International Conference "Water, Environment, Energy and Society" (WEES–2009)
New Delhi, 12–16 January 2009

Artificial Neural Network Modelling of Tawi River Basin

Avinash Agarwal[1], Tanveer Ahmad and R.D. Singh[2]
National Institute of Hydrology
Jal Vigyan Bhawan, Roorkee, Uttarakhand - 247 667, INDIA
E-mail: [1]avinash@nih.ernet.in; [2]rdsingh@nih.ernet.in

ABSTRACT: The rainfall-runoff process is a highly non-linear complex process due the spatial and temporal variability of precipitation patterns and watershed characteristics. The understanding and modelling of rainfall-runoff transformation on watershed scale has attracted the hydrologists for water management, stream flow estimation, water supply, irrigation, drainage, flood control, water quality, power generation, recreation, and wild life protection and propagation. A number of modelling approaches have been developed in past to simulate the process accurately and efficiently.

In this study, a back propagation Artificial Neural Network (ANN) modelling approach has been formulated in FORTRAN language. Model uses the gradient descent optimization technique considering pattern learning process with different normalization techniques and applied for monthly rainfall-runoff modelling of the *Tawi* river catchments up to *Tawi* Bridge at Jammu. The model uses the monthly rainfall and runoff data from 1992 to 2002 which is pre-analyzed for general behavior of process on annual basis. Models were calibrated and validated considering whole data set in four different ways with different normalization techniques and by considering both three and four layers system with different numbers of nodes in hidden layers. The models were also evaluated considering four different statistical testing techniques. Three-layered feed forward network structure was better than a four layered structure. In all four combinations, adopted for the modeling, none was found effective uniformly in all calibration, cross-validation and verification periods and may be probably due to quality of the data.

INTRODUCTION

In system theoretic modeling, Artificial Neural Networks (ANNs) have become very popular for process simulation and forecasting in a number of areas including finance, power generation, medicine, water resources and environmental science. ANNs are able to generalize a relationship from small sample of data, are robust in the presence of noisy or missing inputs and can learn in response to changing environments. ANNs have been applied widely in various aspects of hydrology such as rainfall-runoff modeling, stream flow forecasting, ground water modeling, water quality, water management policy, precipitation forecasting, hydrological time series, and reservoir operations.

The Artificial Neural Network (ANN) models are best described by lumped deterministic black box model. The first successful artificial neural network was developed in 1940s. The literature of ANN with hydrologic application include modeling of the daily rainfall runoff process and snow rainfall process, assessment of stream ecological and hydrological responses to climate change, rainfall and runoff forecasting, sediment transport prediction, ground water quality and ground water remediation. A brief review of the rainfall runoff relationship over watershed using ANN can be seen in ASCE (2000a, b).

The current literature on applications of the ANN in hydrology water resources, especially on runoff includes Anmala (2000), Tokar (2000), Imrie (2000), Wilby (2003) and Rajurkar (2002). Some papers covering runoff forecasting are by Danh (1999), Elshorbagy (2000), Thirumalaiah (2000), Xu (2002), Birikundavyi (2002), Shivakumar (2002), Xiong (2002) and Cigizoglu (2003). Other applications of ANN include unit hydrograph derivation by Lange (1998) and geomorphology-based ANN for watershed by Zhang (2003).

A number of researchers have investigated the potential of ANN in modelling and forecasting of rainfall-runoff processes. Most of the studies normally describing a methodology for development of the ANN model adopting pattern learning process. In literature, the developed ANN-based rainfall-run-off models are compared with the conceptual and/or statistical models and the former are found to be superior to others in performance. Attempts have also been made to simulate runoff hydrographs using different input parameters, such as drainage basin, elevation, average slope, and average annual precipitation. A little attempt has however been made towards ANN extrapolation, representing internal behavior and distributed approach of ANN models that can model storm rainfall, runoff process. In this study, an attempt will be made to study the extrapolation

[1]Conference speaker

properties of ANN using different normalization techniques.

This research paper presents ANN software developed for hydrological applications with following as objectives:

1. Development computer programs for ANN models based on Back Propagation (BP) approach.
2. Application of ANN models using monthly rainfall-runoff on Tawi river basin of J&K.
3. Study of extrapolation properties of ANN through different normalization methods.

METHODOLOGY

Struggle for survival in living creature is supported by the neurobiological system of the body consisting of nerve cells. Numerous nerve cells process the incoming information and the signal is transmitted to nucleus/brain to take decision. The information processing together with several nerve cells is a new way of computing, called 'neurocomputing'.

Neurocomputing is accomplished by highly parallel structures designed to directly process the information coming from external world. To understand the basis behind neurocomputing, it is necessary to understand the neurobiological principles. The neurons are nerve cells and the neural network is network of these cells. Many in medical science believe that a cluster of parallel processing structure perhaps best represents a neural network.

The anatomy of a real biological neuron, comprises of *Dendrites, Soma, Axon* and *Synaptic Buttons*, is shown in Figure 1. The information is picked up at the *Dendrites*. The *Soma* is cell body whereas the *Axon* is long transmission line like structure and the tail end of the *Axon* is called *Synaptic Buttons*. These neurons are so powerful in processing the information, that even a small earthworm with only 302 neurons has a computing power around one thousand times the power of *Pentium II processor*. However, it is estimated that the human brain has 10^9 to 10^{12} neurons with 10^{15} *Synaptic Buttons* (Nav Bharat Times, April 28, 1999). Thus the computation power of parallel processing in neurobiological system is very high. An artificial neuron structure parallel to the real biological neuron is also shown in Figure 1.

In short, an Artificial Neural Network (ANN) is a computational structure inspired by neurobiological systems and learns from the past what has happened. More and better learning generalizes the path to solve complex problems arriving in nature. In other way, ANN is the best tool to represent natural processes. The tasks like survival, search, image and speech recognition, classification, generalization and so on can easily be handled by neurobiological systems while these are quite difficult by conventional digital computing. Mathematically, an ANN may be treated as a universal approximate of a neurobiological system (ASCE, 2000). The ANN learns to solve a problem by developing a memory capable of associating the input and corresponding output sets. Owing to the characteristics of linear and non-linear processes, the flexible ANN mathematical structure is capable of identifying complex non-linear relationships between the input and output data sets (Hsu, 1995; Raman and Sunil Kumar, 1995).

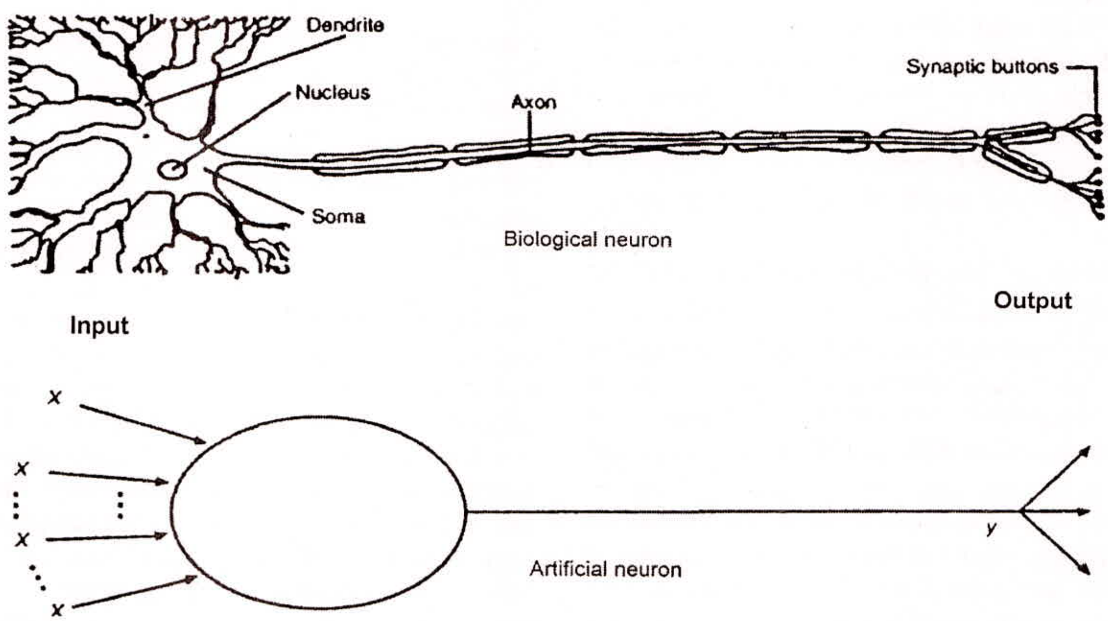

Fig. 1: Anatomy of biological and artificial neurons

ARTIFICIAL NEURAL NETWORK MODEL

The framework or structure of ANN as well as optimization scheme is the basic of ANN. The structure of an ANN consists of many single elements called nodes or neurons. The signal or information is passed between the nodes through inter connection or link and each connecting link has an associated weight that represents the strength of the connection. Each node applies an activation function to its net input to determine output signal.

The structure of a network refers to its framework as well as interconnection scheme. A structure may be either a fully connected network or a layered network. The framework of layered structure is specified by number of layers, such as, the input layer, the hidden layer/s and the output layer. According to inter-connection scheme, a network can be a feed forward connection, a feed backward connection, a lateral connection, a recurrent connection or a time delayed connection.

The multi-layer BPANN is layered parallel processing system consisting of input, output and hidden layer/s. There are many processing elements in each layer called nodes and these are connected by links of different weights. The number of nodes in input and output layers corresponds to the number of input and output variables of the model. There is no fixed rule but some guidelines are available in literature as to how many nodes should be in hidden layer/s. Optimum number of nodes in hidden layer depends on the complexity of modelling problem. An artificial neural network consisting of three layers as j, i, and k with number of nodes in the layer j as $j = 1$ to jj, in the layer i as $i = 1$ to ii and in layer k as $k = 1$ to kk along with the interconnecting weights W_{ij} and W_{ki} are shown in Figure 2.

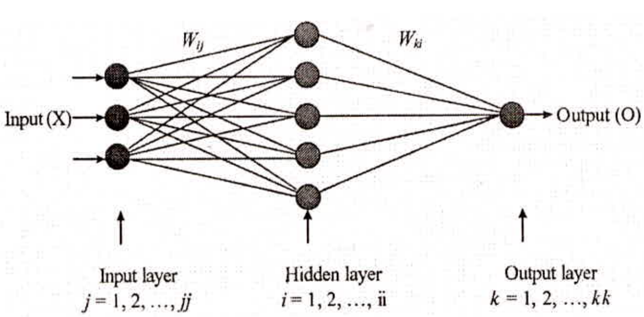

Fig. 2: Structure and notations in a multi-layer BPANN

Back Propagation Artificial Neural Network Model

The calculation in the back propagation scheme (Rumelhart *et al.*, 1986) is based on the generalized delta rule and consists of feed forward and the error back calculations. In feed forward calculation the nodes in the input layer j receive the data (input vector) and each neuron in layer i and k receives the weighted sum of output from the previous layer as input through an activation function.

Normalization of Data

Normalization of data is one of important aspect of ANN model development. The input and output data are normalized betwmen a suitable range, usually 0.0 to 1.0. There are very few studies which evaluates the effect of the normalization scheme on the ANN model performance. Normally, two different normalization methods are adopted in ANN model development and can be expressed using the following equation,

$$Xn_{(I)} = X_{(I)}/X_{\max} \qquad \text{... (1)}$$

$$Xn_{(I)} = (X_{(I)}-X_{\min})/(X_{\max}-X_{\min}) \qquad \text{... (2)}$$

where, Xn is the normalized values; X is the original value; $X_{\min}$ is the minimum value of the variable; $X_{\max}$ is the maximum value of the variable; and I is the index representing the number of data points.

First method converts the original data series into the range ($X_{\min}/X_{\max}$ and 1.0); second method converts the original date series into the range (0.0 and 1.0). The problem with these normalization methods is that they do not provide any flexibility of extrapolation beyond the calibration range. To provide extrapolation ability of ANN beyond the calibration range the thirst method suggested is of the following form as,

$$Xn_{(I)} = a + b \,.\, (X_{(I)} - X_{\min})/(X_{\max}-X_{\min}) \qquad \text{... (3)}$$

The value a and b could be selected such that a maximum flexibility could be given to develop model. Normally the value of a and b is taken as 0.1 and 0.8. This converts the original data series into the range 0.1 to 0.9 respectively and provides equal flexibility on both lower and higher side. Different values of a and b could be selected to provide unequal flexibility to lowered higher side of the model. The present developed FORTRAN code in this study has options to consider all these three methods one at a time.

Learning Process

The formulation of an appropriate artificial neural network model involves learning (training) relationship between input and output and several methods have been suggested in literature. The two alternate approaches for layered network are unsupervised and supervised learning (Figure 2). In unsupervised learning, the weights connecting nodes in the network are given a critic (good or bad), based on the error information supplied to the network to improve the weights. The competitive, reinforced and genetic algorithms are few learning processes that fall in this category. The best known example in reinforced learning is Kohonen network (Kohonen, 1988).

In supervised learning process, the weights connecting nodes in the network are set on the basis of detailed error information supplied to the network by an external 'teacher' as shown in Figure 3. In this process, the network is trained using a set of known input-output pairs to the network and the error between estimated and observed output is used to adjust the connecting weights. The supervised learning process is carried out either by pattern learning or batch learning processes:

(a) *Pattern learning process:* The pattern learning is more commonly used supervised learning process (French *et al.*, 1992) which is governed by the error of each data set taken one by one. The weights continuously get adjusted with the processing of each data set. The processing is slow as it continuously eliminates the biasness, if any, of one data set by the next few data sets coming in the sequence.

(b) *Batch learning process:* In batch learning, the process of learning is governed by the average error of all data sets. Weights could be adjusted in accordance with the set indicating highest error or lowest error or average error. The learning speed in batch learning process varies depending upon the type of error being selected for weight adjustment. In general, the process requires more computational time for each updating of weights. When weights are adjusted in accordance with the highest error, the learning is relatively fast in comparison to the pattern learning process (French *et al.*, 1992).

Network Generalization

The criteria to stop the convergence or the learning of ANN are normally linked with the generalization of the network. The generalization of the network is dependent on the network structure and the size, the learning algorithm or the method of convergence, quality and quantity of training and validation data domain (Fu, 1996). The generalization of the network means how well the learned ANN model performs on the data used for model development, and on the new data set other than the data set used for training the network. The performance of the model is verified through statistical evaluation criteria. Some of the other methods suggested in the literature to improve the generalization are network pruning, weight decay and elimination and weight shearing. Criteria linked with the generalization and used to stop the convergence are as follows:

1. Based on minimum error
2. Based on the minimum error gradient
3. Based on cross validation performance.

The first two methods listed above may lead to over or under, learning and thereby poor generalization. The cross-validation criterion normally avoids over-learning/under-learning and actually improves the generalization performance of the network and has been used in this study. A pre-determined level of accuracy is not required in this method; however, it requires intensive computations and demands much more data and computer space.

Steps in Development of ANN Model

Steps followed in the development of artificial neural network model are summarized as:

1. Identify parsimoniously all physically based input variables with their time memory that influence the output.
2. All input and output sets for the calibration and verification period are normalized.
3. Start with a three-layered ANN model having only one hidden layer. The number of hidden layer could be increased if it is necessary. Start with a minimum number of nodes in hidden layer with approximately double of input nodes (Hipel *et al.*, 1994). The numbers of nodes in the input layer are equal to the number of input variables, whereas, the number of nodes in output layer is equal to the number of output variables.
4. All the interconnecting weights are assigned a small value between –0.5 to +0.5 through a random number generation program.
5. Select fixed or variable values of learning rate and/or momentum term depending upon the algorithm used for optimization.
6. Select the learning process that is either pattern learning or batch learning processes.
7. Execute the program that performs: (a) feed forward calculation, (b) error back propagation in the network and (c) finally change the weight.
8. Estimates output for calibration and verification periods and apply performance evaluation criteria.
9. Perform whole operation for maximum desired iterations.
10. Select the iteration that results in maximum generalization on the basis of performance evaluation criteria.
11. For required generalization, repeat the learning process with assigning more numbers of nodes in the hidden layer or by increasing the numbers of hidden layers.

A FORTRAN code has been developed for accepting any number of layers with capability of considering any number of nodes in any layer. It generates random numbers between zeros to one for all links in the layered system. The program further has capability to normalize the data in all three ways as described earlier. The program uses the pattern learning process with learning and momentum rate as variable or constant. The program can run for any number of iteration. For each iteration it calculates root mean square error, correlation coefficient, coefficient of efficiency, average absolute relative error for both calibration and verification data. The program also calculates threshold statistics for the last iteration specified in the program.

Weaknesses and Limitations

In spite of most popular and successful demonstration of ANN with back propagation scheme, a number of drawbacks exist in this algorithm (Vemuri, 1992).

1. The selection of number of nodes and number of layers in the network is quite arbitrary. It becomes computationally cumbersome if the number of hidden layers is increased.
2. Lack of valid, consistent and efficient criteria for pruning the network. The one successful technique, which emerged to reduce the connectivity, is by starting with a possible smallest network.
3. Selection of parameters such as learning rate (α) and momentum term (β) is totally based on users' experience as, at present, there is no acceptable mathematical basis for this selection.
4. Learning speed, generalization, and converging properties are poorly understood. Frequently, when numbers of hidden units are large, the network begins to memorize, instead of acceptable learning and genera-
lization.
5. The weights of ANN do not reflect any of the system parameters in physical sense.

ANNs have several drawbacks for some applications. They may fail to produce satisfactory results if the data set is insufficient in size or the function is not able to learn the process of the data set. The optimum network geometry as well as the optimum internal network parameters are problem dependent and generally have to be found using a trial and error process. ANNs cannot cope with major changes in the system because they are trained (calibrated) on a historical data set and it is assumed that the relationship learned will be applicable in the future. If there were any major changes in the system, the neural network would have to be adjusted to the new process. Some research studies have been reported in the literature to evolve better algorithms for the generalization of ANN model, i.e., to estimate the extreme events. Imrie *et al.* (2000) added a guidance system to a training algorithm and found that the new algorithm performed better in estimating the extreme events. Guidance system includes cross-checking of the model performance, selection of output transfer function and fixing of the range of normalization.

Performance Evaluation of ANN Model

The performance of a model can be evaluated in terms of several characteristics. The three important characteristics of a good model are accuracy, consistency and versatility (Kachroo, 1992). The term accuracy refers to the ability of the model to reduce the calibration error to the observed data of the calibration period. The presence of level of accuracy and the estimate of the parameter values through different samples of data is referred as the consistency of the model. A versatile model is defined as the model, which is accurate and consistent when used for the diverse applications involving model evaluation criteria not directly based on the objective function used during the calibration of the model. Few performance evaluation criteria selected in the present study are: Root Mean Square Error (RMSE), Correlation Coefficient (CC), Coefficient of efficiency (CF).

While judging the acceptability of a model through evaluation criteria, the ability of each criterion must be properly understood. Root mean error (MSE) shows the measure of mean residual variance summed over the period (Yu, P.S., 1994). Correlation Coefficient (CC) defines the degree of correlation between two variables. The correlation between the observed and estimated value are accounted by the correlation statistic. Coefficient of Efficiency (CE) criterion has the basis of standardization of the residual variance with initial variance (Nash and Sutcliffe, 1970). In this criterion, a perfect agreement between the observed and estimated output yields an efficiency of 1.0. For a zero agreement, all the estimated values must be equal to the observed mean. A negative efficiency represents a lack of agreement worse than if the estimated values are replaced with the observed mean. In this criterion, the value of efficiency strongly depends upon the initial variance of the observed records. Volumetric error (EV) is absolute prediction error. The AARE and TS not only give the performance index in terms of prediction but also the distribution of the prediction error. The criterion can be expressed for different levels of absolute relative error for the model (Nayak, 2004).

STUDY AREA

River Tawi is a major left bank tributary of Chenab River. The river originates from the lapse of Kali Kundi glacier and adjoining area south-west of Bhadarwah in Doda District (J&K), at an elevation of about 4000 m. Initially, it flows in westerly direction for about 16 km then turns towards north-western direction for a distance of about 27 km near Sudh Mahadev. Then it follows a westerly direction for about 5 km up to Chenani, and further down a south-westerly direction unto Udhampur, after which it takes a southerly course for about 24 km. In the lower part of this reach, a number of small streams or *nallahs* join with it. The river then takes south-westerly direction for about 8 km and then north-westerly direction for about 12 km where Jhajjar *nallah* joins it on the right bank. From this point, the river takes a southwesterly direction for about 24 km up to Jammu city, beyond which it flows in the same direction for another 25 km in braided pattern before entering into Pakistan. It has nine tributaries carrying mostly monsoon discharge. Total length of the river is about 141 km. The river in general flows through steep hills on either side excepting the lower reach for about 35 km. Width of the Tawi River is about 300 m. at Jammu bridge site.

Physiography

Tawi river catchment is contained between 32° 35′–33° 5′ N latitude and 74° 35′–75° 45′ E longitude (Figure 3). The catchment area of the river up to Indian border (Jammu) is 2168 km^2, and falls within the districts of Jammu, Udhampur, and a small part of Doda.

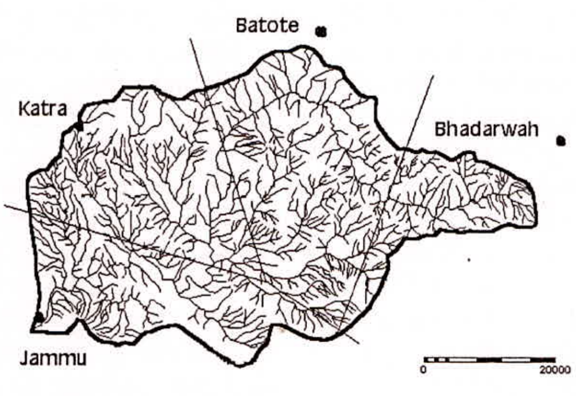

Fig. 3: Index map of Tawi river basin showing observation stations, thiessen polygon and other information

The upper part of the catchment is characterized by rugged mountainous topography whereas lower part consists of low hills and aggradational plains. Elevation in the catchment varies between 400–4000 m, with an average elevation of about 2200 m. The slope of the catchment is from East to West in the upper part and North-east to South-west in the lower part.

The river Tawi originates from the Kaplas granite Dom and cuts across Panjal, Murree, Muran and Krishanpur Thrusts throughout the course up to Jammu. The main tributaries of Tawi, namely, Duddar Khad, Ramnagar Wali Khad, Salam, Naddal, Biruni, Baramani, Juni, Balin, Bamir Khad, Jhajjar, Gamhi, Dhamal Khad, Saro, Sulok, follow strike direction and are controlled by geological structures. The control of structures on the streams is generally evident from the drainage pattern.

Climate

The climate of Tawi catchment is characterized with three distinct features. The north-eastern catchment area comprising of Bhadarwah and adjoining area where climate is extra-tropical mountain type. In this area, winter is severe and influence of south-west monsoon is negligible. Central territory consisting of Udhampur district where, climate is of mountain type, monsoon has sufficient influence. The south-western zone consisting of Jammu district where climate is warm with strong influence of monsoon. The south-west monsoon is active during June to September, and is pronounced in July and August. The upper reaches of the catchment area experience snowfall during December to March. An area of about 200 km^2 of the Tawi catchment is snowfed.

Soils

Soils of Doda district are generally encountered on hill top, mid-hills or plateaus with undulating to partially leveled/terraced lands. The formation is mainly alluvial in nature whereas in the mid-lands or foot-hills the process of colluviation seems predominant. Generally, silty material is brought down from above by the action of water and gets deposited at the base hill. The texture, in general, varies from sandy loam to silty clay loam. The area is considerably bisected and perennial and seasonal *nallahs* are encountered in large numbers.

The major portion of Udhampur district consists of hilly terrain. The soils occurring in plateaus or mid-hills are well managed, properly terraced or bounded. The soils are moderately deep to deep on the mid hills and plateaus whereas deep to very deep at the foot-hills. The texture, in general, is coarse to medium.

Soils of Jammu district are alluvial subtropical having a texture varying between sandy loam to silty clay loam. The lower part is recent alluviam whereas the outer plains are pleistocene. The foot-hills of Siwaliks are moderately

Table 1: Annual Summary of the Data for Tawi River Basin in J&K

Year	*Discharge Data*		*Rainfall Data*					*Ratio*
	Discharge m^3/s	*Discharge mm*	*Jammu mm*	*Katra mm*	*Bhadarwah mm*	*Batote mm*	*Weighted mm*	
1992	18038	718.8	1355.4	2414.0	1359.1	1927.5	1796.3	0.40
1993	24825	989.3	1773.1	2776.2	1739.6	2277.8	2162.7	0.46
1994	23005	916.8	2073.8	2865.4	1441.9	1700.9	2018.8	0.45
1995	19478	776.3	1413.9	2247.7	1540.7	1638.9	1693.3	0.46
1996	29777	1186.7	1970.5	2734.8	1140.6	1830.3	1963.6	0.60
1997	21527	857.9	1739.8	2611.0	1285.6	1822.0	1893.2	0.45
1998	9478	377.7	1214.5	1711.3	1389.6	1715.7	1513.1	0.25
1999	5867	233.8	1127.7	1749.3	835.1	1033.5	1191.3	0.20
2000	6213	247.6	1291.8	2106.9	956.5	1080.4	1355.5	0.18
2001	8256	329.0	1289.4	1714.9	870.3	980.7	1208.1	0.27
2002	7797	310.7	1067.5	1526.0	981.0	1039.4	1141.4	0.27

deep to deep soils with coarse textures having stoney face in general. The lower plains are having a slope gradient of 0 to 3% whereas in the upper reaches it may go as high as 25%.

Data Used

For this study, the monthly rainfall of Jammu, Katra, Bhadarwah, Batote and runoff of the Jammu Bridge for years January, 1992 to December 2002 has been used. The missing rainfall records during September, 1992 to June, 1995 were estimated by developing monthly rainfall runoff relationships for each site from the rest of available records. The monthly data of eleven years is converted to annual records (Table 1) and the runoff rainfall ratio has been estimated. The ratio for first six years is seen in the range of 0.40 to 0.60 however, in the rest five years, it varied in the range 0.18 to 0.27. In first glance, the ratio indicates that the catchment has improved from year 1998, but looking at the quite low rainfall of the same period, no conclusion can be drawn. Under the circumstance, it can be said as the variability of data with the period selected for the study.

MODEL DEVELOPMENT

The development of ANN rainfall runoff model involves the following steps. (1) Identification of effective input variables of the model, (2) Fragmentation of data set for training, testing and validation of the model (3) Model development. The model development involves the following steps, (a) Selection of structure of the model, (b) Normalization of data (c) ANN model procedure (d) Training, testing and validation of model.

Identification of Effective Input Variables of the Model

To describe the physical phenomena of runoff simulation model, the input variable is only the rainfall. Now the criterion for selection of lags for input variable suggests the correlation analysis between output variable and lagged input variables. The cross-correlation analysis is performed using runoff variable at time t and t+1 and the rainfall variables at time t, t–1 and t–2 considering the whole data set (Table 2).

Table 2: Cross-correlation Analysis of Rainfall and Runoff Variables in Time

Variable	*Rain (t–2)*	*Rain (t–1)*	*Rain (t)*	*Run (t)*	*Run (t+1)*
Rain (t–2)	1.00				
Rain (t–1)	0.34	1.00			
Rain (t)	–0.07	0.34	1.00		
Run (t)	0.14	0.43	0.80	1.00	
Run (t+1)	–0.06	0.15	0.43	0.48	1.00

From the correlation matrix, the significant model variables can be identified. It can be seen from the matrix that the runoff at time t is highly correlated to rainfall at t and t–1 (0.80 and 0.43) but not to rainfall at t–2 (0.14). At the same time, the rainfall at time t–1 is not highly related to rainfall at time t–2. It is therefore a simulation model between rainfall and runoff can be developed considering runoff at tine t and rainfall at time t and t–1 as these variables are mutually independent and could be considered highly dependent on output runoff. The simulation model occupies the following form:

$$\text{Runoff}\,(t) = f\,\{\text{Rainfall}\,(t),\ \text{Rainfall}\,(t\text{–}1)\}$$

Fragmentation of Data Set for Training, Testing and Validation

Monthly rainfall of Jammu, Katra, Bhadarwah, Batote and run-off of the Jammu Bridge for years January, 1992 to December, 2002 has been selected for training, testing and validation of the model. The annual summary of data indicates two distinct groups in data; (1) A high run-off rainfall ratio in the range of 0.40 to 0.60 and (2) a low runoff rainfall ratio group in the range 0.18 to 0.27. Therefore, for modelling the data could be divided into three blocks. The first block will be from 1992 to 1995 (four years), the second block will be for years 1996 and 1997 and the third block will be from1998 to 2002 (five years). Keeping for any improvement in the modelling, the modelling with these three blocks could be done in four different ways as shown in the Table 3 referred as case one and two.

Case one is the model developed with a high runoff rainfall ratio (with relatively high rainfall) and verification on low runoff rainfall ratio (with relatively low rainfall). The case two is the model development with low runoff rainfall ratio (with relatively low rainfall) and verification on high runoff rainfall ratio (with relatively high rainfall). In third case, a random selection of the data has been done from the full data set and the model developed has been verified on high runoff rainfall ratio (with relatively high rainfall) and on low runoff rainfall ratio (with relatively low rainfall). In fourth case, the data of years 1998 to 2002 has been considered for model formulation and verification.

Table 3: Data Length Considered for the Cases of ANN Modelling

Case of Model	*Modelling Process*	*Data Considered*	*Years Considered*
Case 1	Calibration	1992 to 1995	4
	Testing	1996 & 1997	2
	Verification	1998 to 2002	5
Case 2	Calibration	1998 to 2002	5
	Testing	1996 & 1997	2
	Verification	1992 to 1995	4
Case 3	Calibration (random 50 %)	1992 to 2002	11
	Verification	1992 to 1995	4
	Verification	1998 to 2002	5
Case 4	Calibration	1998 to 2000	3
	Verification	1901 & 1902	2

Model Development

Three and four layers BPANN models were developed with non-linear sigmoid as activation function uniformly between the layer to simulate rainfall runoff process. While deciding the structure of ANN, nodes in input layer were equal to number of input variables, nodes in hidden layer were approximately double of input nodes (Hipal, 1994) and the node in the output layer was one as the model provides single output. The learning of ANN initiates with the normalisation (re-scaling) of all the data with three methods viz. Normalisation by maximum value (code 1), normalisation by using both maximum and minimum value (code 2), and normalisation by using both maximum and minimum value with flexibility of extrapolation (code 3). All inter-connecting links between nodes of successive layers were assigned random values called weights between +0.5 to –0.5 (Dawson and Wilbay, 1998) and a value decreasing from 0.5 is considered for both learning rate (α) and momentum (β) term (Orri, 1995; Raman, 1995). The learning is done by pattern learning process and is carried out for maximum one thousand iterations using the error of each data set, one by one, and the weight continuously updated with the processing of each data set.

Calibration, Cross-Validation and Verification of Model

In model development, the learning is through generalisation of ANN for which the developed model is simultaneously checked for its improvement in efficiencies during the phase of model formulation. As can be seen that the data of Tawi river basin is highly variable, the modelling has been done considering the data in different ways (Table 3) as to have any applicability of model and its improvement.

Case 1

Model calibration efficiencies AERE, CC, CE and EV reported in Table 4 for four layer system is found relatively better than three layer system. Normalisation by method (code 2) slightly improves the efficiencies in model development over the normalisation by first method (code 1). The third method of normalisation that provides the flexibility of extrapolation (code 3) has resulted adversely on the model efficiencies. In cross-validation, the efficiencies CE and EV are not as good compared to the model development which is quite obvious as the run-off rainfall ratio for years 1996 and 1997 are 0.60 and 0.45 relatively higher compared to the range covered in model development. In verification of the model, again CE and EV are not acceptable and are worse than the cross validation period. Positive values of EV for this period suggested an over estimation of runoff obviously due to low values of run-off rainfall

Table 4: Performance Evaluation of Developed Model with the Data of Case One for Different ANN Structures and Normalisation Methods

Number of Layers	*Normalisation Method*	*Calibration Period, 1992 to 1995*				*Cross-validation Period, 1996 & 1997*				*Verification Period, 1998 to 2002*			
		AARE	*CC,%*	*CE,%*	*EV,%*	*AARE*	*CC,%*	*CE,%*	*EV,%*	*AARE*	*CC,%*	*CE,%*	*EV,%*
3(2,3,1)	1	0.5	82.8	65.8	−8.0	1.3	78.5	36.1	−31.1	1.3	74.3	−100.0	92.7
3(2,4,1)	1	0.5	89.3	79.6	−1.1	1.3	84.5	54.7	−25.9	1.3	74.1	−91.6	90.4
3(2,5,1)	1	0.5	89.2	79.6	−1.0	1.3	84.3	54.9	−25.6	1.3	74.9	−103.5	94.2
4(2,3,3,1)	1	0.5	89.0	79.2	2.3	1.4	84.2	55.3	−24.6	1.5	75.6	−117.4	102.4
4(2,4,4,1)	1	0.5	88.9	78.8	4.6	1.4	83.8	55.4	−22.6	1.5	74.7	−137.7	106.6
4(2,5,5,1)	1	0.5	90.0	80.9	3.4	1.4	85.1	57.8	−22.3	1.5	75.4	−102.7	100.6
3(2,3,1)	2	0.5	82.1	64.5	−8.5	1.4	77.3	35.0	−30.2	1.3	73.4	−104.1	91.2
3(2,4,1)	2	0.5	89.0	79.2	−1.1	1.3	84.1	54.2	−25.2	1.2	73.8	−101.6	89.8
3(2,5,1)	2	0.5	89.0	79.2	−0.2	1.3	83.9	54.5	−25.2	1.3	74.3	−110.5	92.6
4(2,3,3,1)	2	0.5	88.8	78.8	2.4	1.3	84.0	55.2	−23.4	1.4	74.8	−131.5	100.6
4(2,4,4,1)	2	0.5	88.7	78.5	4.6	1.4	83.4	54.8	−22.1	1.5	74.0	−141.9	105.7
4(2,5,5,1)	2	0.5	89.9	80.6	3.4	1.4	84.8	57.5	−22.0	1.4	75.0	−111.6	100.1
3(2,3,1)	3	1.1	84.0	60.8	24.0	2.5	79.7	41.1	−1.4	2.9	70.1	−384.4	184.9
3(2,4,1)	3	1.1	87.6	66.2	26.1	2.5	84.5	50.1	−0.1	3.0	71.0	−362.9	185.1
3(2,5,1)	3	1.1	87.9	66.9	26.3	2.5	84.8	51.1	0.3	2.9	71.4	−363.4	184.8
4(2,3,3,1)	3	1.2	87.2	64.5	28.3	2.6	84.2	49.6	1.3	3.1	70.8	−388.5	191.4
4(2,4,4,1)	3	1.1	82.3	57.1	23.8	2.6	77.7	37.2	−0.6	2.9	69.9	−398.1	188.9
4(2,5,5,1)	3	1.1	88.4	66.2	29.0	2.6	85.1	52.2	2.3	3.0	72.0	−391.0	191.8

ratio during this period. Model calibration, cross-validation and verification efficiencies suggests that the whole eleven years data does not form a homogeneous group and modelling and verification with such data is highly difficult.

Case 2

Data considered in the modelling of case two is opposite of the first as discussed earlier and reported in Table 5 The model development efficiencies are not better than case 1. However, the results in respect of CE and EV are not acceptable and are similar as obtained in case one. This case of model calibration, cross-validation and verification efficiencies again suggests that the whole eleven years data does not form a homogeneous group and modelling and verification with such selection of data is again difficult.

Case 3

Since, the variability in data is high; it was decided to develop the model considering 50 percent of data randomly selected from the full domain. The developed model is not cross validated but directly verified on the lower and upper bonds of the data domain. The results (Table 6) were similar as obtained in case 1 and case 2. The verification results for period 1992–1995 were slightly better than the period 1998–2002.

Case 4

In this case, the data considered is almost homogeneous in respect of runoff rainfall ratio. First three years are considered for the model developed and rest two years for model verification. Even in this case the results (Table 7) in respect of efficiencies are not better than case 1, 2 and 3.

Overall the variability in data of Tawi river basin is too high. Using this data no significant model could be developed and therefore the impact of normalization by three methods could not be significantly tested.

CONCLUSIONS

An ANN model for the rainfall-runoff process was developed for Tawi catchment up to Tawi Bridge at Jammu. Three-layered feed forward network structure was used to model the process. The monthly rainfall and discharge values of the years 1992 to 2002 were considered for the development of ANN model, cross validation and for validation purpose.

Table 5: Performance Evaluation of Developed Model with the Data of Case Two for Different ANN Structures and Normalisation Methods

Number of Layers	Normalisation Method	Calibration Period, 1998 to 2002				Cross-Validation Period, 1996 & 1997				Verification Period, 1992 to 1995			
		AARE	CC,%	CE,%	EV,%	AARE	CC,%	CE,%	EV,%	AARE	CC,%	CE,%	EV,%
3(2,3,1)	1	0.4	78.5	59.9	–10.7	0.7	71.0	–7.4	–69.1	0.5	71.9	–3.5	–58.9
3(2,4,1)	1	0.4	78.5	60.0	–11.0	0.7	71.0	–7.4	–69.1	0.5	71.9	–3.5	–58.9
3(2,5,1)	1	0.4	81.9	66.9	–2.5	0.7	78.3	2.1	–66.4	0.4	72.8	12.3	–52.3
4(2,3,3,1)	1	0.4	82.1	67.4	0.9	0.7	78.1	2.3	–65.9	0.4	73.0	13.1	–51.4
4(2,4,4,1)	1	0.4	82.4	67.9	0.9	0.8	78.3	1.5	–66.7	0.4	73.2	12.2	–51.9
4(2,5,5,1)	1	0.4	82.2	67.5	0.4	0.7	78.6	2.5	–66.0	0.4	73.0	13.2	–51.5
3(2,3,1)	2	0.4	75.7	55.7	–8.7	0.7	69.7	–8.8	–67.5	0.5	70.8	–6.6	–58.6
3(2,4,1)	2	0.4	75.7	55.7	–8.6	0.7	69.7	–8.8	–67.5	0.5	70.8	–6.6	–58.6
3(2,5,1)	2	0.4	80.9	65.3	–1.5	0.7	77.3	1.5	–65.2	0.4	72.5	11.2	–52.1
4(2,3,3,1)	2	0.4	81.7	66.7	1.7	0.7	77.9	2.4	–65.0	0.4	72.8	13.0	–51.0
4(2,4,4,1)	2	0.4	82.4	67.9	1.5	0.8	78.2	1.5	–66.6	0.4	73.2	12.3	–51.0
4(2,5,5,1)	2	0.4	82.0	67.1	1.4	0.7	78.6	2.7	–65.3	0.4	72.8	13.4	–51.0
3(2,3,1)	3	0.8	79.1	52.1	27.5	1.0	75.7	2.3	–56.5	0.4	73.1	13.9	–43.8
3(2,4,1)	3	0.8	79.1	52.1	27.4	1.0	75.7	2.2	–56.5	0.4	73.2	13.8	–43.8
3(2,5,1)	3	0.8	79.3	52.2	28.0	1.0	75.9	2.8	–56.3	0.4	73.1	14.7	–43.4
4(2,3,3,1)	3	0.8	79.9	51.7	30.2	1.0	76.1	3.4	–55.9	0.4	73.5	15.8	–42.7
4(2,4,4,1)	3	0.8	81.8	54.9	29.4	1.0	77.6	4.8	–56.9	0.5	74.0	17.8	–42.6
4(2,5,5,1)	3	0.8	80.0	51.9	30.3	1.0	76.4	3.9	–55.8	0.4	73.4	16.5	–42.4

Table 6: Performance Evaluation of Developed Model with the Data of Case Three for Different ANN Structures and Normalisation Methods

Number of Layers	Normalisation Method	Calibration Period, 1992 to 2002				Verification Period, 1992 to 1995				Verification Period, 1998 to 2002			
		AARE	CC,%	CE,%	EV,%	AARE	CC,%	CE,%	EV,%	AARE	CC,%	CE,%	EV,%
3(2,3,1)	1	0.6	79.4	60.3	–12.4	0.4	82.4	61.0	–23.8	0.7	79.2	–6.7	41.1
3(2,4,1)	1	0.6	79.4	60.3	–12.1	0.4	82.5	61.0	–23.6	0.7	79.2	–6.0	41.9
3(2,5,1)	1	0.6	86.6	74.8	–5.7	0.4	86.1	69.3	–8.8	0.6	78.0	–18.2	45.1
4(2,3,3,1)	1	0.6	87.1	75.7	–3.5	0.4	86.0	69.7	–6.5	0.7	77.7	–9.1	51.7
4(2,4,4,1)	1	0.6	86.9	75.3	–4.3	0.4	86.0	69.7	–7.3	0.7	76.9	–6.4	49.5
4(2,5,5,1)	1	0.6	87.4	76.1	–3.4	0.4	86.2	69.6	–5.9	0.7	77.5	–1.5	50.6
3(2,3,1)	2	0.7	78.5	59.2	–12.4	0.5	82.2	60.3	–24.1	0.7	78.7	–9.8	40.3
3(2,4,1)	2	0.6	86.1	73.8	–7.1	0.4	86.4	70.7	–10.9	0.6	78.5	–28.3	42.4
3(2,5,1)	2	0.6	86.3	74.2	–6.7	0.4	86.2	69.7	–9.9	0.6	77.9	–20.6	42.4
4(2,3,3,1)	2	0.6	87.1	75.7	–4.2	0.4	86.2	70.0	–7.2	0.7	77.8	–8.3	49.7
4(2,4,4,1)	2	0.6	86.9	75.2	–5.4	0.4	86.2	70.1	–8.3	0.7	76.9	–5.3	47.0
4(2,5,5,1)	2	0.6	87.4	76.1	–4.2	0.4	86.4	69.9	–6.6	0.7	77.6	–0.6	48.6
3(2,3,1)	3	1.8	75.4	41.7	37.1	0.9	81.6	57.4	8.1	2.3	74.8	–208.0	148.4
3(2,4,1)	3	107	83.8	57.3	38.3	0.8	86.2	71.0	15.6	2.2	76.3	–200.5	142.2
3(2,5,1)	3	1.6	84.0	58.0	37.7	0.8	86.3	71.3	15.4	2.2	76.3	–195.2	140.8
4(2,3,3,1)	3	1.6	84.6	59.5	34.5	0.8	86.6	72.2	12.8	2.1	76.2	–173.1	136.2
4(2,4,4,1)	3	1.6	84.8	59.6	34.4	0.8	86.7	72.3	12.7	2.2	76.3	172.0	136.3
4(2,5,5,1)	3	1.7	86.1	61.6	35.9	0.8	86.	72.3	15.5	2.2	76.8	–173.5	138.3

Table 7: Performance Evaluation of Developed Model with the Data of Case Four for Different ANN Structures and Normalisation Methods

Number of Layers	*Normalisation Method*	*Calibration Period, 1998 to 2000*				*Verification Period, 2001 & 2002*			
		AAERE	*CC,%*	*CE,%*	*EV,%*	*AARE*	*CC,%*	*CE,%*	*EV,%*
3(2,3,1)	1	0.3	72.4	52.3	–2.9	0.5	90.2	55.1	–24.6
3(2,4,1)	1	0.3	72.3	52.1	–2.9	0.5	90.2	55.2	–24.7
3(2,5,1)	1	0.3	72.8	53.0	–1.9	0.5	91.3	58.2	–23.1
4(2,3,3,1)	1	0.3	72.8	52.9	–2.6	0.5	91.7	57.2	–23.7
4(2,4,4,1)	1	0.3	72.3	52.0	–3.0	0.5	90.5	53.6	–24.6
4(2,5,5,1)	1	0.3	73.1	53.3	–2.7	0.5	91.6	57.8	–23.8
3(2,3,1)	2	0.3	70.6	49.7	–2.8	0.6	86.3	49.7	–24.7
3(2,4,1)	2	0.3	70.6	49.6	–2.8	0.6	86.3	49.7	–24.7
3(2,5,1)	2	0.3	72.1	51.9	–1.8	0.5	89.6	55.9	–23.2
4(2,3,3,1)	2	0.3	72.0	51.7	–2.7	0.5	87.8	55.2	–24.1
4(2,4,4,1)	2	0.3	71.9	51.5	–2.3	0.5	89.5	52.7	–23.8
4(2,5,5,1)	2	0.3	72.3	52.2	–2.9	0.5	89.7	55.4	–24.1
3(2,3,1)	3	0.6	69.5	36.0	24.2	1.0	85.0	46.3	2.5
3(2,4,1)	3	0.6	69.3	35.7	24.0	1.0	84.5	45.3	2.4
3(2,5,1)	3	0.5	71.1	38.9	24.2	0.9	89.0	52.5	3.3
4(2,3,3,1)	3	0.5	70.8	38.7	23.5	0.9	87.7	50.6	2.3
4(2,4,4,1)	3	0.6	70.9	38.6	23.7	0.9	88.0	50.0	2.4
4(2,5,5,1)	3	0.5	71.1	39.6	23.1	0.9	88.4	51.6	2.1

Four different combinations of rainfall and run-off were considered as input to the network and trained by BP algorithm with different error tolerance, learning parameter, number of cycles and number of hidden layers. It was observed from the training results that the combination of rainfall (t–4), rainfall (t–3), rainfall (t–2), rainfall (t–1), runoff (t–1) and rainfall (t) as input and runoff (t) as output was the best combination compared to other combinations as it results in high coefficient of correlation (0.988), and low root mean square error (305). But the validation of the model did not give as good result as in the case of the training. The coefficient of correlation and root mean square error for the validation run was 0.624 and 1680, respectively. In the training of ANN models, the main objective is to achieve a global minimum error on the whole length of the data. Training the model with long record of data, which contain more extreme events, can reduce the large variations in the ANN model. Inclusion of priori knowledge such as antecedent precipitation index (API), φ index, temperature, humidity, slope, cover condition in the training stage itself can improve the performance of the model.

REFERENCES

ASCE (2000a). "Artificial neural networks in hydrology-I: Preliminary concepts." *Journal of Hydrologic Engineering*, ASCE, 5(2), 115–123.

ASCE (2000b). "Artificial neural networks in hydrology-II: Hydrologic applications." *Journal of Hydrologic Engineering*, ASCE, 5(2), 124–137.

Burian, S.J., Durrans, S.R., Nix, S.J. and Pitt, R.E. (2001). "Training artificial neural networks to perform rainfall disaggregation." *Journal of Hydrologic Engineering*, ASCE, 6(1), 43–50.

Bevan, K.J., Lamb, R, Quinn, P.F., Romanowicz, R. and Freer, J. (1995). "TOPMODEL in Singh, V.P. (Ed). Computer models of watershed hydrology." *Water Resources Publications*, 627–668.

CWC (1988). Outline plan report for irrigation development in Chenab basin. Central Water Commission, New Delhi.

Central Water Commission (1990). Discharge of Chenab Basin (1967–1989), Northern Investigation Circle, Projects Investigation Organization, Central Water Commission, Jammu.

Central Water Commission (1991). Water Year Book (1989–1990), Chenab Basin, Northern Circle, Project Investigation Organization, Central Water Commission, Jammu.

Chow, V.T., Maidment, D.R. and Mays, L.W. (1988). *Applied hydrology*, McGraw-Hill Book Company.

Daniell, T.M. (1991). "Neural Networks-applications in hydrology and water resources engineering." *Int. Hydro. and Water Resour. Sump.*, Inst. of Engineers, Perth, Australia, 797–802.

Danish Hydraulic Institute (1988). "Hydrological computerized modeling system (SHE)." Agern Alles, DK-2970 Horsholm, Denmark.

Dawson, C.W. and Wilby, R. (1998). "An artificial neural network approach to rainfall-run-off modeling." *Hydrological Sciences Journal*, 43(1), 47–66.

Elshorbagy, A., Simonovic, S.P. and Panu, U.S. (2000). "Performance evaluation of artificial neural networks for runoff prediction." *Journal of Hydrologic Engineering*, 5(4), 424–433.

Fernando, A.K. and Jayawardena, A.W. (1998). "Runoff forecasting using RBF networks with OLS algorithm." *Journal of Hydrologic Engineering*, ASCE, 3(3), 203–209.

Goyal, V.C. (2000–2001). Low-flow analysis of Tawi river (Jammu & Kashmir). National Institute of Hydrology, (Under publication).

Hsu, K-L., Gupta, H.V. and Sorooshian, S. (1995). "Artificial neural network modeling of the rainfall-runoff process." *Water Resources Research*, 31(10), 2517–2530.

Jain, S.K. (1991–1992). Land use mapping of Tawi catchment using satellite data. National Institute of Hydrology, CS-72 (Unpublished).

Kachroo, R.K. (1992). "River flow forecasting, Part I. A discussion of the principles." *Journal of Hydrology*, 133, 1–15.

Maier, H.R. and Dandy, G.C. (2000). "Neural networks for the prediction and forecasting of water resources variables: A review of modeling Issues and applications". *Environmental Modelling & Software*, 15, 101–124.

McCulloch, W.S. and Pitts, W. (1943). "A logic calculus of the ideas immanent in nervous activity." *Bull. of Math. Biophys.*, 5, 115–133.

Nash, J.E. and Sutcliffe, J.V. (1970). "River flow forecasting through conceptual models:1. A discussion of principles." Journal of Hydrology, 10, 282–290.

Patwary, B.C. and Kamal, Kumar 1991–92. Water availability study of river Tawi (J&K State). National Institute of Hydrology, CS-86 (Unpublished).

Raman, H. and Sunil Kumar, N. (1995). "Multivariate modeling of water resources time series using artificial neural networks." *Hydrological Sciences Journal*, 40(2), 145–163.

Ramasastri, K.S. (1992–1993). Hydrological network for Tawi (J&K). National Institute of Hydrology, TR-161 (Unpublished).

Rumelhart, D.E., McLelland, J.L. and the PDP Research Group (1986). *Parallel distributed processing, explorations in the micro structure of cognition, vol. I: Foundations.* MIT Press, Cambridge, Mass.

Shanker, Ravi, Kumar, G. and Saxena, S.P. (1989). Geological Map of the Himalaya- Western Sector (Sheet No. 1).

Sajikumar, S. and Thandaveswara, B.S. (1999). "A non-linear rainfall-runoff model using an artificial neural network." *Journal of Hydrology*, 216, 32–55.

Smith, J. and Eli, R.N. (1995). "Neural network models of rainfall-runoff process." *Journal of Water Resources Planning and Management*, ASCE, 121(6), 499–508.

Snyder, F.F. (1938). "Synthetic unit-graphs", *Trans. Am. Geophys. Union*, Vol. 19, 447–454.

Sudheer, K.P., Gosain, A.K. and Ramasastri, K.S. (2001). "Selection of appropriate input vector to neural network based rainfall-runoff models: A statistical approach." *International Conference on Civil Engineering*, Bangalore, July 2001, 464–471.

Thirumalaiah, K. and Deo, M.C. (1998). "River stage forecasting using artificial neural networks." *Journal of Hydrologic Engineering*, ASCE, 3(1), 26–32.

Thirumalaiah, K. and Deo, M.C. (2000). "Hydrological forecasting using neural networks." *Journal of Hydrologic Engineering*, ASCE, 5(2), 180–189.

Tokar, A.S. and Johnson, A. (1999). "Rainfall-runoff modeling using artificial neural networks". *Journal of Hydrologic Engineering*, ASCE, 4(3), 232–239.

Tokar, A.S. and Markus, M. (2000). "Precipitation-Runoff modeling using artificial neural networks and conceptual models." *Journal of Hydrologic Engineering*, ASCE, 5(2), 156–161.

Zealand, C.M., Burn, D.H. and Simonovic, S.P. (1999). "Short term streamflow forecasting using artificial neural networks." *Journal of Hydrology*, 214, 32–48.

Zhang, B. and Govindaraju, R.S. (2000). "Prediction of watershed runoff using Bayesian concepts and modular neural networks." *Water Resources Research*, 36(3), 753–762.

Zhu, M.-L. and Fujita, M. (1994). "Comparisons between fuzzy reasoning and neural network methods to forecast runoff discharge." *Journal of Hydroscience and Hydraulic Engineering*, 12(2), 131–141.