), 143–161. doi:10.1023/B:WARM.0000024727.94701.12

- Lohani, A. K., Goel, N. K., & Bhatia, K. K. S. (2006). Takagi–Sugeno fuzzy inference system for modeling stage–discharge relationship. *Journal of Hydrology*, 331(1-2), 146–160. doi:10.1016/j.jhydrol.2006.05.007
- Mukarji, A., Chatterjee, C., & Raghuwanshi, N. S. (2009). Flood Forecasting Using ANN, Neuro-Fuzzy, and Neuro-GA Models. *Journal of Hydrologic Engineering*, 14(6), 647–652. doi:10.1061/(ASCE)HE.1943-5584.0000040
- Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models: 1. A discussion of principles. *Journal of Hydrology (Amsterdam)*, 10(3), 282–290. doi:10.1016/0022-1694(70)90255-6
- Rajurkar, M. P., Kothiyari, U. C., & Chaube, U. C. (2002). Artificial neural networks for daily rainfall–runoff model. *Hydrological Sciences Journal*, 47(6), 865–877. doi:10.1080/02626660209492996
- Rajurkar, M. P., Kothiyari, U. C., & Chaube, U. C. (2004). Modeling of the daily rainfall–runoff relationship with artificial neural network. *Journal of Hydrology (Amsterdam)*, 285(1-4), 96–113. doi:10.1016/j.jhydrol.2003.08.011
- Rumelhart, D. E., Hinton, G. E., & McClelland, J. L. (1986). *A general framework for parallel distributed processing*. In *Parallel Distributed Processing: explorations in the microstructure of cognition* (pp. 45–76). Cambridge, UK: MIT Press.
- Salas, J. D., Delleur, J. W., Yevjevich, V., & Lane, W. L. (1980). *Applied Modeling of Hydrological Time Series*. Denver, CO: Water Resources Publications.
- Singh, P. & Kumar, N. (1997b). Effect of orography on precipitation in the western Himalayan region. *Journal of Hydrology*, 199, 183-206.
- Smith, J., & Sli, R. N. (1995). Neural network model for rainfall runoff process. *Journal of Water Resources Planning and Management*, 121(6), 49–508. doi:10.1061/(ASCE)0733-9496(1995)121:6(499)
- Solaimani, K. (2009). A Study of Rainfall Forecasting Models Based on Artificial Neural Network. *Asian Journal of Applied Sciences*, 2(6), 486–498. doi:10.3923/ajaps.2009.486.498
- Srinivasulu, S., & Jain, A. (2006). A comparative analysis of training methods for artificial neural network rainfall–runoff models. *Applied Soft Computing*, 6(3), 295–306. doi:10.1016/j.asoc.2005.02.002
- Sudheer, K. P., Gosain, A. K., & Ramasastri, K. S. (2002). A data-driven algorithm for constructing artificial neural network rainfall-runoff models. *Hydrological Processes*, 16(6), 1325–1330. doi:10.1002/hyp.554
- Suhaimi, S. and Bustami, R.A. (2009). Rainfall Runoff Modeling using Radial Basis Function Neural Network for Sungai Tinjar Catchment, Miri, Sarawak. *UNIMAS E-Journal of Civil Engineering*, 1(1).
- Takagi, T., & Sugeno, M. (1985). Fuzzy identification of systems and its applications to modeling and control. *IEEE Transactions on Systems, Man, and Cybernetics Systems*, 15(1), 116–131.
- Tayfur, G., & Singh, V. P. (2006). ANN and fuzzy logic models for simulating event-based rainfall–runoff. *Journal of Hydrologic Engineering*, 132(12), 1321–1330. doi:10.1061/(ASCE)0733-9429(2006)132:12(1321)
- Vernieuwe, H., Georgieva, O., Baets, B. D., Pauwels, V. R. N., Verhoest, N. E. C., & Troch, D. F. P. (2005). Comparison of data-driven Takagi–Sugeno models of rainfall–discharge dynamics. *Journal of Hydrology (Amsterdam)*, 302(1-4), 173–186. doi:10.1016/j.jhydrol.2004.07.001
- Vivekanandan, N. (2014). Prediction of Rainfall Using MLP and RBF Networks. *International Journal of Advanced Networking and Applications*, 5(4), 1974–1979.

Athar Hussain is working as an Associate Professor in Department of Civil Engineering and Head, Environmental Engineering Department at Ch. B. P. Govt. College of Engineering, New Delhi. Dr. Hussain completed his Ph.D. in Environmental Engineering from IIT Roorkee in year 2007. During the research period Dr. Hussain worked as SRF in CSIR and MoEF sponsored projects until December 2006. Thereafter, Dr. Hussain started his career as a lecturer in civil engineering department at Dr. B.R. Ambedkar National Institute of Technology, Jalandhar until May 2007. Thereafter, he joined as an Assistant Director with Environment Management Group at National Productivity Council, New Delhi an autonomous body under Ministry of Commerce & Industry and worked there until October 2010. During the tenure he was involved in various national level and policy level projects sponsored from reputed government agencies and organizations such as MoEF, CPCB, SPCBs, etc., in the field of environmental engineering. Thereafter he worked as an Assistant Professor, civil engineering department in School of Engineering at Gautam Buddha University until December 2015. Thereafter he joined as Associate Professor in Department of Civil Engineering at Ch. B. P. Govt. College of Engineering, New Delhi. His research area includes industrial wastewater treatment and solid waste management with special focus on anaerobic treatment of wastewater and solid waste. He is supervising many Ph.D. Research Scholars in diversified areas of environmental engineering. Dr. Hussain has more than 15 years of experience in the area of teaching, research and consultancy. He has also guided more than 60 M.Tech dissertations. He has published more than 60 papers in international referred journals and international conferences. Dr. Hussain is also serving as a reviewer for many international journals like Taylor and Francis, Inderscience Publishers, Elsevier Science, etc. He is also an editorial board member of many journals some of which include Elsevier science and Inderscience publications. He is also an Associate Member of The Institution of Engineers (IE) India, Member of Institution of valuers" (IV), N. Delhi, India and Member of Indian Society of Wind Engineering, India. He has also handled various national and policy level projects with government and private organizations.

Jatin Kumar Singh is a M.Tech. student in Civil Engineering Department, Environmental Engineering Section, School of Engineering, Gautam Buddha University, Greater Noida, India.

A. R. Senthil Kumar is working as a Senior Scientist at National Institute of Hydrology, Roorkee, District Haridwar, Uttarakhand, India.