I.J.E.M.S., VOL.4 (2) 2013: 108-112 ISSN 2229-600X

INTERNATIONAL JOURNAL OF ENGINEERING AND MANAGEMENT SCIENCES

www.scienceandnature.org

ARTIFICIAL NEURAL NETWORKS BASED DAILY RAINFALL-RUNOFF MODEL FOR AN AGRICULTURAL HILLY WATERSHED

[1]Singh, P. V., [1]Kumar Akhilesh, [2*]Rawat J. S., & [1]Kumar Devendra
[1]Department of Soil & Water Conservation Engineering, College of Technology, GBPUAT, Pantngar-263145
[2]*Director, NRDMS, Kumaon University, SSJ Campus, Almora-263601

ABSTRACT
Runoff prediction through rainfall-runoff modeling is required for many engineering applications including efficient water management planning. Conceptually, the rainfall-runoff transformation process is non-linear phenomenon and highly complex in nature as it involves large number of variables pertaining to mainly rainfall, soil system, morphology and vegetation. Various approaches have been adopted to represent rainfall-runoff process. However, recently the concepts of Artificial Neural Networks (ANN) have been applied successfully for modeling non-linear relationships in different fields of Engineering and Science. In this study, ANN based daily rainfall-runoff model has been established for an agricultural hilly watershed known as Khunt micro-watershed located in the district of Almora, Uttarakhand, India. In the development of the model, daily rainfall and runoff data for the period 1st June to 31th October for years 2005-2009 were used to train the ANN, and for the years 2010 and 2011 were used for model validation purpose. The performance of the developed model was assessed based on parameters like root mean square error (RMSE) and correlation coefficient (R). A network structure resulting in highest value of correlation coefficient and simultaneously in the lowest value of RMSE was designated as the best performing. Based on these considerations, it was observed that the performance of the model based on one day lag and two days lag time were found satisfactory for the study area. However, the model based on the network 5-5-1 structure with 2 days lag was found to has an edge over the model based on 3-3-1 model structure with one day lag.

KEYWORDS: Rainfall, runoff modelling, artificial neural network, Almora, India

INTRODUCTION

Water is essential for integrity and sustainability of the earth's system (UN-world water development report, 2003). The availability and distribution of water is highly erratic both in space and time in India and in most of the part of world. The water requirement for ever-increasing population of the world even aggregated the gap between the availability and the demand of the water. The problem to some extent can be mitigated by adopting a very scientific way of water management. The authentic information on availability of amount of runoff at different times from a watershed is the key for any water management planning process.

The transformation of rainfall into runoff is a very complex process due to highly erratic occurrence and distribution pattern of rainfall as well as heterogeneous watershed system. This makes quite difficult to ascertain the quantitative and qualitative response of the catchment in response to a certain rainfall input or a series of inputs. Basically, the hydrological processes are non-linear time and space dependent in nature. The linear representation of these processes is over simplification and may not provide a true representation. Further, hydrological models may be lumped or black box models and distributed models. Distributed models being more complex in nature also need exhaustive data base for their development. The required data information may not always be available, particularly in remote hilly areas, and the hydrologic models (rainfall-runoff models) are attempted considering watershed as a lumped system where no consideration is given for space distribution and output from the watershed system is simulated in terms of total input.

ANN models which are basically lumped model (black box model) have been used successfully to model complex non-linear input-output relationships. Many studies (Minns and Hall, 1996; Shamseldin, 1997; Dawson and Wilby, 1998; Campolo et al. 1999; Zhang et al.2000;) have demonstrated that the ANNs are excellent tools to model the complex rainfall–runoff process and can perform better than the conventional modeling techniques. Therefore, in view of the importance of the relationship between rainfall and runoff, the present study was undertaken with the main objective to develop rainfall-runoff models using feed-forward multilayer neural network - back propagation technique for prediction of daily runoff for an agricultural hilly watershed.

Study Area

The present study was carried out in Khunt micro-watershed in Almora district of Uttarakhnad, India. The Khunt watershed is located in the Lesser Himalayan terrain of Kumaun Division of Uttarakhand in India encompassing an area of 137 ha lies between 29o36'7.45"N to 29o36'57.23"N latitudes and 79o35'34.84"E to 79o36'36.1"E longitudes. The location map of Khunt watershed is shown in Fig. 1. The catchment is semicircular in shape which circularity ration stands at 0.63 with length width (L/W) ratio as 2.7. The climate of watershed is cool temperate with annual maximum, minimum and average temperature of the watershed stand the average annual rainfall stands at 913.4mm (Rawat, 2012). The elevation in the watershed varies in between 1220 m and 1480 m above mean sea level. About 75% area of the watershed is used for cultivation of rain fed crops and remaining 25% is under barren condition. The average slope of the watershed stand at 19.5 degree which varies between 1 on terraced fields to 70 degree on steep scarps. The daily rainfall and runoff data of the study area was provided by the Centre of Excellence for Natural Resources Data Management System, Department of Geography, Kumaun University SSJ Campus Almora.

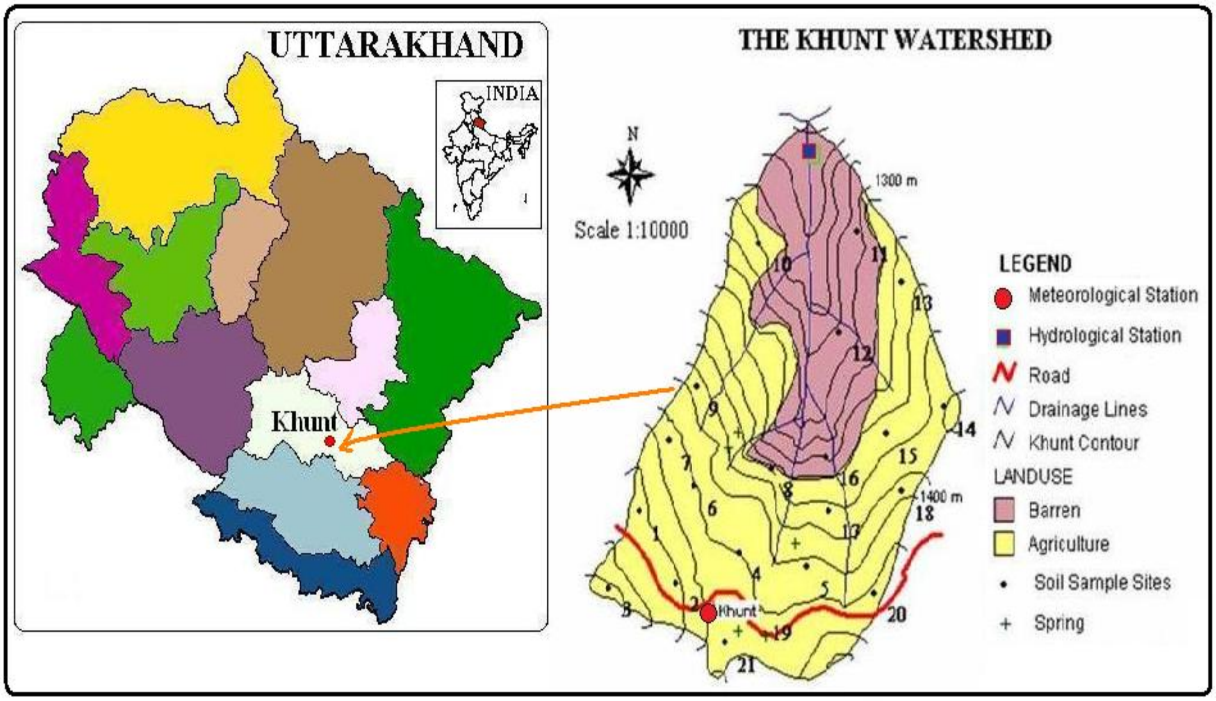


Fig.1: Location map of the study area

ARTIFICIAL NEURAL NETWORK

An ANN is an information-processing system composed of many nonlinear and densely interconnected processing elements or neurons, which are analogous to the biological neurons in the human brain. The main function of the ANN paradigms is to map a set of inputs to a set of outputs. The most important attribute of a multilayer feed-forward network is that it can learn a mapping of any complexity (Zurada. 1992).

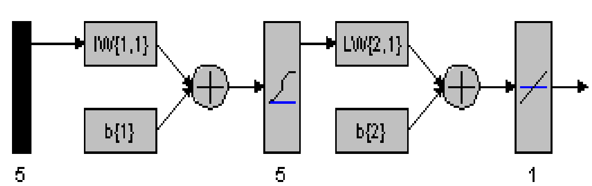
Model Development

The watershed system being dynamic in nature is memory based system. The runoff from a watershed system on any day not only depends on the rainfall on that day but also on previous rainfall and runoff. Many researchers such as Minns et al. (1996) and Campolo et al. (1999) that the rainfall information alone is not sufficient to compute the runoff. Chua and Wong (2011) also compared the performance of different ANN developed model based on rainfall alone and rainfall and runoff together as input for an asphalt plane and found that taking both rainfall and runoff as input provide the better result. The influence and extent of previous data (lag) would depend upon the persistence of memory component of the watershed system. Generally large watersheds have strong memory component and consider large value of input data lag. In this study an effort is made to develop ANN based daily rainfall-runoff model for prediction of daily runoff using daily rainfall (P) and daily runoff (Q) data of duration Ist June to 30th October during the years 2005-2009 for the study area. It has been reported that to develop ANN based models, in general, a three days lag is sufficient. In present case, the study area being small in size only maximum two days lag have been considered to develop rainfall-runoff model as shown in the Table1.

Table 1. Model description and model detail

Model No.	Model description	No. of neurons in the hidden layer
1.	$Qt= f(P_t, P_{t-1}, Q_{t-1})$	2,3
2	$Qt= f(P_t, P_{t-1}, Q_{t-1}, P_{t-2}, Q_{t-2})$	2,3,4,5,

A log sigmoid transfer function which varies between zero and one was used in the hidden layer and a pure linear transfer function was used at the output layer. A neural network structure for 2 days lag considering 5 neurons in hidden layers and one neuron in output layer is shown below,



Size of Hidden Layers

The number of neurons in the hidden layer is mainly responsible for capturing the dynamic and complex relationship among various input and output variables considered in developing an ANN. The most popular and effective strategy for selecting the appropriate number and size(s) of the hidden layer(s) is trial-and-error procedure. Using too few neurons results in under fitting and too many results in several problems such as over fitting. A thumb rule i.e. the number of hidden neurons should be in the range between the size of the input layer and the size of the output layer is used in the study for determining the number of neurons in the hidden layer. A number of networks with one hidden layer were trained with different combination of hidden neurons and a network is selected which results in the minimum root mean square error (RMSE) and maximum correlation coefficient (R).

Normalization of data

In the present study log sigmoid function is used as transfer function in the hidden layer and a pure linear function is used in the output layer of an ANN. The log sigmoid function ranges between zero and one. Therefore, the rainfall and runoff data are normalized using the following relationship.

$$Z = \frac{Z_i - Z_{\min}}{Z_{\max} - Z_{\min}} \quad \text{... (1)}$$

where Z is normalized variable, Zi is the real value of the variable applied at node i, Zmax is the maximum value of the variable; Zmin is the minimum value of the variable.

Model Performance

The performance of the developed model was evaluated by determining the statistical indices like correlation coefficient (R) and root mean square error (RMSE) which are mathematically expressed as, .

Correlation coefficient (R)

The correlation coefficient (R) is an indicator of degree of closeness between observed and predicted values. The correlation coefficient is determined by the following relationship.

$$R = \frac{\sum_{i=1}^{n}\left\{\left(Q_{oi} - \bar{Q}_o\right)\left(Q_{pi} - \bar{Q}_p\right)\right\}}{\sum_{i=1}^{n}\left(Q_{oi} - \bar{Q}_o\right)^2 \sum_{i=1}^{n}\left(Q_{pi} - \bar{Q}_p\right)^2} \quad \text{... (2)}$$

where, $\bar{Q}_o$ is the arithmetic mean of observed values of runoff , $\bar{Q}_p$ is the arithmetic mean of predicted values of runoff, Qoi is the ith observed value of the runoff, Qpi is the ith predicted value of the runoff and n is the number of observations.

Root mean square error (RMSE)

The root mean square error (RMSE) is used as a measure to assess the prediction accuracy of the developed model. It always produces positive values by squaring the errors. The RMSE is zero for perfect fit and increased values indicate higher deviations between predicted and observed values. The root mean square error between observed and predicted values is determined by the following relationship.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(Q_{oi} - Q_{pi}\right)^2}{n}} \quad \text{...(3)}$$

RESULTS AND DISCUSSION

The daily data of rainfall and runoff of the period 1st June to 30th October during years 2005-2009 were used for the training of the model, whereas the daily data of years 2010 and 2011 were used for verification of the model. Based on the minimum value of the root mean square error (RMSE) and maximum value of correlation coefficient (R), the best model structure has been selected (Table 2). From Table 2, it is clear that model considering 2 days lag is better in comparison to 1 day lag which is also evident by Fig 2 and 3. Scatter plots (Fig 3 and 4) during both training and validation indicate that only for extreme values the model over predicts

Table 2. Calculated values of statistical indices for the best possible structures

Model No.	Input variables	Best network structure	Epoch	Correlation coefficient (R)		Root Mean Square Error (RMSE)	
				Training	Validation	Training	Validation
1	P_t, P_{t-1}, Q_{t-1}	3-3-1	17	0.8925	0.9060	0.0383	0.0405
2	P_t, P_{t-1}, Q_{t-1}, P_{t-2}, Q_{t-2}	5-5-1	13	0.9348	0.9589	0.0302	0.0287

CONCLUSIONS

The study reveals that a feed-forward artificial neural network with back-propagation algorithm having a network 5-5-1 with two days lag and network 3-3-1 with one day lag are able to model the rainfall-runoff transformation quite accurately. The qualitative evaluation of the model performance on the basis of graphical comparison revealed a close agreement between observed and predicted values of runoff. Under quantitative evaluation, the higher values of correlation coefficient (R) and lower value of root mean square error (RMSE) for network also confirm the model's ability to predict daily runoff with reasonable accuracy in the study area. The networks with two days lag with network structure 5-5-1 performed slightly better than the networks with one day lag having network structure as 3-3-1 and the former networks requires a less amount of computational efforts for training as well as testing.

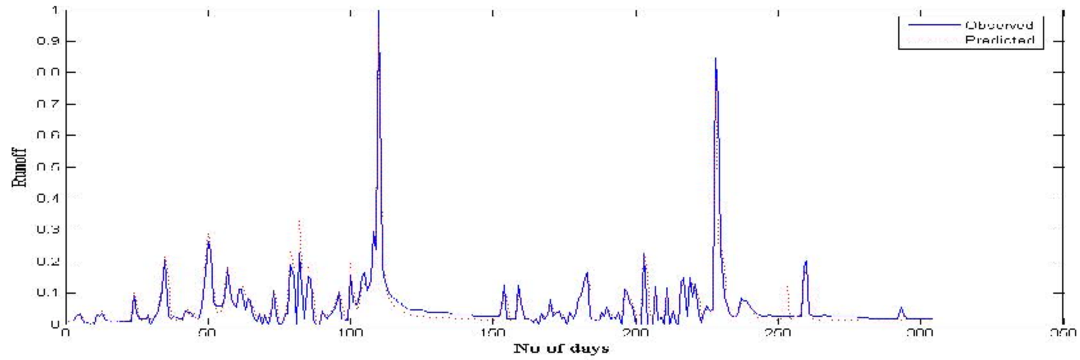


Fig 2. Observed and predicted daily runoff (dimensionless) with model structure 3-3-1 for 1 day lag during validation

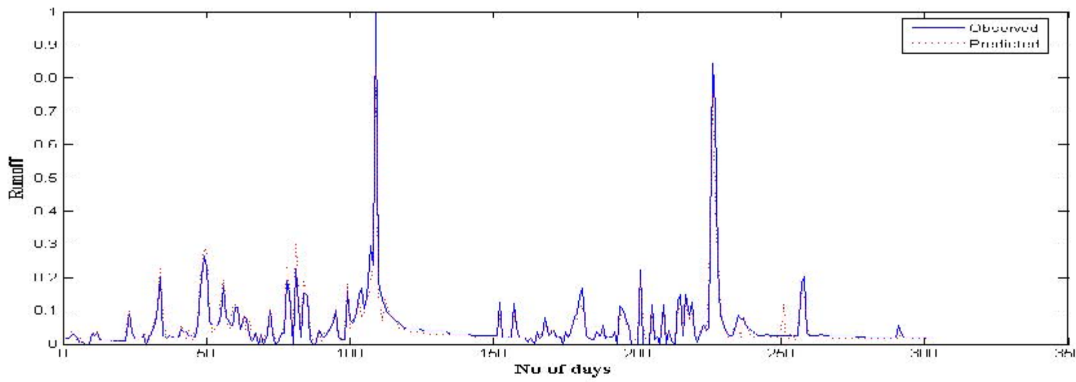


Fig 3. Observed and predicted daily runoff (dimensionless) with model structure 5-5-1 for 2 day lag during validation

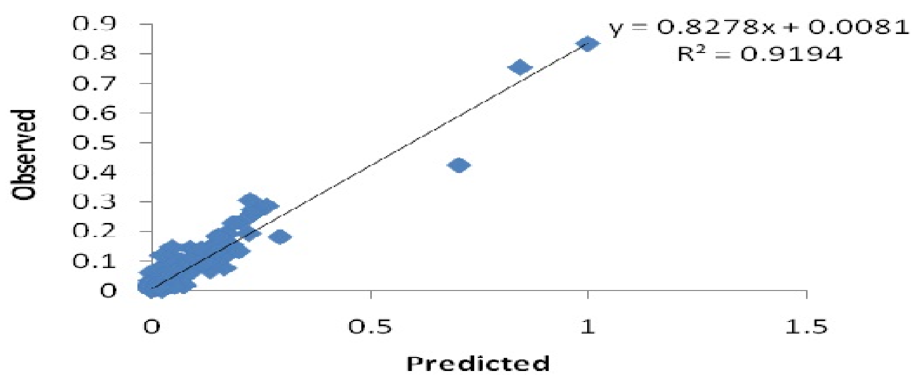


Fig-4. Validation set's scatter plot with model structure 5-5-1 for 2 day lag

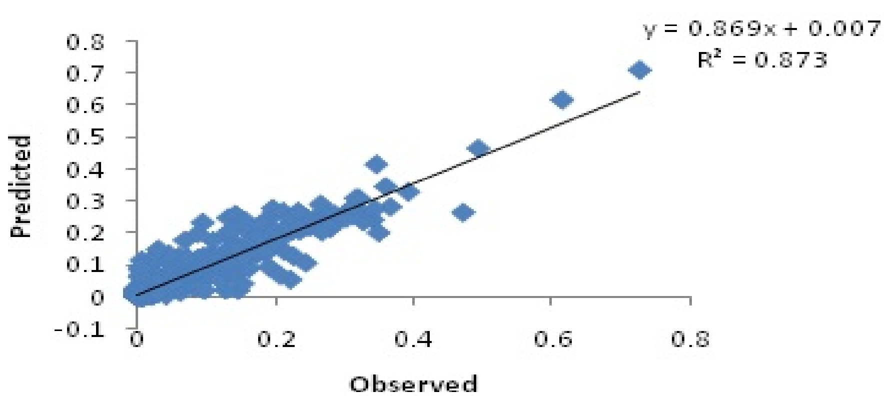


Fig-5. Training set's scatter plot with model structure 5-5-1 for 2 day lag

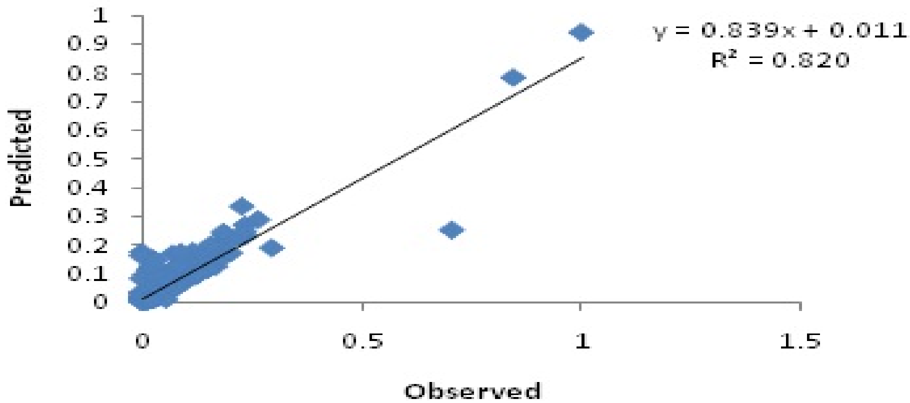


Fig- 6. Validation set's scatter plot with model structure 3-3-1 for 1 day lag

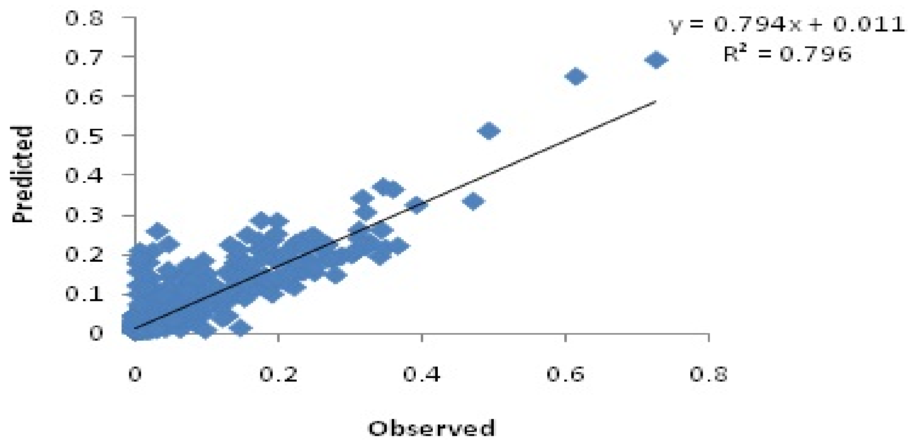


Fig-7. Training set's scatter plot with model structure 3-3-1 for 1 day lag

REFERENCES

Campolo, M. and P. Andreussi (1999) A. Soldati River flood forecasting with neural network model, Water Resources Research 35 (4), 1191–1197.

Dawson, C. W. and R Wilby (1998) An artificial neural network approach to rainfall runoff modeling. Hydrological Science Journal, 43(1), 47-66.

Lloyd H.C. Chua and Tommy S.W. Wong (2011) Runoff forecasting for an asphalt plane by Artificial Neural Networks and comparisons with kinematic wave and autoregressive moving average models. Journal of Hydrology. 397, 191–20.

Minns, A, W. and M. J. Hall (1996) Artificial neural networks as rainfall-runoff models. Hydrological Science Journal, 41(3), 399-417.

Rawat,J.S. (2012) Weather Informatrics of Khunt, Centre of Excellence for Natural Resources Data Management System in Uttarakhand, Department of Geography, Kumaun University, SSJ Campus Almora-263601, pp.1-.32.

Shamseldin, A. Y. (1997) Application of a neural network technique to rainfall-runoff modeling. Journal of Hydrology, 199, 272-294.

Zurada, M. Jacek (1992) Introduction to artificial neural systems. West publishing company.

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