Available online at www.elixirpublishers.com (Elixir International Journal)

Agriculture

Elixir Agriculture 87 (2015) 35609-35612

ISSN: 2229-712X

Non-linear modeling of rainfall runoff in Bearma Sub-Basin, Bundelkhand Using ANN

Parmanad[1,*], T.Thomus[2], R.M Singh[1] and Mangal Yadav[1]

[1]Department of Farm Engineering, Institute of Agricultural Sciences, Banaras Hindu University, Varanasi-221 005, Uttar Pradesh India.

[2]National Institute of Hydrology, Regional Centre, Walmi Campus, Bhopal-462 00l, Madhya Pradesh, India.

ARTICLE INFO

Article history:
Received: 14 August 2015;
Received in revised form:
27 September 2015;
Accepted: 05 October 2015;

Keywords
Artificial neural network (ANN),
Non-linear model,
Rainfall runoff relationship etc.

ABSTRACT

Water is one of the important natural resource available to mankind. Proper utilization of this resource requires assessment and management of the quantity and quality both spatially and temporally. A mathematical model provides quantitative mathematical description of the processes which includes a collection of mathematical equations expressing relationships between input and output variables through establishing and estimating the relevant parameters. The ANN models have been used successfully to model the complex non-linear input-output relationship. An ANN can be defined as data processing system consisting of a large number of samples. Artificial neural networks (ANN) have found increasing applications in various aspects of hydrology. The study revealed that a feed-forward artificial neural network with back propagation algorithm having a single hidden layer with two neurons in the hidden layer was able to model the rainfall-runoff transformation quite accurately. The correlation coefficient during the training varies between 0.88 and 0.93 and during testing varies between 0.78 and 0.95 respectively whereas the model efficiency varies between 73.70% and 85.77% with an overall efficiency of 81.18% during training and between 52.62 % and 90.01 % with an overall efficiency of 66.71% during testing.

Introduction

Water has a unique role in development planning processes due to its benefits and problems created by its excesses, shortages and quality deterioration. Rainfall runoff models are used to study the basin response to catchment rainfall and can be used to establish non-linear relationship between the rainfall and runoff. The rainfall-runoff can be performed on annual, monthly or daily time scale. For flood studies the event based rainfall-runoff modelling can be carried out and relationships established between rainfall and runoff for specific event for arriving at the flood regime during high intensity rainfall events (Hsu, K. L., Gupta, V.H., and Sorooshian, S.,1995). The most important step in application of a conceptual model to a catchment is model calibration. The calibration process requires a procedure of comparison of the simulated and observed runoff and adjustment of the parameters to minimise the errors between the observed and simulated runoff. The ANN models have been used in a non-linear mathematical structure, which is capable of representing arbitrarily complex non-linear processes that relate the inputs and outputs of any system (Sajikumara,, B.S. Thandaveswarab., 1999). The present study examines to develop a non-linear model using ANN for the daily runoff training of the neural network and selection of appropriate network architecture and testing the developed model with independent data set for the Bearma sub-basin in Bundelkhand region of Madhya Pradesh.

Materials and Methods

Architecture of the network

A neural network can be viewed as a weighted direct graph in which the artificial neurons are nodes and directed weighted edges represent connections between neurons. The architecture in which the local group of neurons can be connected may be either of (a) feed forward architecture – in which the network has no loops, (b) feedback architecture – in which loops occur in the network because of feedback connections (Anderson, J.A., 1995). Different network architectures yield different behavioural patterns of varying complexity.

General Description of Bearma Sub-basin up to Gaisabad

The river Bearma is one of the important tributaries of river Ken. This sub-basin is located between north latitudes of 230 07′ and 240 18′ and east longitudes of 780 54′ and 800 00. The Bearma sub-basin is bounded by Sonar basin on the west, the Ken sub-basin on the east and by the Vindhyan ranges on the south and lies completely in Madhya Pradesh. Parts of Narsinghpur, Jabalpur, Sagar, Damoh and Panna districts are drained by river Bearma. Some of the important tributaries are Lamti, Sun, Bamner, Guraiya, Godhar and Mala. The total catchment area of the basin is 5890 km2 and it is a leaf shaped elongated basin. The G/D site is located at Gaisabad and the catchment area up to Gaisabad is about 5807 km2.

Processing and analysis of data

The discharge data at the gauging site were processed to check for errors and inconsistencies. A rainfall runoff relation was developed for all the years under consideration. The annual runoff coefficient for the basin varied from 0.28 to 0.50. The graph showing the comparison of the annual rainfall and annual runoff is given in Fig.1. Similarly, the comparison of monthly rainfall and runoff in Bearma sub- basin is given in Fig. 2.

Model development

The basic structure of a network usually consists of three layers: the input layer, where the data are introduced to the network, the hidden layer or layers, where data are processed; and the output layer, where the results of given input are produced. When applying neural networks to modelling, a

Tele:
E-mail addresses: mangalchamps@gmail.com

number of decisions must be made. It is imperative to choose an appropriate neural network structure in terms of input vector and output vector, apart from the hidden neurons (Avinash Agrawal and R.D Singh.,2004). Determination of appropriate network architecture is one of the most difficult tasks in the model building process. The steps involved in the identification of a dynamic model of a system are:

(a) Selection of the input and output data suitable for calibration and validation

(b) Selection of a model structure and estimates of its parameters, and

(c) Validation of the identified model.

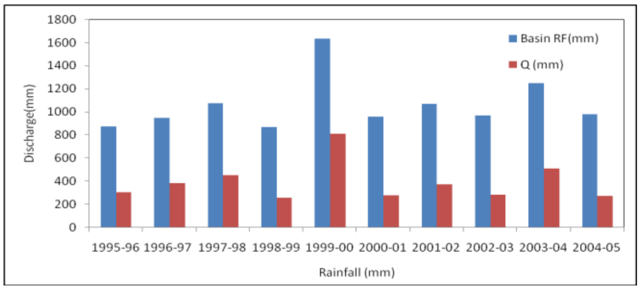

Fig 1. Annual rainfall and runoff in Bearma sub-basin

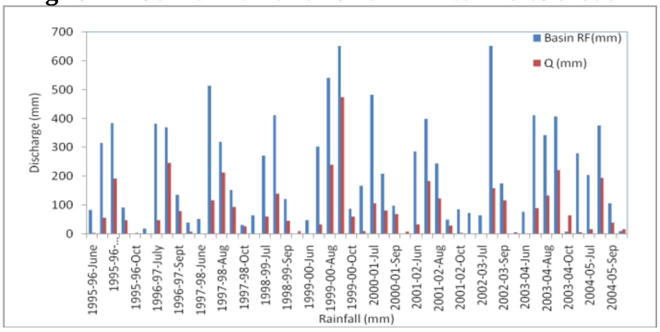

Fig 2. Monthly rainfall and runoff in Bearma sub-basin

Standardization of time series

A time series may often contain periodic components that tend to repeat over a period of time intervals, due to astronomic cycles. The behaviour of time series is known as a periodicity, which means that the statistical characteristics change periodically within the year. The periodic component can be removed from the time series as given by,

$$z_{t,\tau} = \frac{x_{t,\tau} - \bar{x}_\tau}{\sigma_\tau} \qquad \text{.........................(i)}$$

where,

zt,□ = standardized time series

and are the mean and standard deviation of the th day ;

□ = time interval within the year

Autocorrelation function

The autocorrelation function expresses the degree of dependency among neighbouring observations. It is a process of self-comparison expressing the linear correlation between an equally spaced series and the same series at a specified lag. Let z0 , z1, z2, ... zN-1 be a realization of a stationary stochastic process, then the population autocorrelation function can be defined as the quotient of the population auto covariance, cov(zt, zt+h) and variance, var(zt):

$$\rho(h) = \frac{Cov(z_t, z_{t+h})}{Var(z_t)} \qquad \text{............................ (ii)}$$

where,

□(h) = auto correlation function

zt = the value of the variable at tth time

h = time lag

Since the series analyzed is just one particular realization out of an infinite number of realizations of a stochastic process produced by the underlying probabilistic mechanism, the population autocorrelation function can be estimated using the simple autocorrelation function r(h):

$$r(h) = \frac{\sum_{t=1}^{N-h}(z_{t+h} - \bar{z})(z_t - \bar{z})}{\sum_{t=1}^{N-h}(z_t - \bar{z})^2} \; -1 \le r(h) \le 1 \qquad \text{.................... (iii)}$$

where,

$\bar{z}$ = sample mean

The 95% confidence band for the sample autocorrelation function given by Anderson and Jenkins, 1970:

$$r(h) = 0 \pm \frac{1.96}{\sqrt{n}} \ge \left[1 + 2\sum_{j=1}^{q} r_j^2\right]^{\frac{1}{2}} \; h>q \qquad \text{................(iv)}$$

where,

q = order of the process

n = number of observation in the series.

Model performance indicators

A mean squared error (MSE) is one of the most commonly used performance measure in hydrological modeling. Many researchers used MSE or its root (RMSE) as an accuracy measure (Carpenter and Brathelemy 1994; Bastarache et al., 1997; Shamseldin 1997). The MSE and RMSE are given by,

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(Q_o(i) - Q_c(i)\right)^2 \qquad \text{................. (v)}$$

$$RMSE = (MSE)^{0.5} \qquad \text{.................. (vi)}$$

Results and Discussion

Identification of the network architecture

It can be observed that the RMSE during training and testing gets reduced as the number of neurons in the hidden layer increases from one to two. Thereafter, with the increase in the number of neurons in the hidden layer, the RMSE gets reduced during training whereas it increases during testing. This indicates the model is able to reproduce the results well on the independent test data as long as the number of neurons in the hidden layer is limited to two. Similarly the efficiency of the model is maximum during training and testing with two neurons in the hidden layer is 81.18% and 66.71% respectively. Even though the efficiency of the model increases significantly during the training with the increase in the number of neurons in the hidden layer, but during testing it drops down to significantly lower values. The efficiency during training and testing with 2 neurons in the hidden layer is indicative of the fact. So the network with 4 neurons in the input layer, two neurons in the hidden layer and one neuron in the output layer with log sigmoid transfer function in both hidden layer and output layer has been finalised for modelling the rainfall-runoff process in Bearma sub-basin. The effect of the number of neurons in the hidden layer is presented in Table 1.

Performance of the model

In the present study, based on the performance in representing the rainfall-runoff process, the feed-forward neural network with two neurons in the single hidden layer was finalized. The comparison of the observed and simulated runoff during the training period from 1998 to 2000 is given in Fig.3 to Fig.5. It is observed that the model is able to simulate the flows reasonably well except for two peak flow in 1998 and one peak flow each during 1999 and 2000. Overall the model was able to simulate the flows with a fair degree of accuracy.

Table 1. Goodness of fit statistics for the effect of number of neurons in hidden layer

Year	Efficiency (%)	RMSE	Correlation coefficient	Year	Efficiency (%)	RMSE	Correlation coefficient
Training				***Testing***			
1995	79.44	100.75	0.89	2001	60.72	167.33	0.78
1996	77.36	198.40	0.88	2002	52.62	197.92	0.93
1997	85.77	132.27	0.93	2003	65.96	196.25	0.82
1998	73.70	123.72	0.89	2004	90.01	78.38	0.95
1999	82.95	267.93	0.91				
2000	84.76	73.03	0.92				

Table 2. Efficiency of the trained and tested model during 1995-96 to 2004-05

No. of neurons in the hidden layer	RMSE		Efficiency (%)	
	Training	Testing	Training	Testing
1	191.08	191.72	75.58	56.23
2	163.26	167.21	81.18	66.71
3	130.83	202.74	88.53	51.06
4	119.25	194.57	90.49	54.92
5	95.16	206.92	93.94	49.02

The model was tested on the independent test data from 2001 to 2003. The comparison of the observed and simulated runoff during the testing period is given in Fig.6 to Fig. 8. The model trained with the input vector is able to simulate the flows with the independent test data with reasonable accuracy. The model tends to have smaller value of RMSE during training. The value of RMSE is found slightly deteriorating during validation. During training as well as testing the model matches the peak flow to a reasonable degree of accuracy. The efficiency of the model is measure of the performance of the model in predicting the output values. According to this statistic, the model predictions were fairly good during training. The model efficiency during training and validation is 81.18% and 66.71% respectively. The efficiency of the model, root mean squared error, and correlation coefficient during the various years of training and testing is given in Table 2. It can be seen that the efficiency of the model is rather very poor during the training period of year 1998 as it is not able to simulate the peak flow on July 9, 1998 accurately (observed peak : 1290.62 cumecs; simulated peak: 348.99 cumecs). However the other results during training is satisfactory.

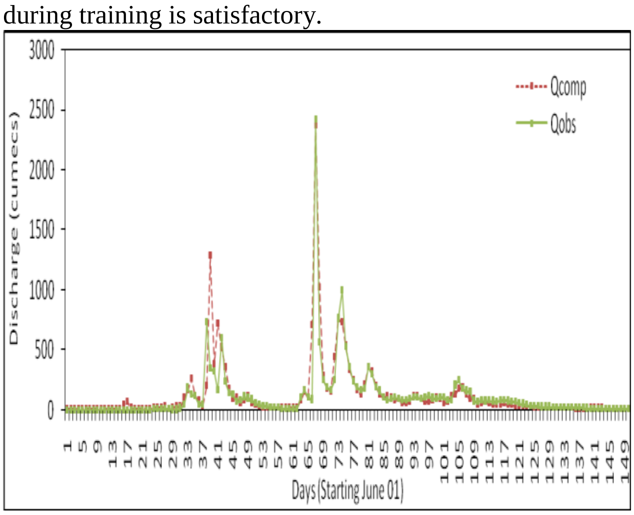

Fig 3. Comparison of the observed and computed flows during calibration (1998)

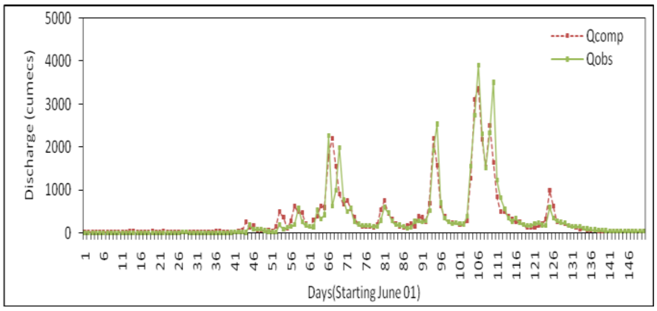

Fig 4. Comparison of the observed and computed flows during calibration (1999)

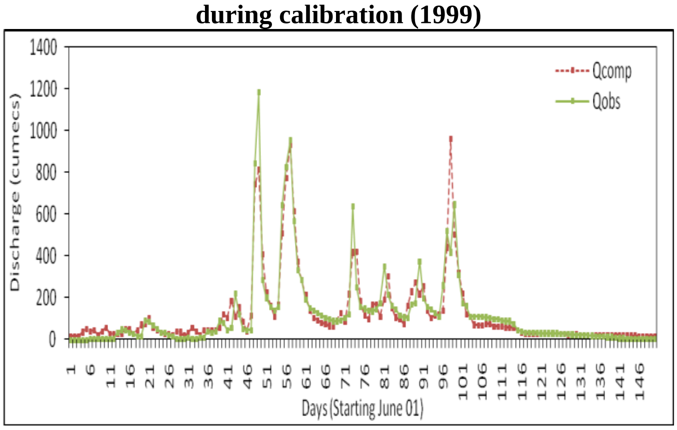

Fig 5. Comparison of the observed and computed flows during calibration (2000)

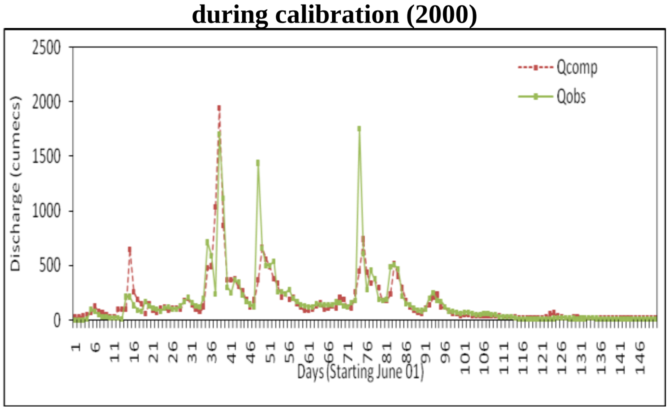

Fig. 6. Comparison of the observed and computed flows during validation (2001)

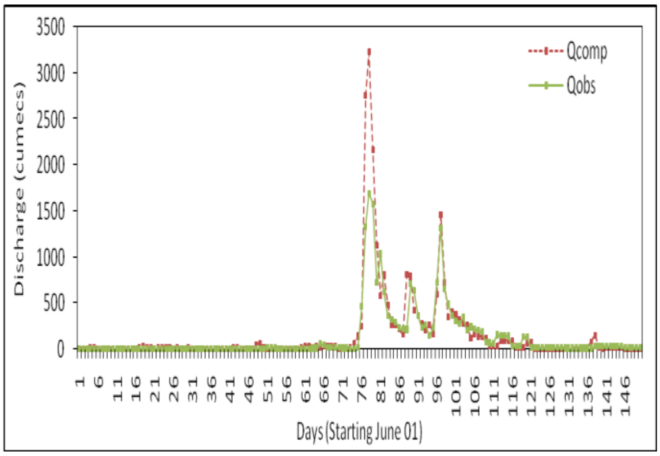

Fig 7. Comparison of the observed and computed flows during validation (2002)

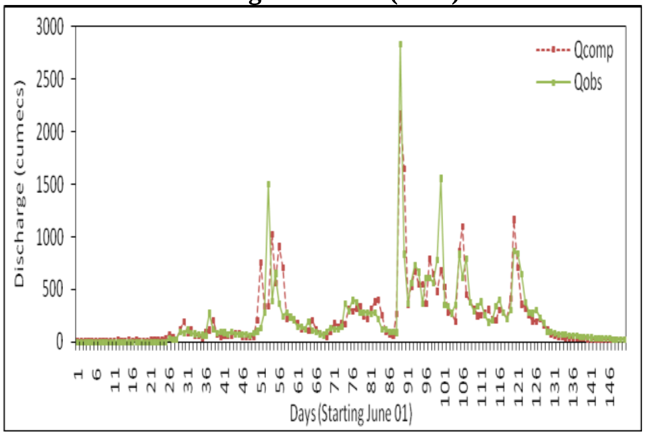

Fig 8. Comparison of the observed and computed flows during validation (2003)

Conclusion

An attempt has been made to apply the artificial neural network techniques to develop a rainfall-runoff model for the Bearma sub-basin in the Bundelkhand region falling in Madhya Pradesh. Earlier application of artificial neural networks in water resources revealed that the approach of neural computations was very effective in developing the required model, due to its various advantages. Therefore it was decided to apply the ANN to model the precipitation-runoff relationship in the Bearma sub-basin. The architecture of the feed-forward back-propagation network was determined based on a trial and error procedure and after examining the goodness of fit statistics. An auto correlation and partial auto correlation analysis of the standardized daily flow series suggested that the flow at time'ť' was highly correlated to previous one days flow. These parameters were included in the input vector of the network, apart from prior 2-days rainfall series, at which the flow was to be predicted. The number of rainfall patterns in the input vector was finalized by cross-correlation analysis and subsequently by trial and error.

The study demonstrates that ANN can model accurately the non-linear relationship between rainfall and runoff and provides a systematic approach and shortened time spent on training of models compared to development and calibration of the conceptual models. Hence it is concluded that the ANN model demonstrates a high potential for application of neural networks to various precipitation-runoff modelling scenarios. Feed-forward back propagation network model developed for rainfall-runoff process in the Bearma sub-basin might be employed for water resources planning and management in the basin.

References

Anderson, J.A., (1995). An introduction to neural networks, MIT Press, Cambridge, Mass.pp.47.

Avinash Agrawal and R.D Singh., (2004). Runoff modelling through back propagation artificial neural network with variable rainfall-runoff data.Water Resources Management 18:285-300.

Carpenter, W.C., and Barthelemy, J.F., (1994). Common misconceptions about neural networks as approximators. J. Comp. in Civ. Engg., ASCE, 8 (3), pp. 345-358.

Hsu, K. L., Gupta, V.H., and Sorooshian, S., (1995). Artificial neural network modelling of the rainfall-runoff process. Water Resour. Res., 31 (10), pp. 2517-2530.

Sajikumara,, B.S. Thandaveswarab., (1999). A non-linear rainfall–runoff model using an artificial neural network. Journal of Hydrology, 216 , pp. 32–55.