INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X
Impact factor: 4.295
(Volume 3, Issue 6)
Available online at www.ijariit.com

Rainfall Runoff Modelling Using Artificial Neural Network

Danish Hussain
Azad Institute of Engineering & Technology, Lucknow, Uttar Pradesh
danishhusain969@gmail.com

Ashraf Usmani
Azad Institute of Engineering & Technology, Lucknow, Uttar Pradesh
ashrafusmani11@gmail.com

Deeak Kumar Verma
Azad Institute of Engineering & Technology, Lucknow, Uttar Pradesh
deepak00515@gmail.com

Farooq Jamal
Azad Institute of Engineering & Technology, Lucknow, Uttar Pradesh
farooqjamal637@gmail.com

Maaz Allah Khan
Azad Institute of Engineering & Technology, Lucknow, Uttar Pradesh
danishhusain969@gmail.com

***Abstract:** The use of an artificial neural network (ANN) is becoming very common nowadays due to its ability to analyse complex nonlinear events. An ANN has a flexible, convenient and easy mathematical structure to identify the nonlinear relationships between input and output data sets. This capability could efficiently be employed for the different hydrological models such as rainfall-runoff models, which are inherently nonlinear in nature and therefore, representing their physical characteristics is challenging. In this paper, the influences of back propagation algorithm and their efficiencies which affect the input dimensions on rainfall runoff model have been demonstrated. The capability of the Artificial Neural Network with different input dimensions has been attempted and demonstrated with a case study on Sarada River Basin. The ANN models developed were able to map the relationship between input and output data sets used. The model developed on rainfall and runoff pattern have been calibrated and validated. The significant input variables for the training of ANN models were selected based on statistical parameters like cross-correlation, autocorrelation, and partial autocorrelation function. It was found that those models considering rainfall lag rainfall and discharge as inputs were performing better than those considering rainfall alone. It was found that the neural network model developed was performing well. It can be inferred from the developed model that the Neural Network model was able to predict runoff from rain fall data fairly well for a small semi-arid catchment area considered in the present study.*

***Keywords:** Rainfall-Runoff Model, Artificial Neural Network, Cross-correlation, Auto-correlation.*

INTRODUCTION

In recent years, artificial neural network (ANNs) models have been widely referred as black box models which were successfully used for modelling complex hydrological processes, such as rainfall-runoff that have been shown as viable tools in hydrology. ANN models are built upon the input and output observations. These models have the capability that even without the detailed understanding of the complex physical laws that govern the process under investigation, they are able to provide reasonably accurate results. The application of the method was widely adopted in hydrology. Researchers (M.A. Kaltech 2008) have also compared the performance of developed ANN models with other methods successfully and demonstrated their approach.

The merits and shortcomings of this methodology have also been discussed in the review by the ASCE task committee on application of ANNs in hydrology (ASCE, 2000a, b). They have described that rainfall-runoff modelling has received maximum attention from ANN models. In a preliminary study, Halff et al. (1993) designed a three-layer feed-forward ANN using the rainfall hyetographs as input and hydrograph as output. This study opened several possibilities for a rainfall-runoff application using neural networks. The studies by Smith and Eli (1995) and Kaltech (2008) may be viewed as a 'proof of concept' for the analysis for ANNs in rainfall runoff modelling. Subsequently, a number of studies have shown that employed neural networks for rainfall runoff modelling (Hsu et al., 1995; Tokar and Johnson, 1999; Abrahart and See, 2000). The rainfall runoff process lends itself well to ANN applications.

The nonlinear nature of the relationships, availability of long historical records, and the complexity of the physical based models are some of the factors that have attracted researchers to consider alternative models in which, ANNs have been a one of the viable alternative choices.

Neural Network Model

Artificial Neural Networks (ANNs) are information-processing systems that imitate the functions of the human brain. The ANN structure consists of a number of processing elements, which are called neurons, and the interconnections between the neurons are called weights (Kisi et al., 2013). In the architecture of ANN, neurons are classified in groups, and are called layers. Artificial neural networks uses a mathematical simulation approach that adopts a biological system in order to process the acquired information and derive the output after the network has been trained properly for pattern recognition. The main theme of ANN model is, it considers the brain as a parallel computational device for various computational tasks that were performed relatively poorly by traditional serial computers. The neural network structure in the present study uses a three-layer learning network consisting of an input layer, a hidden layer and an output layer consisting of output variable(s) as shown in Figure 1. The input nodes pass on the input signal values to the nodes in the hidden layer unprocessed. The values are distributed to all the nodes in the hidden layer depending on the connection weights Wij and Wjk (Najjar, Y., Ali, H., 1998) between the input node and the hidden nodes. Connection weights are the interconnecting links between the neurons in successive layers. Each neuron in a certain layer is connected to every other single neuron in the next layer by links having an appropriate and an adjustable connection weight.

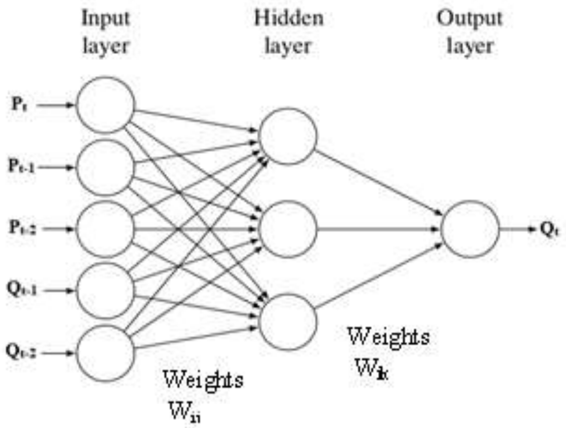

Figure 1. Architecture of the Neural Network Model used in this Study

In this study, the Feed Forward Back Propagation (FFBP) algorithm was used for training using Levenberg–Marquardt optimization technique. This optimization technique is reported to be more powerful than the conventional gradient descent techniques (Y. Najjar and H. Ali, 1998). The study showed that the Marquardt algorithm is very efficient when training networks which has few hundred weights. Although the computational requirements are much higher in iterations of the Marquardt algorithm its efficiency is higher. This is especially true, when high precision is required. The Feed Forward Back Propagation (FFBP) distinguishes itself by the presence of one or more hidden layers, whose computation nodes are correspondingly called hidden neurons or hidden units. The function of hidden neurons is to intervene between the external input and the network output in useful manner.

Method of Application of ANN for Rainfall-Runoff Modelling

The runoff from a watershed outlet is a complex phenomenon and mainly related to the current rainfall rate and also may be to the past rainfall and runoff situations and several hydrological processes. In any discrete or lumped hydrological system, the rainfall-runoff relationship can be generally expressed as per equation 1. (M.T. Hagan, 1994; S.J. Riad, 2004).

Q(t) ¾F[R(t), R(t-Δt),..., R(t-n_xΔt),Q(t-dt),...Q(t n_yΔt)] ......(1)

where R represents rainfall, Q represents runoff at the outlet of the watershed, F is any kind of model structure (linear or nonlinear), Δt is the data sampling interval, and nx and ny are positive integers numbers reflecting the memory length of the watershed. In this study, the Simplex search method is used to find a set of optimum values for those weights used in the ANN, which are denoted by Wij, $0 \leq i \leq n, 1 \leq j \leq i$ and those by Wjk, $0 \leq j \leq 1, 1 \leq k \leq 1, 1 \leq k \leq m$. The estimated runoffs, denoted by Q (t), are determined as a function of those optimum weights of the ANN, which is expressed equation 2.

Q(t) ¾F[R(t), R(t-Δt),..., R(t-n_xΔt),Q(t-dt),...Q(t n_yΔt)] $|$W ij,W jk$|$ ......(2)

When the ANN is implemented to approximate the above relationship between the watershed, average rainfall and runoff there will be a number of n = nx+ny+1 nodes in the input layer, n = nx+ny+1, and one node in the output, i.e. m=1.

The database collected for this study represents ten years daily sets of rainfall-runoff values for the Sarada River Basin.

The length of the data used for calibration of any model depends on data sequence length of the study area and also several factors depending on the model. In the present study, seven years (2001-2008) data were used for calibration and balance three years data was used for validation. The training phase of ANN model will be terminated with the mean squared error (RMSE) and later testing has been performed to achieve minim error. The runoff flow estimation has been carried out in two steps. Initially, only rainfall data has been employed to the input layer. Later, the previous daily flow value has been incorporated into an input data.

Study Area

The Sarada River Basin is located within 82013‟0” E & 83005‟0” longitude and 170 25‟ 0” & 180 17‟ 0” N and latitude. The total area of the study basin is around 1252.99 km2. The Sarada river basin forms a part of Survey of India (SOI) sheets Nos. 65 O/1, 2, 3 and 6 and 65 K/13, 14 and 15 with a scale of 1:50000. The Index map of the study area of Sarada River Basin has been given in Figure 2. After the reconnaissance survey, the watersheds were delineated on the basis of the drainage line, land slope, and outlet point. Furthermore, on the basis of drainage channels and land topography, the Sarada River Basin is subdivided into five sub basin Viz., K.Kotapadu, Madugula, Chodavaram, Kasimkota, and Anakapalli.

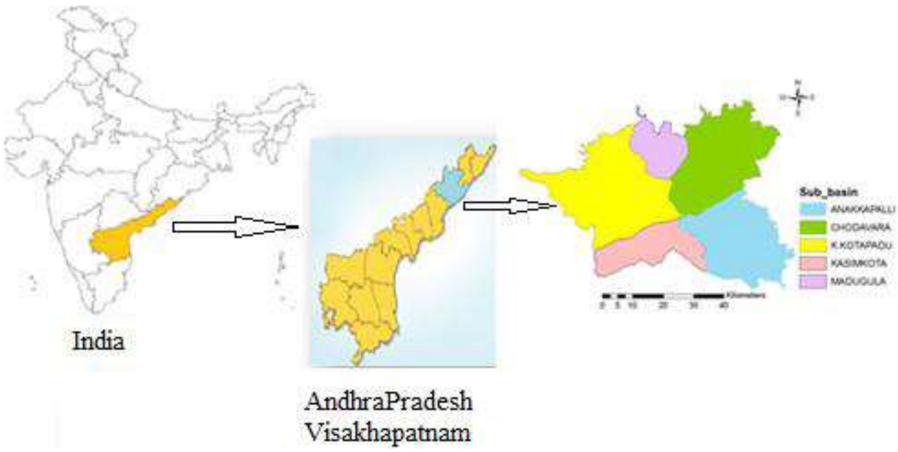

Figure 2. Index Map of the Study Area

Model Performance

The selected basin performance has been evaluated with five performance measures to evaluate the model performance. The performance measures are Nash-Sutcliffe coefficient efficiency (ENS), root means square error (RMSE), mean absolute error (MAE), the coefficient of determination (R2) and the difference in peak (DP). The mathematical expressions of the goodness of fit indices used in the study are represented in equations 3 to 7.

(i). Nash-Sutcliffe coefficient efficiency (ENS): It is expressed as:

$$E=1-\frac{\sum_{i=1}^{n}(X_{obs,i}-X_{model,i})^2}{\sum_{i=1}^{n}(X_{obs,i}-\overline{X_{obs,\iota}})^2} \quad (3)$$

(ii). Correlation Coefficient (r)

$$r=\frac{\sum_{i=1}^{n}(X_i-\overline{X_\iota}).(Y_i-\overline{Y_\iota})}{\sqrt{\sum_{i=1}^{n}(X_i-\overline{X_\iota})^2.(Y_i-\overline{Y_\iota})^2}} \quad (4)$$

(iii). Root Mean Square Error (RMSE)

$$RMSE=\sqrt{\frac{\sum_{i=1}^{n}(X_{obs,i}-X_{model,i})^2}{n}} \quad (5)$$

(iv). Mean absolute error (MAE): It is expressed as:

$$MAE = \frac{\sum_{i=1}^{n}(X_{model,i} - X_{obs,i})}{n} \tag{6}$$

(v). The difference in Peak: It is expressed as:

$$\mathbf{DP = (o_i) - Max\ (p_i)} \tag{7}$$

Mean Areal Rainfall

Thiessen polygon weightage method has been adopted for the analysis of rainfall data and runoff data by considering basin as five sub basins. The runoff data available at Anakapalli gauging station was used to make a rainfall-runoff relationship. The runoff data is up to Anakapalli was used to identify the rain gauge station that contributes to mean annual rainfall of the basin. The Arc-GIS software has been used to develop polygons (Figure 3) and to calculate the area of polygons for better accuracy. The Theissen weightage for each raingauge station was calculated and used to calculate mean areal rainfall over the area. The statistic of Theissen polygon of Sarada River Basin is presented in Table 2.

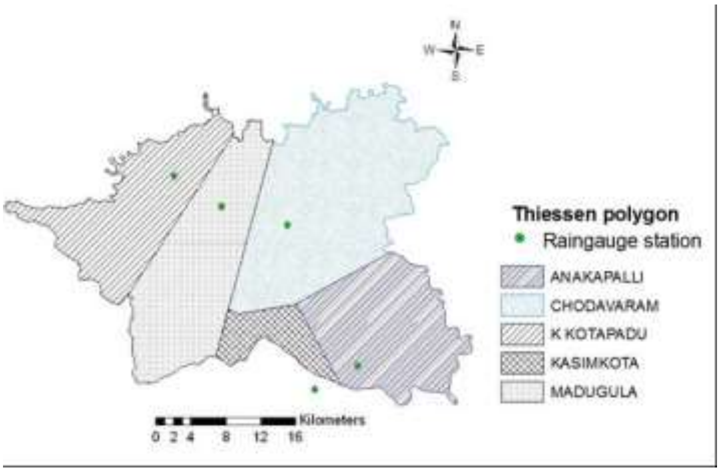

Figure 3. Theissen Polygon for the Study Area

Table 1. Theissen Polygon Statistics

S. No.	Station Name	Area (km^2)	Weightage Factor
1.	Chodavaram	346.6	0.276
2.	Madugula	88.54	0.07
3.	K.Kotapadu	379.1	0.302
4.	Anakapalli	243.24	0.194
5.	Kasimkota	195.51	0.156

RESULTS AND DISCUSSIONS

All the networks selected were calibrated with different combinations of input. Daily rainfall and runoff values are used as input to ANN models. The length of the data used for calibration is 7 years (2556 days) and for validation is 3 years (1096 days). Networks are tested with a different number of hidden neurons and the model structures with least root mean square error (RMSE) is considered as the best structure. The ANN structure was tested for 1 to 6 hidden neurons. It can be observed that by adding hidden neurons RMSE decreases up to a certain value and again increases. Accordingly, selection of hidden neurons is done by comparing RMSE of the network. A total of six combinations of input variables were investigated for the Sarada river basin. Simulated runoff has been compared with that of observed values using performance functions like as Nash Sutcliffe efficiency and RMSE as per the equation 3.

The Goodness-of-fit statistic of each of the six developed ANN models during calibration and validation is presented in Table 2 and Table 3 respectively.

Table 2. Goodness-of-fit Statistics for the Observed and Predicted Daily Runoff for Gauging Station of Sarada River Basin during Calibration Period (2001-2007)

MODEL	E_{NS}	RMSE	R^2	MAE	DP
A	9.202	40.603	0.093	20.557	475.210
B	13.916	39.535	0.196	24.777	444.227
C	22.616	37.484	0.288	16.979	342.672
D	77.644	20.148	0.783	7.733	146.131
E	78.373	19.816	0.815	8.465	153.577
F	85.868	16.019	0.877	7.941	158.418

Table 3. Goodness-of-fit Statistics for the Observed and Predicted Daily Runoff for Gauging Station of Sarada River Basin during Validation Period (2008-2010)

MODEL	E_{NS}	RMSE	R^2	MAE	DP
A	7.116	16.837	0.096	9.070	148.684
B	9.631	16.608	0.159	8954	93.858
C	11.921	16.396	0.129	9.988	139.033
D	71.700	9.128	0.754	5.040	48.112
E	76.684	8.236	0.783	4.407	33.818
F	67.997	9.983	0.729	4.827	41.077

CONCLUSION

The ANN model simulated daily runoff has fairly matched with the observed values. Statistical analyses have also been performed to compare the simulated daily runoff with its measured counterpart. The high coefficient of determination (R2) values of 0.815 and model efficiency of 78.37 % shows the close agreement between the measured and simulated runoff value during the calibration period. The coefficient of determination (R2) values of 0.783 and model efficiency of 76.68 % also shows the close agreement between the measured and simulated runoff value during the validation period. The low RMSE value for selected ANN model during calibration and validation also shows the better prediction of peak runoff value. In this study, the results obtained show clearly that the artificial neural networks are capable of model rainfall-runoff relationship in the small semi-arid catchments in which the rainfall and runoff are very irregular, thus, confirming the general enhancement achieved by using neural networks in many other hydrological fields. The ANN approach could provide a very useful and accurate tool to solve problems in water resources studies and management.

REFERENCES

1. R.J. Abrahart and L.See, “Comparing neural network and autoregressive moving average techniques for the provision of continuous river flow forecasts in two contrasting catchments”. Hydrological Processes, vol.14, (2000), pp.2157–2172.

2. ASCE Artificial neural networks in hydrology–I: preliminary concepts. Journal of Hydrologic Engineering, ASCE task committee on application of ANNS in hydrology, vol.15, no.2, (2000a), pp.115–123.

3. ASCE Artificial neural networks in hydrology–II: hydrologic applications. Journal of Hydrologic Engineering, vol.5, no.2, (2000b), pp. 124–137.

4. C.W. Dawson and R.L. Will, “An artificial neural network approach to rainfall-runoff modelling”, Journal des sciences hydrologiques, vol.43, no.2, (1998), pp.47-65.

5. M.T. Hagan and M.B. Menhaj, “Training feedforward techniques with the Marquardt algorithm”, IEEE Transactions on Neural Networks, vol.5, no.6, (1994), pp. 989-993.

6. A.H. Halff and H.M. Halff Azmoodeh, “Predicting Runoff from Rainfall using Neural Networks”, In Engineering Hydrology, Kuo CY (ed.). Proceedings of the Symposium sponsored by the Hydraulics Division of ASCE, San Francisco, CA, July 25–30, ASCE, New York; (1993),pp.760–765.

7. K.L. Hsu, H.V. Gupta and S. Sorooshian, “Artificial Neural Network Modeling of the Rainfall runoff Process”, Water Resources Research, vol.31, no.10, (1995), pp.2517–2530.

8. D.I. Jeong and Oh.Y. Kim, “Rainfall-Runoff Models Using Artificial Neural Networks for Ensemble Stream Flow Prediction”. Hydrological Process .vol.19, (2005), pp.3819-3835.

9. M.A. Kaltech, “Rainfall-Runoff Modeling Using Artificial Neural Network modelling and understanding”, Caspian Journal of Environmental Sciences, vol.6, (2008), pp.153-158.

10. Y. Najjar and H. Ali, “On the Use of BPNN in Liquefaction Potential Assessment Tasks”. In Artificial Intelligence and Mathematical Methods in Pavement and Geotechnical Systems, (Edited by Attoh-Okine), (1998), pp. 55-63.

11 Y. Najjar, X. Zhang, “Characterizing the 3D Stress-strain Behaviour of Sandy Soils”: A Neuro-mechanistic Approach. In ASCE Geotechnical Special Publication Number 96, (Edited by G. Filz and D. Griffiths), (2000), pp.43-57.

12. S.J. Riad, L. Mania, Y. Bouchaou, and Najjar, "Predicting Catchment Flow in Semi-arid Region via Artificial Neural Network Technique", Hydrological Process, vol.18, ., (2004), pp. 2387-2393.
13. A.Y. Shamseldin, "Application of Neural Network Technique to Rainfall-runoff modelling". Journal of Hydrology, vol.199, (1997), pp. 272-294.
14. J. Smith, R.N. Eli, Neural-network Models of Rainfall runoff Process. Journal of Water Resources Planning and Management, vol.121, no.6, (1995), pp. 499–508.
15. A.Z. Tokar and P.A. Johnson, Rainfall-runoff modelling using the artificial neural network. Journal of Hydrologic Engineering, ASCE, vol.4, no.3, (1999), pp.232–239.